Using Deep Learning to Detect the Presence and Location of Hemoperitoneum on the Focused Assessment with Sonography in Trauma (FAST) Examination in Adults

Megan M. Leo[1,2] · **Ilkay Yildiz Potter**[3] · **Mohsen Zahiri**[3] · **Ashkan Vaziri**[3] · **Christine F. Jung**[4,5,6] · **James A. Feldman**[1,2]

Abstract

Abdominal ultrasonography has become an integral component of the evaluation of trauma patients. Internal hemorrhage can be rapidly diagnosed by finding free fluid with point-of-care ultrasound (POCUS) and expedite decisions to perform lifesaving interventions. However, the widespread clinical application of ultrasound is limited by the expertise required for image interpretation. This study aimed to develop a deep learning algorithm to identify the presence and location of hemoperitoneum on POCUS to assist novice clinicians in accurate interpretation of the Focused Assessment with Sonography in Trauma (FAST) exam. We analyzed right upper quadrant (RUQ) FAST exams obtained from 94 adult patients (44 confirmed hemoperitoneum) using the YoloV3 object detection algorithm. Exams were partitioned via fivefold stratified sampling for training, validation, and hold-out testing. We assessed each exam image-by-image using YoloV3 and determined hemoperitoneum presence for the exam using the detection with highest confidence score. We determined the detection threshold as the score that maximizes the geometric mean of sensitivity and specificity over the validation set. The algorithm had 95% sensitivity, 94% specificity, 95% accuracy, and 97% AUC over the test set, significantly outperforming three recent methods. The algorithm also exhibited strength in localization, while the detected box sizes varied with a 56% IOU averaged over positive cases. Image processing demonstrated only 57-ms latency, which is adequate for real-time use at the bedside. These results suggest that a deep learning algorithm can rapidly and accurately identify the presence and location of free fluid in the RUQ of the FAST exam in adult patients with hemoperitoneum.

Keywords Focused Assessment with Sonography in Trauma · Deep learning · Artificial intelligence · Emergency ultrasound · Point-of-care ultrasound · Trauma

✉ Megan M. Leo
megan.leo@bmc.org

1 Boston University School of Medicine, Boston, MA, USA

2 Department of Emergency Medicine, Boston Medical Center, BCD Building, 800 Harrison Ave1St Floor, Boston, MA 02118, USA

3 BioSensics, LLC, Newton, MA, USA

4 Division of Emergency Ultrasound, Department of Emergency Medicine, John H. Stroger Jr. Hospital of Cook County, Chicago, IL, USA

5 Department of Emergency Medicine, Chicago Medical School of Rosalind, Franklin University of Medical Sciences, Chicago, IL, USA

6 Department of Emergency Medicine, Rush Medical College, Chicago, IL, USA

Introduction

The use of point-of-care ultrasonography (POCUS) has gained wide acceptance in diverse practice settings including in the emergency department, prehospital, military, and austere settings because of the capability of this technology to rapidly and accurately and noninvasively rule-in serious injury by detecting free intra-abdominal, intra-thoracic, and intra-pericardial free fluid [1–10]. The Focused Assessment with Sonography in Trauma (FAST) exam has become standard of care for the evaluation of patients with trauma by decreasing time to operative management and has been accepted as a part of the American Trauma Life Support (ATLS) protocol in the USA [11, 12]. The FAST exam interpretation requires training and can have differences in interpreter accuracy, with sensitivity and specificity of the FAST exam ranging from 87 to 98%

and 99 to 100%, respectively in detecting intraperitoneal fluid in patients who suffer from blunt trauma [13]. The primary limitation to the generalized use of the FAST exam has been the lack of trained and experienced operators in all types of clinical settings, including those prehospital and in resource-limited settings. While expert readers can achieve high agreement, experience affects the diagnostic performance and the interpretation of the FAST exam. The human and financial resources needed to adequately train each clinician to achieve minimum standards for competency as recommended by the American College of Emergency Physicians and other national societies are substantial and they have been reported as a major barrier to more consistent use of POCUS in emergency and critical care settings [14, 15]. Other image factors such as the volume of free fluid and recognition of structures such as the stomach and vascular structures that could be misclassified as free fluid support the value of an artificial intelligence application to aide less experienced providers interpret the exam [16].

Teleultrasound refers to instruction provided virtually by a trainer and has been shown to be feasible and non-inferior to in-person instruction for image acquisition and interpretation [17, 18]. However, this still requires a trainer or expert to be available in real time to provide this feedback. We propose that the development of artificial intelligence to assist novice operators in the real-time interpretation of the FAST exam may be another way to bridge this gap. Many studies have examined the role of deep learning and artificial intelligence to aide in the interpretation of ultrasound applications including cardiac echocardiography, pregnancy evaluation, specific organ diagnoses, and many other applications [19–24]. Three recent studies have examined the use of deep learning in pediatric and adult FAST [25–27], attaining up to 97% accuracy in hemoperitoneum identification over right upper quadrant exams [26]. Despite high prediction performance, these approaches did not provide location of identified abdominal free fluid. To address this, Lin et al. [28] described the use of a deep learning algorithm to perform pixel-by-pixel segmentation of free fluid and diagnosed ascites for each exam based on the segmentation result, achieving sensitivity and specificity of 94% and 68% for Ascites-1, respectively.

We believe that our study addresses important gaps and limitations in the literature on the use of artificial intelligence to analyze the FAST exam. To begin with, recent studies on deep learning-based free fluid detection such as Cheng et al. [26] have included non-trauma causes of abdominal free fluid in their analysis of the "FAST" exam. As there can be differences in the appearance of traumatic hemoperitoneum such as complex images with clot and free blood, as well as the amount of free fluid and appearance of other abdominal organs in non-trauma causes of free fluid

(such as cirrhosis of the liver with ascites), we believe that it is important that a clinical application designed for the interpretation of the FAST exam in trauma including ours should be tested in the intended population.

Our study also advances an application that could potentially identify both the presence and location of free fluid in real time. Providing the location of free fluid identified by the application allows the clinician to visualize and overread the interpretation, similar to the functionality of an automated electrocardiogram interpretation provided on each report. Moreover, as FAST exam is used in emergency care settings, high accuracy in hemoperitoneum detection should be paired with the capability of rapid inference. Motivated by this, the purpose of this study was to test the performance of a deep learning algorithm that could be performed on a laptop computer (alternatively desktop computer, smart phone, or tablet) to *rapidly and accurately* identify the presence and location of hemoperitoneum in adult patients. Our approach is based on the YoloV3 algorithm [29], which exhibits fast and well-established prediction performance for object detection in various domains [30]. We assess each FAST exam image-by-image using YoloV3 and determine hemoperitoneum presence for the exam using the detection with highest confidence score. In doing so, free fluid is also localized on individual images by drawing a bounding box around the detection, rather than pixel-by-pixel segmentation as proposed by previous works. Our approach attained 95% sensitivity, 94% specificity, 95% accuracy, and 97% in hemoperitoneum detection, significantly outperforming three recent methods we implemented. At the same time, average time to process each ultrasound image was only 57 ms, equivalent to 18 ultrasound images processed per second, which confirmed our motivation of rapid detection.

Materials and Methods

Study Design and Setting

We performed a retrospective analysis of archived POCUS images obtained from 2010 to 2018 in adult patients 18 years or older who had FAST examinations performed by emergency physicians in the Emergency Department of an urban Level 1 Trauma Center with approximately 140,000 visits per year. We identified cases with confirmed hemoperitoneum as well as an equal number of negative controls. We cross referenced the exams to the Trauma Registry to link exams to specific patient medical records and reviewed the electronic medical record (Epic Systems software, Verona, WI) to confirm that hemoperitoneum was identified by additional imaging such as computerized tomography or operative report. To have an equal number of male and female

positive cases, we also included women who had confirmed hemoperitoneum from ruptured ectopic pregnancy or ovarian cyst at surgery. We abstracted demographic data, injury mechanism, confirmatory test results, ultrasound device manufacturer, and probe that was used for the FAST study on a password protected secure computer.

The right upper quadrant (RUQ) view of the abdomen is the most sensitive region of the FAST exam for free fluid [31]. Therefore, for this proof-of-concept study, we analyzed only the RUQ quadrants from clinically obtained FAST exams. We obtained negative FAST exams from the ultrasound image archives that were coded and confirmed as negative studies. Each ultrasound exam was exported as a video clip in.mp4 format. The study was deemed exempt from review by the institutional review board at the study site.

Two expert readers performed all of the image analysis and labeled all FAST studies. The first expert is an emergency point-of-care ultrasound fellowship-trained expert with over 10 years of experience and oversees the quality assurance and image review for the emergency department. The second expert reader who analyzed the images was among the inception cohort of one of the first emergency point-of-care ultrasound credentialing programs in the country; has met initial credentialing standards set by the American College of Emergency Physician (ACEP) guidelines, including continuous hospital privileging in the interpretation of the FAST exam; and has more than 20 years of experience and imaging quality assurance review [32].

Studies identified within the emergency ultrasonography archive were exported and deidentified. These were assigned a unique identifier prior to sharing images with the industry collaborator.

Data Annotation

The deidentified video clips (.mp4) were used for analysis using a deep learning algorithm. For videos that belong to positive cases, clinical experts in collaboration manually segmented all regions in each frame that indicate a clearly visible area of free fluid using PhotoPad Image Editor for Windows (NCH Software). Each image was coded by the two reviewers as a collaborative process to reach consensus on all areas on a unique image where free fluid was detected. Each image was coded independently of other images because free fluid may not be seen in each image of a FAST exam due to the operator scanning through the entire area of interest in the body. Although the physicians who coded the images were aware of the overall categorization of the study (positive or negative), each still frame had to be coded independently for the presence and location of

free fluid because this can vary by location of the ultrasound probe and the amount of free fluid.

To prepare these ground-truth labels for the free fluid detection algorithm, we mapped free fluid regions to bounding boxes by fitting a tight rectangular bounding box to fully contain each free fluid region. Each case contained on average 91 ($\pm$ 62) ground-truth free fluid boxes, reaching up to 269 boxes per case. Ground-truths also exhibited a wide range of sizes: A free fluid box occupies on average 2.7% ($\pm$ 2.6) of a video frame, with the smallest box occupying only 0.6% and the largest box occupying up to 14.8% of their corresponding video frames.

Data Preprocessing

The resulting dataset of 94 cases was partitioned into training, validation, and test cases via fivefold cross validation in a stratified manner, keeping a uniform ratio of positive and negative cases in each set. For each fold, 72% of the cases are used for training our algorithm, 8% of the cases were used for validation, and the remaining 20% of the cases were held out for testing. Data partitioning was based on cases rather than video frames, ensuring that a subject that is included in training is *not* included in validation or testing. We note that using 5 folds in cross validation are typical in machine learning and deep learning literature and has been shown to attain high prediction performance [33–36]. While it is possible to further increase the number of folds, each training fold requires training a new deep learning model, introducing a computational overhead.

As described in the "Study Design and Setting" section, each ultrasound exam was exported as a video clip in.mp4 format for analysis, deidentified, from a secure imaging archival system used by the hospital. We prepared each image from each video clip using the same data preprocessing steps. Images preprocessed from training, validation, and test cases are then used for training, validation, and testing, respectively. Each video frame contains data acquisition information (such as frequency) in the margins, in addition to the relevant image data. Thus, each video frame was first cropped by 5% from left and right sides to zoom into the region of interest, while ensuring that image data and free fluid regions are not discarded. The resulting frames were resized to obtain a uniform size of 256 $\times$ 256 pixels for each video frame. Finally, to aid neural network training, pixel values were normalized between 0 and 1 [37]. To do so, the maximum pixel value in each video was computed and each pixel value within the video was divided by this value. This process is common in object detection literature and further aims to mitigate the common lighting and contrast variations across different ultrasound videos in our application [29].

Automated Free Fluid Detection Methodology

We employed and extended the YoloV3 algorithm to automatically detect free fluid in ultrasound videos, relying on its fast and well-established prediction performance for object detection in various domains [29, 30] that range from detecting people, vehicles, and common objects to clinical conditions from medical images [38]. Yolo is considered one of the most common choices in real-life deployment for object detection due to its simple architectural design, low complexity, and easy implementation [38]. The initial version of the Yolo algorithm exhibited localization errors, particularly in detecting the smaller sized objects. The third version of Yolo (YoloV3) that we use was designed to overcome the aforementioned drawbacks and improved computational efficiency.

YoloV3 identifies specific objects in images using latent features learned by a deep convolutional neural network to detect an object; the architecture of this network is summarized in Fig. 1. Unlike classifier-based object detection approaches that perform inference on multiple candidate locations and scales to find high confidence detections, YoloV3 algorithm employs the same convolutional network on image region partitions, and predicts bounding boxes and probabilities for each region, significantly accelerating inference compared to former methods. These bounding boxes are weighted by the predicted probabilities to form the final detections.

In our application, we analyze each FAST exam frame-by-frame using YoloV3 and determine hemoperitoneum presence for the exam via a cumulative decision based on detections from *all frames*. For each 2D video frame input, YoloV3 makes two sets of predictions: (i) rectangular bounding boxes that circumscribe potential free fluid regions, and (ii) a confidence score in the range [0,1] that is associated with each free fluid detection. Free fluid confidence score governs the likelihood of free fluid existence and is thresholded in post-processing stages to determine

hemoperitoneum presence for each exam. In particular, we determine hemoperitoneum presence for each exam using the detection with the highest confidence score across all images. The primary reason of the choice of initially analyzing still images prior to a cumulative diagnosis per video clip is employing YoloV3, as it requires a 2-dimensional input. YoloV3 is selected due to its rapid and accurate object detection performance in various domains, as also discussed above. Crucially, using the highest confidence score across all images for final diagnosis considers the fact that all images in the video are correlated. In doing so, we reduce the effect of detected boxes that may be false negatives or positives for specific still images, as they typically attain lower confidence scores.

To aid learning from our small dataset of 94 cases, we employed transfer learning [39] by initializing the weights of the neural network with weights pre-trained on a benchmark object detection dataset named Common Objects in Context (COCO) [40]. Following initialization, we trained YoloV3 on the pairs of video frames and corresponding ground-truth free fluid boxes that belong to positive training cases only. Free fluid detection was optimized by minimizing the binary cross-entropy loss between confidence score predictions and ground-truth (positive or negative) labels, while free fluid localization was optimized by minimizing the mean-squared error between center coordinates, widths, and heights of predicted and ground-truth boxes. Training lasted for a maximum of 100 epochs and was stopped when the loss function evaluated over the validation set stopped changing significantly. Neural network weights were optimized via Adam optimization with a learning rate of 10^{-3} [41]. To better generalize the performance over unseen cases, input frames in training were perturbed via Gaussian distributed additive noise with standard deviation of 0.2 [42] and neural network weights were regularized via weight decay with regularization level of 0.02 [43].

To assess each FAST exam in the test set, we first applied the trained model on each video frame to detect free fluid
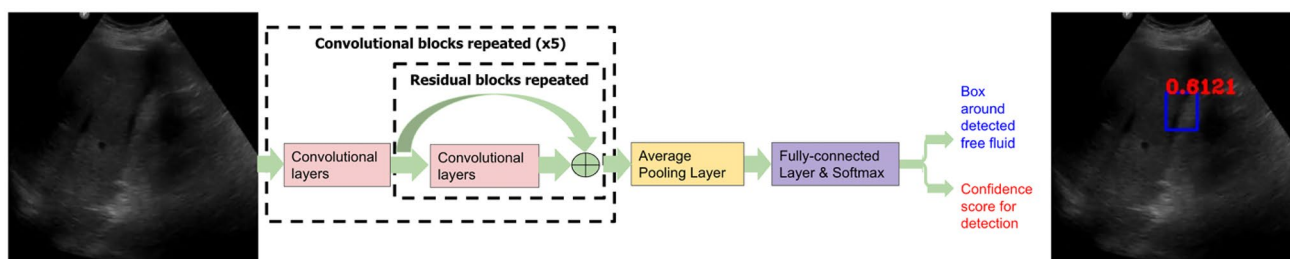

Fig. 1 Summary of the convolutional network architecture within YoloV3. This network extracts latent features via repeated convolutional and residual layers, and aggregates learned features via a fully connected block to make two sets of predictions: (1) rectangular bounding boxes that circum-

scribe potential free fluid regions, and (2) the confidence score in the range [0,1] that is associated with each free fluid detection. Each FAST exam is analyzed image-by-image and free fluid presence in the exam is determined from the most confident detection

boxes and corresponding confidence scores. If there are multiple detections in one frame, we kept the detection with the highest confidence. To combat false positive detections, we also only kept the free fluid detections that occur in at least two consecutive video frames. Having recorded the resulting detections and confidence scores for each exam, we represented each exam with the highest confidence score across all frames; we used this score to classify the case as positive or negative.

Statistical Analysis

We determined the threshold for positive or negative classification of each case based on the validation set performance. To do so, we first applied our algorithm on each exam in the validation set to obtain the highest confidence score. We determined the detection threshold as the score that maximizes the geometric mean of sensitivity and specificity over the validation set [44]. Then, we applied our algorithm on each exam in the test set to obtain the highest confidence score and thresholded each score at the best validation threshold to classify the exam as positive or negative.

We used descriptive statistics to analyze the study population and to evaluate the free fluid detection performance of our algorithm via several metrics, using the highest confidence score for each test case and the best validation threshold described above. Prior to thresholding of confidence scores, we evaluated area under the receiver operating characteristic curve (AUC) and average precision (AP). After binary classification of each case based on thresholding, we computed sensitivity, specificity, likelihood ratios, and predictive values for positive and negative classes, and accuracy.

We reported the average of each metric, along with its 95% confidence interval (CI), over 5 test folds. Detections on positive cases in each test fold were made by the corresponding trained model that has *not* seen these cases in training. As negative cases were *not* used in training, we applied the trained model that leads to the least false positive detections on the negative validation and test cases, since this model has the best potential to be deployed.

For positive cases, we also assessed the localization performance of our approach using the Intersection over Union (IOU) metric [45], computed as the ratio of the overlap between a ground-truth free fluid box (GT) and the corresponding predicted box (PR) to the area of the union of the two. Denoting the number of frames as N for a true positive case, we computed the IOU on the video frame in which the free fluid box was localized the best:

$$\text{IOU} = \max_{i \in [1, N]} \frac{GT_i \cap PR_i}{GT_i \cup PR_i}.$$

We also captured the amount of time it takes for each video frame to be analyzed automatically. In doing so, we computed the inference time for each frame in milliseconds (ms) as averaged over all cases and also reported the inference speed as the number of processed frames per second (fps). Statistical analysis was performed using the Scikit-Learn library of the Python 3.7 programming language to compute all quantitative performance metrics [46]. We also used Statistical Software for analysis REDCap hosted at clinical investigation site by an internal grant and SAS 9.4 (SAS Institute, Cary, NC) [47].

Qualitative Analysis

In addition to the quantitative analysis, we visually analyzed the free fluid localization performance of our approach. To do so, we plotted the ground-truth vs. predicted free fluid boxes on several video frames from different cases, and visually compared the locations and sizes of the boxes corresponding to the same video frames side to side. All visual results were reviewed in collaboration by the same experts as data annotation.

Results

We analyzed 44 ultrasound videos collected from patients' hemoperitoneum who had free fluid in the RUQ of the FAST exam and 50 negative controls. There were 23 free fluid positive female cases, 21 positive male cases, 25 negative female cases, and 25 negative male cases. Table 1 summarizes the demographics of our dataset. All cases were acquired by either a Philips Sparq Ultrasound System or Zonare zone ultra and with either a curvilinear or phased array probe. Table 2 summarizes the number of cases acquired by each probe type. All male positive cases had an injury mechanism of trauma while female cases are reported in Table 2 as a percentage of trauma due to a portion of cases of hemoperitoneum having a gynecologic etiology of ruptured ectopic pregnancy or rupture hemorrhagic ovarian cyst.

Our automated free fluid detection algorithm had a 95% sensitivity, 94% specificity, 95% accuracy, and 97% AUC in distinguishing positive and negative cases (see Table 3) and missed only 2 cases from male subjects out of all 44 positive cases, demonstrating the discrimination capability of the confidence scores of detected boxes. We found no differences in the performance of the algorithm based upon sex.

Along with classification of each case as positive or negative, free fluid was localized on the video frames of positive cases due to employing the YoloV3 algorithm. The IOU metric computed over the positive cases is affected by both localization performances, as well as the sizes and aspect ratios of

Table 1 Demographics. Due to missing values in age, mean and SD values are based on sample size of 92

	Negative FAST ($n=50$)	Positive FAST ($n=44$)
Age, median (IQR)	40 (29–53)	30 (25–39)
Age, mean (SD)	42.59 (17.73)	34.3 (13.21)
Gender, n (%)		
Male	25 (50)	21 (47.73)
Female	25 (50)	23 (52.27)
N/A	0 (0)	0 (0)
Race, n (%)		
White	20 (40)	16 (36.36)
Black or African American	19 (38)	18 (40.91)
Asian	1 (2)	2 (4.55)
Other	0 (0)	0 (0)
Unknown	10 (20)	8 (18.18)
Ethnicity, n (%)		
Hispanic or Latino	7 (14)	7 (15.91)
Not Hispanic or Latino	42 (84)	36 (81.82)
N/A	1 (2)	1 (2.27)
Trauma, n (%)		
Yes	30 (60)	25 (56.82)
No	19 (38)	18 (40.91)
N/A	1 (2)	1 (2.27)
Device, n (%)		
C5-2	0 (0)	5 (13.64)
C6-2	49 (98)	36 (81.82)
S4-2	1 (2)	1 (2.27)
SP5-1	0 (0)	0 (0)
N/A	0 (0)	1 (2.27)
Probe type, n (%)		
Mindray	0 (0)	0 (0)
Philips	50 (100)	42 (95.45)
Zonare	0 (0)	1 (2.27)
N/A	0 (0)	1 (2.27)

the detected boxes. As demonstrated visually (Figs. 2 and 3), our algorithm exhibited strength in good localization in finding the free fluid region of interest, while the detected box sizes and aspect ratios varied. This phenomenon resulted in 56% IOU averaged over all positive cases. Most importantly, our algorithm exhibited only 57-ms latency in image processing on a MacBook Pro laptop, equivalent to processing 18 ultrasound images per second in fps.

Table 2 Acquisition device specifications

Manufacturer	Device probe type	Number of cases
Philips	C5-2	6
Philips	C6-2	85
Philips	S4-2	2
Zonare	C6-2	1

Competing Methods

Inspired by the recent deep learning literature on automated free fluid detection from ultrasound, we implemented the 2D U-Net [28] and MaskRCNN [48] for free fluid detection and localization, as well as ResNet [26] for classifying each exam as positive or negative for free fluid presence. In particular, MaskRCNN is an object detection method similar to YoloV3, with the extension of adding a small fully connected branch for predicting a segmentation mask within each detected bounding box. Following Lin et al. [28], the U-Net architecture involved 4 encoder-decoder blocks with 64, 128, 256, and 512 channels and 4 residual units, respectively. For head-to-head comparison with YoloV3, we initialized the MaskRCNN [48] with weights pre-trained over the COCO dataset and used detected bounding box scores to classify each case for free fluid presence. Following Cheng et al. [26], we used the ResNet-50 version with weights

Table 3 Classification and localization performance of free fluid detection via YoloV3

Predicted	Actual	
	Present	Absent
Positve	42	3
Negative	2	47
Statistic	Value	95% CI
Sensitivity	95.45%	84.53 to 99.44%
Specificity	94.00%	83.45 to 98.75%
Positive likelihood ratio	15.91	5.30 to 47.75
Negative likelihood ratio	0.05	0.01 to 0.19
Disease prevalence	46.81%	36.44 to 57.39%
Positive predictive value	93.33%	82.35 to 97.68%
Negative predictive value	95.92%	85.83 to 98.92%
Accuracy	94.68%	88.02 to 98.25%
AUC	97%	88.1 to 99.8%
AP	97%	88 to 99%
IOU	56%	

AUC area under the curve, *AP* average precision, *IOU* Intersection over Union

pre-trained over the ImageNet dataset and fine-tuned the last two convolutional blocks and final fully connected block during training. For each competing method, we followed the same data preprocessing and statistical analysis steps as YoloV3, using the highest free fluid probability over all frames and thresholding this value to classify each case, with the threshold tuned over the validation set. To assess YoloV3 performance against each competing method for statistical significance, we computed the *p*-value under the two-sided *t*-test between the pair of detection metrics estimated over 94 cases along with 95% confidence intervals.

Table 4 compares the performances of YoloV3 against U-Net, MaskRCNN, and ResNet. YoloV3 consistently and significantly outperforms all competing methods in free fluid detection with *p*-value < 0.0001. Meanwhile, the free fluid localization performance assessed by IOU over the true positive detections is lower than U-Net and MaskRCNN. While MaskRCNN IOU is higher, its predicted confidence scores are not discriminative for free fluid detection, leading to significantly lower detection performance than YoloV3. As we also discuss qualitatively below, our method performs localization by drawing a box around the free fluid region to be reviewed by clinicians, rather than pixel-by-pixel exact segmentation. As a result, while the general location and region of interest of detected boxes are correct compared to the ground-truths, detected box sizes and aspect ratios vary and lower the average IOU. Our approach via YoloV3 is designed by acknowledging this trade-off, as efficient

Fig. 2 Example ground-truth vs. detected free fluid boxes for correct classification. Ground-truth image displays a box around the expert-confirmed free fluid region, while the predicted image displays a box around the predicted free fluid region by YoloV3. Next to each box, a score that indicates the algorithm confidence is displayed
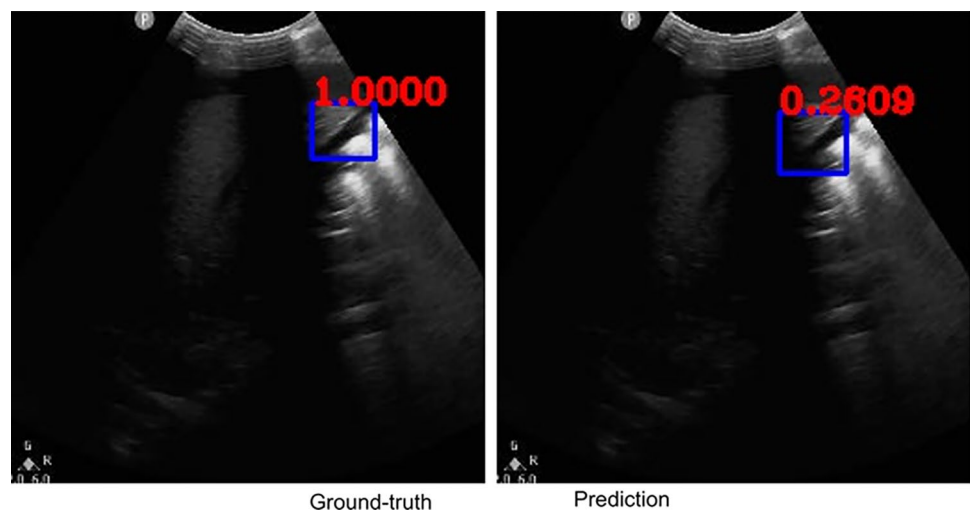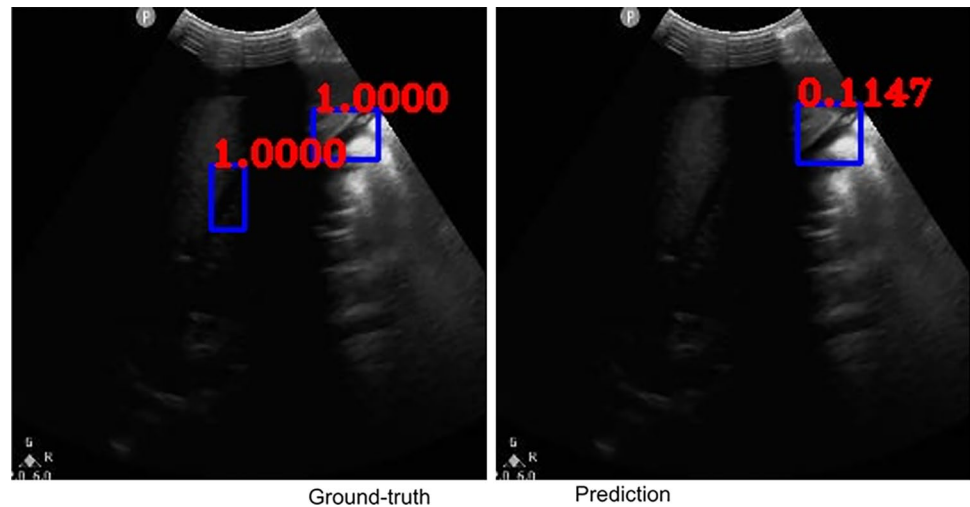
Fig. 3 Example ground-truth vs. detected free fluid boxes for correct classification

Ground-truth   Prediction

and accurate free fluid detection is the priority in point-of-care applications, rather than the exact shape and size of free fluid. Efficiency against exact segmentation is further confirmed by the time required to process each ultrasound frame: MaskRCNN processes each frame on average in 195 ms (5 fps), running 3 times slower than its object detection counterpart YoloV3 requiring only 57 ms (18 fps).

Qualitative Results

Figures 2, 3, 4, 5, 6, 7, 8, and 9 provide examples of the image analysis used to classify each case from all corresponding images. Each ground-truth image displays a box around the coded free fluid region confirmed by experts. The predicted image corresponding to each ground-truth image displays a box provided by the automated algorithm that indicates the predicted free fluid region. Next to each box, a score that varies between 0 and 1 is shown, which indicates the confidence on the coded free fluid region. As expected, for expert-confirmed ground-truth boxes, confidence score is the highest value of 1. Quantitative examples validate that our algorithm exhibits strength in good localization, while the detected box sizes and aspect ratios vary within reason, as in Figs. 6, 7, 8, and 9. Naturally, confident detections correspond to the free fluid boxes around which

there is high visual contrast, as in Fig. 6. Figures 3, 4, and 5 demonstrate the cases in which there are multiple ground-truth free fluid boxes, while our algorithm focuses on its most confident detection.

We reviewed the 2 discordant cases that were falsely labeled as negative by the algorithm (Figs. 10 and 11). Both videos had large areas of imaging artifacts from rib or other shadowing, which is often a pitfall for human operators and interpreters also. Particularly, for the case represented in Fig. 10, the larger area of free fluid was overlapped with shadows, while the higher contrast free fluid occupied a much smaller (less than 1% of the video frame) and was in a location that is harder to detect. For the case represented in Fig. 11, despite the larger size of the free fluid, it can be noted that the contrast was lower than typical examples shown in Figs. 2, 3, 4, 5, 6, 7, 8, and 9.

Figures 12 and 13 compare the localization performances of YoloV3 against U-Net and MaskRCNN on example frames from 4 different cases in each row and their corresponding ground-truth free fluid boxes. The case in Fig. 12 row 1 exhibits high contrast and is accordingly localized the best by YoloV3 in terms of size and shape, with 80% confidence. While the average IOU of MaskRCNN is higher than YoloV3, Fig. 12 row 2 and Fig. 13 row 1 exhibit the cases for which both YoloV3 and

Table 4 Classification and localization performances of the proposed approach via YoloV3 vs. U-Net, MaskRCNN, and ResNet. For each detection metric of U-Net, MaskRCNN, and ResNet, we present the p-value of the statistical difference against YoloV3 performance

	Specificity (p-value)	Sensitivity (p-value)	Accuracy (p-value)	AUC (p-value)	AP (p-value)	IOU
YoloV3	94%	95%	95%	97%	97%	56%
U-Net	70% (6×10^{-37})	64% (7×10^{-66})	70% (2×10^{-66})	74% (2×10^{-47})	75% (3×10^{-48})	60%
MaskRCNN	64% (3×10^{-50})	45% (2×10^{-95})	55% (2×10^{-105})	44% (3×10^{-100})	51% (10^{-90})	89%
ResNet	100% (2×10^{-7})	59% (8×10^{-69})	81% (5×10^{-19})	69% (2×10^{-43})	77% (10^{-25})	N/A

Fig. 4 Example ground-truth vs. detected free fluid boxes for correct classification

Ground-truth   Prediction
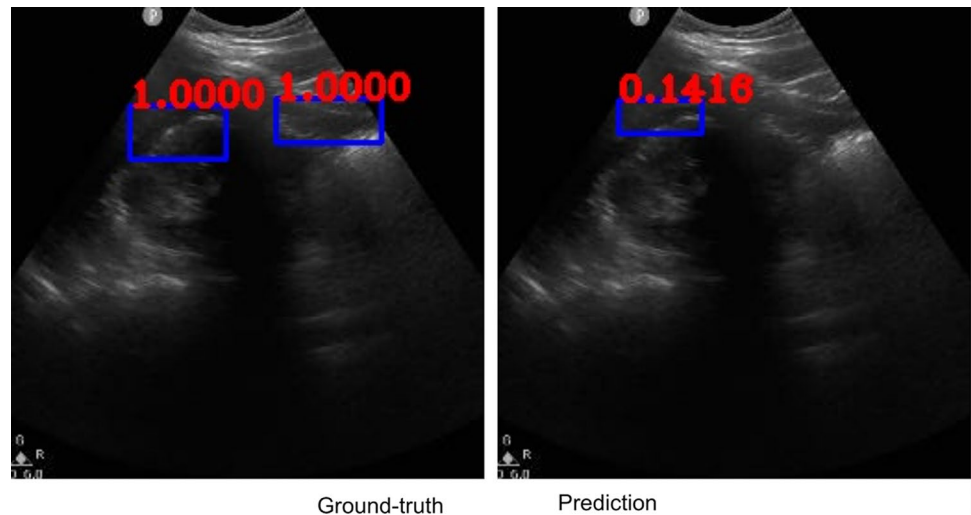
MaskRCNN cannot capture the exact shape and size of the ground-truth free fluid. In particular, MaskRCNN underestimates the free fluid size for Fig. 12 row 2 and includes part of the shadowing in addition to free fluid for Fig. 13 row 1. U-Net includes background regions in free fluid segmentation for most cases, including Fig. 12 and Fig. 13 row 1. Figure 13 row 2 exhibits a case where YoloV3 considerably underestimates free fluid size compared to MaskRCNN and U-Net, while the general location and region of interest are correct compared to the ground-truth, similar to Figs. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, and 12. As also assessed quantitatively above, YoloV3 performs rapid detection and localization by drawing a box around the free fluid, rather than the less efficient pixel-by-pixel segmentation. This design choice demonstrates significantly higher accuracy in free fluid detection than all competing methods, with a trade-off in estimating the exact shape and size of free fluid around the correctly localized region of interest.

Discussion

In this proof-of-concept exploratory study, we have demonstrated that a deep learning algorithm can rapidly and accurately identify the presence and location of free fluid in the RUQ of the FAST exam in adult patients with hemoperitoneum using, e.g., a personal laptop computer. Our model does support the hypothesis that such an application could potentially be used in real time as an aide to health care providers in diverse practice settings to identify patients who have a positive FAST from hemoperitoneum. Although there are cases in which there are multiple ground-truth free fluid boxes, our algorithm focuses on its most confident detection. As our end goal is to correctly classify each case as a binary result of positive vs. negative, this design choice provides sufficient information for free fluid detection with high prediction performance. Moreover, our algorithm can detect free fluid boxes with various sizes and aspect ratios.

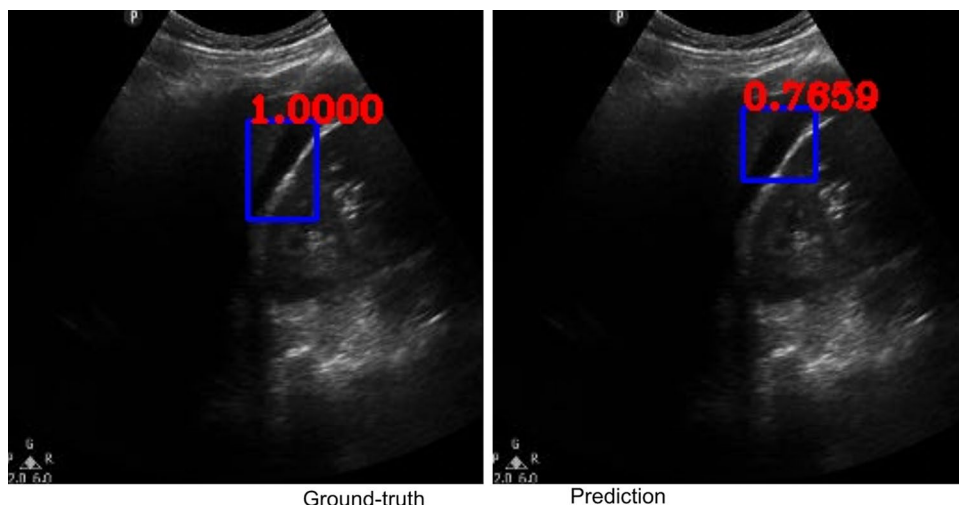**Fig. 5** Example ground-truth vs. detected free fluid boxes for correct classification
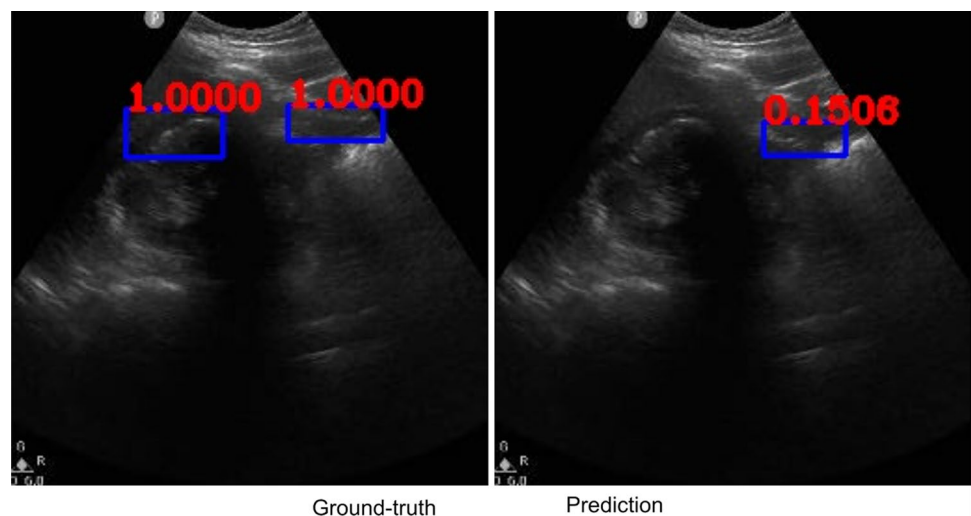
Ground-truth   Prediction

Fig. 6 Example ground-truth vs. detected free fluid boxes for correct classification

Ground-truth   Prediction

We expect the same results for all quadrants of the FAST exam and plan to expand our work to include those.

A relevant pilot study using image segmentation and shape analysis in 20 subjects with MATLAB reported computerized analysis to detect free fluid in the right upper quadrant that had a sensitivity (95% confidence interval), 100% (69.2–100%), and specificity (95% confidence interval), 90.0% (55.5–99.8%), as compared to expert review [49]. Our study instead uses deep learning and artificial intelligence to produce good test characteristics for interpretation of the RUQ FAST exams. Our approach and results can be compared to three published deep learning FAST studies [24–26] that have used deep learning for the FAST exam via ResNet and VGG networks as classification methods. In particular, ResNet-50 was used to detect free fluid from FAST exams, albeit including both trauma and non-trauma cases [25]. This deep learning algorithm receives a video frame and predicts which category the frame belongs to, e.g., the FAST exam view. Instead, our approach involving YoloV3 receives a video frame and makes a pair of predictions: location of a detected object, i.e., free fluid, as well as a confidence score for this detection. We then use the cumulative scores in each video to classify the video. In contrast, classification methods such ResNet cannot unravel the location of free fluid after classification. Moreover, we demonstrate that YoloV3 consistently outperforms ResNet-50 against all free fluid detection metrics (Table 4).

Regarding computational differences, the base neural network within Yolo has been shown to attain similar performance to ResNet, while making computations two times faster [29]. Both approaches have been so far tested on personal desktop or laptop workstations. There is also potential for point-of-care applications to transfer and run inference algorithms to compatible computers that are connected remotely to ultrasound scanners and provide this real-time feedback result. Meanwhile, our approach has the unique capability of classifying FAST exam cases with respect to free fluid existence, as well as visualizing the location of

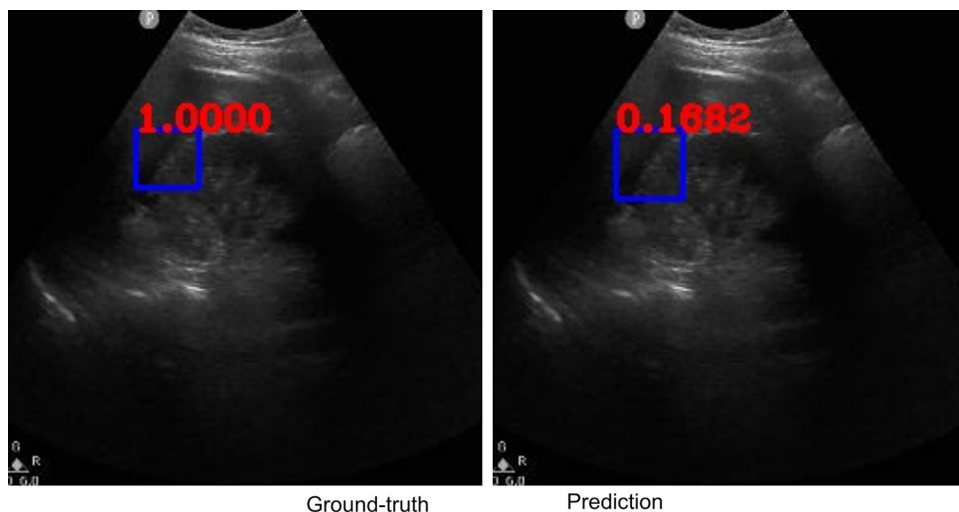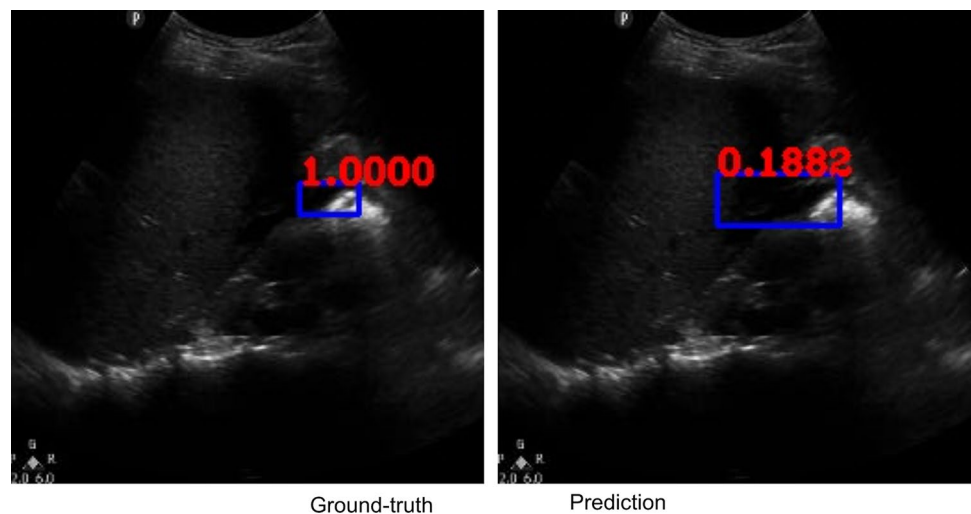**Fig. 7** Example ground-truth vs. detected free fluid boxes for correct classification

Ground-truth   Prediction

Fig. 8 Example ground-truth vs. detected free fluid boxes for correct classification


Ground-truth   Prediction

free fluid for positive cases, providing a further step towards informative point-of-care applications. These deep learning FAST studies also differ from our study because they included cases where the free fluid was not caused by hemoperitoneum. It is important for any interpretive algorithm to be tested in the population of intended use as the amount and characteristics of blood in the abdomen and other aspects of imaging could differ from other causes of free fluid such as ascites, dialysis fluid, and congestive heart failure. In our study, the algorithm was trained, validated, and tested by using only cases of hemoperitoneum.

Recently, Lin et al. described the use of a deep learning algorithm to detect ascites [28]. They used U-Net for pixel-by-pixel segmentation of free fluid. Based on the segmentation result, they diagnosed each FAST exam for ascites. The reported performance of this algorithm was lower than we observed. They achieved a sensitivity and specificity of 94.38% and 68.13% for Ascites-1 (images containing the liver or spleen). Similarly, Wu et al. [48] employed

the MaskRCNN deep learning architecture for detection and segmentation of free fluid, albeit on echocardiography images rather than FAST exams. Classification using a pixel-by-pixel approach is a more costly approach than localizing the free fluid by drawing a box, as we demonstrated in our results that MaskRCNN inference is 3 times slower than YoloV3. Crucially, our approach via YoloV3 consistently outperforms both U-Net and MaskRCNN against all free fluid detection metrics (Table 4), with a trade-off in finding the exact shape of localized free fluid.

There are several factors that can affect the performance of an algorithm intended to detect free fluid on ultrasound imaging. The amount of free fluid present and imaging artifact including shadowing and poor image quality can lead to errors in free fluid detection. YoloV3 is known to not detect small objects as accurately as larger objects due to its default field of view, while there are recent improvements in other object detection applications to potentially alleviate this effect [50, 51]. Moreover, image quality has been

Fig. 9 Example ground-truth vs. detected free fluid boxes for correct classification
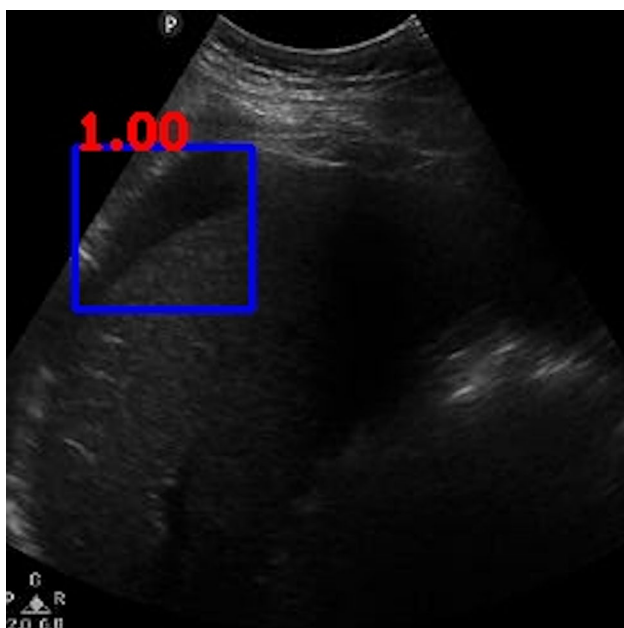

Ground-truth   Prediction
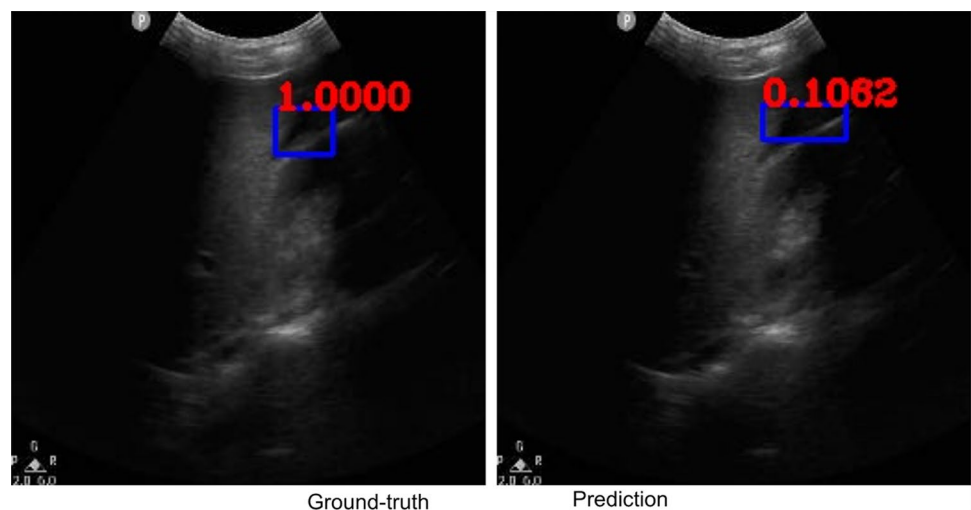
Fig. 10 Ground-truth boxes visualized on positive male case that was not detected by the algorithm

shown to affect the performance of deep learning approaches to image analysis in several clinical applications by introducing artifacts that can confuse automated algorithms in distinguishing objects of interest, which can be potentially dealt with additions of automated denoising methods prior to classification and detection [52–55]. However, our algorithm
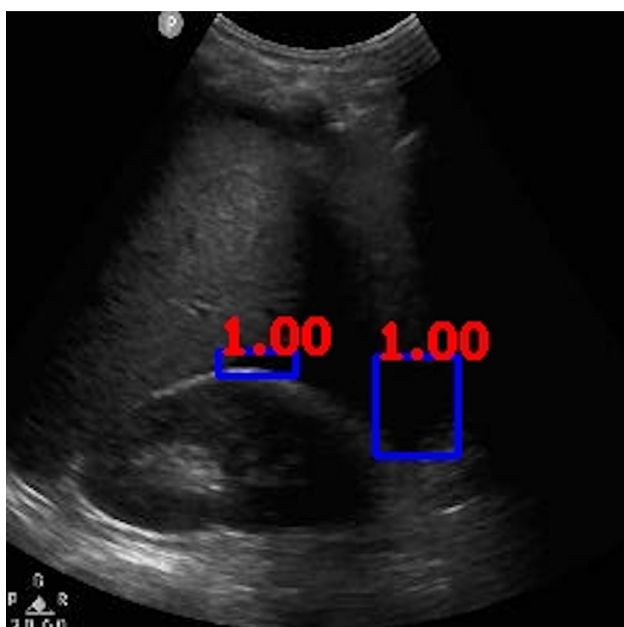

Fig. 11 Ground-truth boxes visualized on positive male case that was not detected by the algorithm

performs well with a high specificity, which is a desirable feature for the FAST exam. Like a human operator, minimizing false positives is important when this interpretation could lead to an emergent surgery or other clinical intervention. A false negative interpretation is clinically more acceptable because further diagnostic testing (CT scan or other radiologist performed imaging) would be appropriate if there is a high suspicion for injury.

The proposed technology represents a significant improvement over the state-of-the-art approaches to increasing access to POCUS, which include a broader application of POCUS training and/or telemedicine. Ultrasound image transfer for off-site interpretation has been shown to be feasible [56–59]. However, this requires not only a reliable data connection but also an on-site technician who is trained to acquire images correctly. Robotic tele-manipulation of an ultrasound probe by an off-site expert sonographer has also been explored [60], but robotic tele-manipulation inherently increases equipment cost, decreases portability, and may be difficult to use in the setting of trauma. An alternate approach is to broaden FAST training to include paramedics and non-physician personnel in rural emergency departments and urgent care settings (such as nurse practitioners and physician assistants) [61]. The cost and time burdens of such training cannot easily be estimated, but the lack of uptake to date suggests that there are many logistical hurdles. Even with a broader application of POCUS training, an automated system could provide a valuable secondary assessment. To our knowledge, our study is the first to describe an automated algorithm that is able to rapidly indicate the location of accurate positive findings to aid both untrained operators and trained clinicians in their interpretation. Further studies will be needed to understand the impact of this intervention on clinician use and accurate interpretation of the FAST exam and patient outcomes.

Our study has several limitations. The confidence intervals of our study reflect the small sample size available for analysis. Even in a large busy urban trauma center, the number of available positive FAST exams for hemoperitoneum is limited. As others have noted, many FAST exams may not be recorded, billed, or available for analysis because of the priority of quickly completing the study in a critically injured patient [62]. While we could include non-trauma causes of free fluid in the RUQ to increase the sample size of cases with a positive FAST from other causes of free fluid in the abdomen, we believe that it is important to confirm the diagnostic performance of any deep learning approach in the intended population of use (e.g. having hemoperitoneum). It will also be necessary to demonstrate that our approach could perform in a real-world setting including with an unstable patient to confirm that the analysis could be provided in a clinically useful time to result.

Fig. 12 Comparison of localization results from the proposed approach via YoloV3 (**b**, **f**) vs. U-Net (**c**, **g**) and MaskRCNN (**d**, **h**) against example ground-truth images (**a**, **e**)
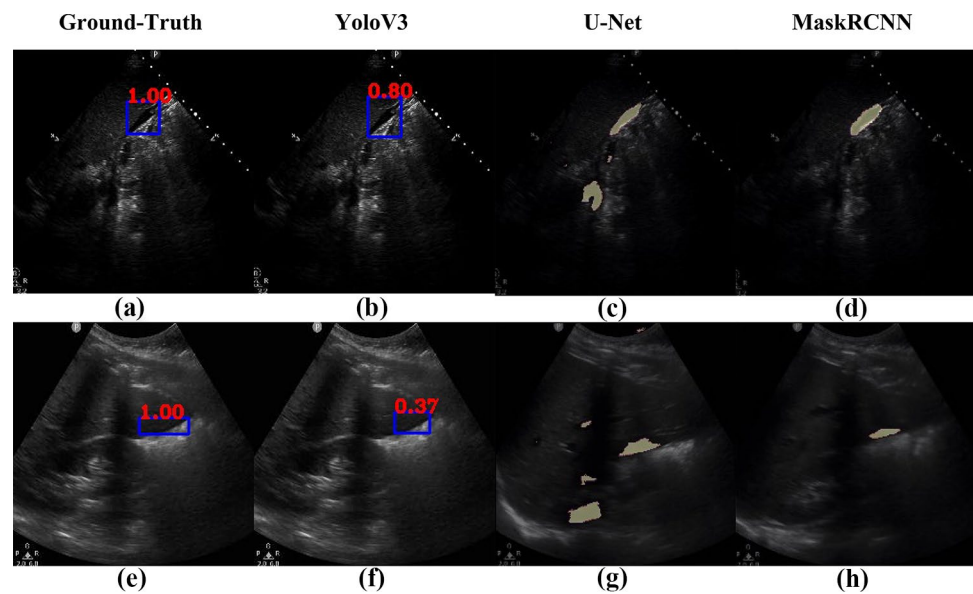
Fig. 13 Comparison of localization results from the proposed approach via YoloV3 (**b**, **f**) vs. U-Net (**c**, **g**) and MaskRCNN (**d**, **h**) against example ground-truth images (**a**, **e**)
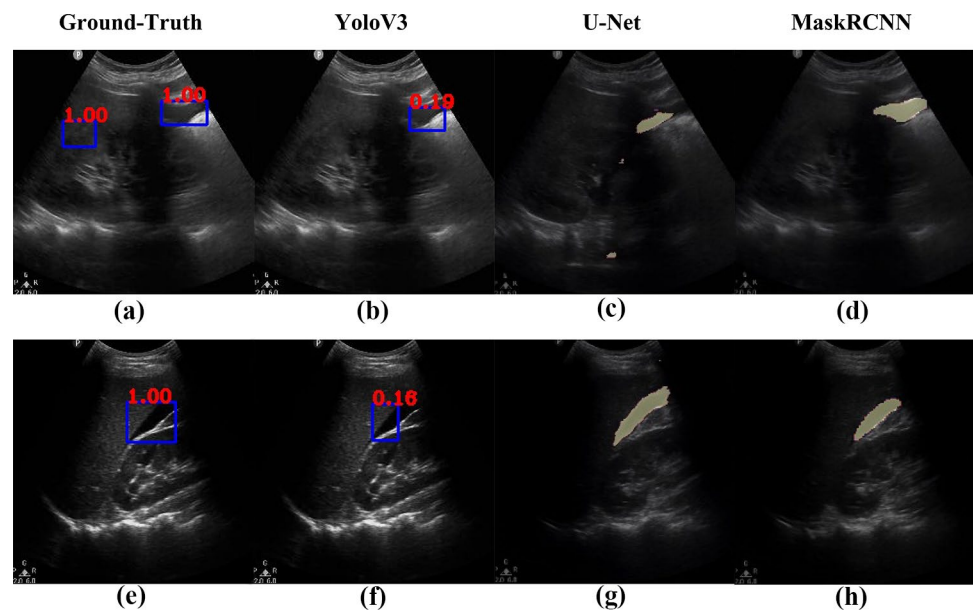
Conclusion

The accurate use of the FAST exam for the evaluation of trauma patients in a wide variety of practice environments can expedite clinical decisions that save lives. We have demonstrated that the use of a deep learning algorithm can identify the presence and location of blood in the abdomen with POCUS FAST exams in adult patients with rapid processing time to result via, e.g., a standard personal laptop. We analyzed 94 cases, which were partitioned into training, validation, and test cases, and found that the algorithm had a 95.4% sensitivity (84.5%, 99.4%), 94% specificity (83.4%, 98.7%), 94.7% accuracy (88.02%, 98.25%), and 97% AUC (89%, 99%) in free fluid detection. The algorithm also exhibited strength in localization and exhibited only 57-ms latency on average in image processing. Our results support the hypothesis that it is feasible to develop a deep learning algorithm that is accurate and provides rapid results using readily available technology that would be relevant to the diverse clinical settings where the FAST is used and can assist providers who have limited expertise.

Declarations

References

1. Savell SC, Baldwin DS, Blessing A, Medelllin KL, Savell CB, Maddry JK. Military use of point of care ultrasound (POCUS). J Spec Oper Med. 2021 Summer;21(2):35–42. https://doi.org/10.55460/AJTO-LW17.

2. Myers MA, Chin EJ, Billstrom AR, Cohen JL, Van Arnem KA, Schauer SG. Ultrasound at the Role 1: An analysis of after-action reviews from the prehospital trauma registry. Med J (Ft Sam Houst Tex). 2021 Jul-Sep;(PB 8–21–07/08/09):20–24.

3. Liang T, Roseman E, Gao M, Sinert R. The Utility of the Focused Assessment With Sonography in Trauma Examination in Pediatric Blunt Abdominal Trauma: A Systematic Review and Meta-Analysis. Pediatr Emerg Care. 2021 Feb 1;37(2):108-118. https://doi.org/10.1097/PEC.0000000000001755.

4. Fornari MJ, Lawson SL. Pediatric Blunt Abdominal Trauma and Point-of-Care Ultrasound. Pediatr Emerg Care. 2021 Dec 1;37(12):624-629. https://doi.org/10.1097/PEC.0000000000002573.

5. Chaijareenont C, Krutsri C, Sumpritpradit P, Singhatas P, Thampongsa T, Lertsithichai P, Choikrua P, Poprom N. FAST accuracy in major pelvic fractures for decision-making of abdominal exploration: Systematic review and meta-analysis. Ann Med Surg (Lond). 2020 Oct 24;60:175-181. https://doi.org/10.1016/j.amsu.2020.10.018.

6. van der Weide L, Popal Z, Terra M, Schwarte LA, Ket JCF, Kooij FO, Exadaktylos AK, Zuidema WP, Giannakopoulos GF. Prehospital ultrasound in the management of trauma patients: Systematic review of the literature. Injury. 2019 Dec;50(12):2167-2175. https://doi.org/10.1016/j.injury.2019.09.034. Epub 2019 Sep 28.

7. Qi X, Tian J, Sun R, Zhang H, Han J, Jin H, Lu H. Focused Assessment with Sonography in Trauma for Assessment of Injury in Military Settings: A Meta-analysis. Balkan Med J. 2019 Dec 20;37(1):3-8. https://doi.org/10.4274/balkanmedj.galenos.2019.2019.8.79. Epub 2019 Oct 9.

8. Netherton S, Milenkovic V, Taylor M, Davis PJ. Diagnostic accuracy of eFAST in the trauma patient: a systematic review and meta-analysis. CJEM. 2019 Nov;21(6):727-738. https://doi.org/10.1017/cem.2019.381.

9. Stengel D, Leisterer J, Ferrada P, Ekkernkamp A, Mutze S, Hoenning A. Point-of-care ultrasonography for diagnosing thoraco-abdominal injuries in patients with blunt trauma. Cochrane Database Syst Rev. 2018 Dec 12;12(12):CD012669. https://doi.org/10.1002/14651858.CD012669.pub2.

10. Quinn AC, Sinert R. What is the utility of the Focused Assessment with Sonography in Trauma (FAST) exam in penetrating torso trauma? Injury. 2011 May;42(5):482-7. https://doi.org/10.1016/j.injury.2010.07.249.

11. ATLS Subcommittee; American College of Surgeons' Committee on Trauma; International ATLS working group. Advanced trauma life support (ATLS®): the ninth edition. J Trauma Acute Care Surg. 2013 May;74(5):1363–6. https://doi.org/10.1097/TA.0b013e31828b82f5.

12. Melniker LA, Leibner E, McKenney MG, Lopez P, Briggs WM, Mancuso CA. Randomized controlled clinical trial of point-of-care, limited ultrasonography for trauma in the emergency department: the first sonography outcomes assessment program trial. Ann Emerg Med. 2006 Sep;48(3):227-35. https://doi.org/10.1016/j.annemergmed.2006.01.008. Epub 2006 Mar 24.

13. Pearl WS, Todd KH. Ultrasonography for the initial evaluation of blunt abdominal trauma: A review of prospective trials. Ann Emerg Med. 1996 Mar;27(3):353-61. https://doi.org/10.1016/s0196-0644(96)70273-1.

14. Stowell JR, Kessler R, Lewiss RE, Barjaktarevic I, Bhattarai B, Ayutyanont N, Kendall JL. Critical care ultrasound: A national survey across specialties. J Clin Ultrasound. 2018 Mar;46(3):167-177. https://doi.org/10.1002/jcu.22559. Epub 2017 Nov 13. PMID: 29131347.

15. Lewiss RE, Saul T, Del Rios M. Acquiring credentials in bedside ultrasound: a cross-sectional survey. BMJ Open. 2013 Aug 30;3(8):e003502. https://doi.org/10.1136/bmjopen-2013-003502. PMID: 23996824; PMCID: PMC3758970.

16. Ma OJ, Gaddis G. Anechoic stripe size influences accuracy of FAST examination interpretation. Acad Emerg Med. 2006 Mar;13(3):248-53. https://doi.org/10.1197/j.aem.2005.09.012. Epub 2006 Feb 22. PMID: 16495421.

17. Boniface KS, Shokoohi H, Smith ER, Scantlebury K. Tele-ultrasound and paramedics: real-time remote physician guidance of the Focused Assessment With Sonography for Trauma examination. Am J Emerg Med. 2011 Jun;29(5):477-81. https://doi.org/10.1016/j.ajem.2009.12.001. Epub 2010 Apr 13.

18. Drake AE, Hy J, MacDougall GA, Holmes B, Icken L, Schrock JW, Jones RA. Innovations with tele-ultrasound in education sonography: the use of tele-ultrasound to train novice scanners. Ultrasound J. 2021 Feb 14;13(1):6. https://doi.org/10.1186/s13089-021-00210-0.

19. Pokaprakarn T, Prieto JC, Price JT, et al. AI estimation of gestational age from blind ultrasound sweeps in low-resource settings. NEJM Evidence 2022 March 28;1(5). https://doi.org/10.1056/EVIDoa2100058

20. Laumer F, Di Vece D, Cammann VL, et al. Assessment of artificial intelligence in echocardiography diagnostics in differentiating takotsubo syndrome from myocardial infarction. JAMA Cardiol. 2022 May 1;7(5):494-503. https://doi.org/10.1001/jamacardio.2022.0183.

21. Van Sloun RJG, Cohen R, Eldar YC., Deep learning in ultrasound Imaging. Proc IEEE 2020;108(1). https://doi.org/10.1109/JPROC.2019.2932116.

22. Diniz PHB, Yin Y, Collins S. Deep Learning strategies for ultrasound in pregnancy. Eur Med J Reprod Health. 2020 Aug;6(1):73-80. Epub 2020 Aug 25.

23. Liu S, Wang Y, Yang X, et al. Deep learning in medical ultrasound analysis: a review. Engineering 2019;5(Generic):261–275. https://doi.org/10.1016/j.eng.2018.11.020.

24. Akkus Z, Cai J, Boonrod A, et al. A Survey of deep-learning applications in ultrasound: artificial intelligence-powered ultrasound for improving clinical workflow. J Am Coll Radiol. 2019 Sep;16(9 Pt B):1318–1328. https://doi.org/10.1016/j.jacr.2019.06.004.

25. Kornblith AE, Addo N, Dong R, et al. Development and validation of a deep learning strategy for automated view classification of pediatric focused assessment with sonography for trauma. J Ultrasound Med. 2022 Aug;41(8):1915-1924. https://doi.org/10.1002/jum.15868. Epub 2021 Nov 6.

26. Cheng CY, Chiu IM, Hsu MY, Pan HY, Tsai CM, Lin CR. Deep learning assisted detection of abdominal free fluid in morison's pouch during focused assessment with sonography in Trauma. Front Med (Lausanne). 2021 Sep 23;8:707437. https://doi.org/10.3389/fmed.2021.707437.

27. Taye M, Morrow D, Cull J, Smith DH, Hagan M. Deep Learning for FAST Quality Assessment. J Ultrasound Med. 2023 Jan;42(1):71-79. https://doi.org/10.1002/jum.16045. Epub 2022 Jun 30. PMID: 35770928

28. Lin Z, Li Z, Cao P, et al. Deep learning for emergency ascites diagnosis using ultrasonography images. J Appl Clin Med Phys. 2022 Jul;23(7):e13695. https://doi.org/10.1002/acm2.13695. Epub 2022 Jun 20.

29. Redmon J, Farhadi A. Yolov3: An incremental improvement. *arXiv preprint arXiv:180402767* 2018.

30. Diwan, T., Anirudh, G. and Tembhurne, J.V., 2022. Object detection using YOLO: Challenges, architectural successors, datasets and applications. Multimedia Tools and Applications, pp.1–33.

31. Lobo V, Hunter-Behrend M, Cullnan E, et al. Caudal Edge of the Liver in the Right Upper Quadrant (RUQ) view is the most sensitive area for free fluid on the FAST exam. West J Emerg Med. 2017 Feb;18(2):270-280. https://doi.org/10.5811/westjem.2016.11.30435. Epub 2017 Jan 19.

32. Ultrasound Guidelines: Emergency, Point-of-Care and Clinical Ultrasound Guidelines in Medicine. Ann Emerg Med. 2017 May;69(5):e27-e54. https://doi.org/10.1016/j.annemergmed.2016.08.457.

33. Yadav, S. and Shukla, S., 2016, February. Analysis of k-fold cross-validation over hold-out validation on colossal datasets for quality classification. In 2016 IEEE 6th International conference on advanced computing (IACC) (pp. 78–83). IEEE.

34. Azadi, S. and Karimi-Jashni, A., 2016. Verifying the performance of artificial neural network and multiple linear regression in predicting the mean seasonal municipal solid waste generation rate: A case study of Fars province, Iran. Waste management, 48, pp.14-23.

35. Lu, H.J., Zou, N., Jacobs, R., Afflerbach, B., Lu, X.G. and Morgan, D., 2019. Error assessment and optimal cross-validation approaches in machine learning applied to impurity diffusion. Computational Materials Science, 169, p.109075.

36. Sejuti, Z.A. and Islam, M.S., 2023. A hybrid CNN–KNN approach for identification of COVID-19 with 5-fold cross validation. Sensors International, 4, p.100229.

37. Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. International conference on machine learning: PMLR; 2015. p. 448–456.

38. Han, R., Liu, X. and Chen, T., 2022, October. Yolo-SG: Salience-Guided Detection Of Small Objects In Medical Images. In 2022 IEEE International Conference on Image Processing (ICIP) (pp. 4218–4222). IEEE.

39. Bengio Y, Courville A, Vincent P. Representation learning: A review and new perspectives. *IEEE transactions on pattern analysis and machine intelligence* 2013;35(8):1798-1828. https://doi.org/10.1109/TPAMI.2013.50.

40. Lin T-Y, Maire M, Belongie S, et al. Microsoft coco: Common objects in context. European conference on computer vision: Springer; 2014. P. 740-755.

41. Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv preprint arXiv:14126980 2014.

42. Dey P, Gopal M, Pradhan P, et al. On robustness of radial basis function network with input perturbation. Neural Comput & Applic 31, 523–537 (2019). https://doi.org/10.1007/s00521-017-3086-5.

43. Krogh A, Hertz J. A simple weight decay can improve generalization. Advances in neural information processing systems. Proceedings of the 4th International Conference on Neural Information Processing Systems, 950–957 (1991).

44. Fawcett T. An introduction to ROC analysis. *Pattern recognition letters* 2006;27(8):861-874.

45. Redmon J, Farhadi A.YOLO9000: Better, Faster, Stronger. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 6517–6525. https://doi.org/10.1109/CVPR.2017.690.

46. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine learning in Python. *the Journal of machine Learning research* 2011;12:2825–2830.

47. Harris PA, Delacqua G, Taylor R, Pearson S, Fernandez M, Duda SN. The REDCap Mobile Application: a data collection platform for research in regions or situations with internet scarcity. *JAMIA Open* Jul 2021;4(3):ooab078.

48. Wu, C.C., Cheng, C.Y., Chen, H.C., Hung, C.H., Chen, T.Y., Lin, C.H.R. and Chiu, I.M., 2022. Development and validation of an end-to-end deep learning pipeline to measure pericardial effusion in echocardiography. medRxiv, pp.2022–08.

49. Sjogren AR, Leo MM, Feldman J, Gwin JT. Image segmentation and machine learning for detection of abdominal free fluid in focused assessment with sonography for trauma examinations: a Pilot study. J Ultrasound Med Nov 2016;35(11):2501–2509. https://doi.org/10.7863/ultra.15.11017.

50. Jiang P, Ergu D, Liu F, Cai Y, Ma B. A review of yolo algorithm developments. Procedia Computer Science 2022;199:1066-1073.

51. Liu M, Wang X, Zhou A, Fu X, Ma Y, Piao C. UAV-YOLO: Small object detection on unmanned aerial vehicle perspective. Sensors (Basel). 2020 Apr 15;20(8):2238. https://doi.org/10.3390/s20082238.

52. Hou Y, Li Q, Zhang C, et al. The state-of-the-art review on applications of intrusive sensing, image processing techniques, and machine learning methods in pavement monitoring and analysis. *Engineering* 2021;7(6):845-856.

53. Hosseinzadeh Kassani S, Hosseinzadeh Kassani P. A comparative study of deep learning architectures on melanoma detection. Tissue Cell. 2019 Jun;58:76-83. https://doi.org/10.1016/j.tice.2019.04.009. Epub 2019 Apr 22.

54. Mishra D, Chaudhury S, Sarkar M, Manohar S, Soin AS. Segmentation of vascular regions in ultrasound images: A deep learning approach. 2018 IEEE International Symposium on Circuits and Systems (ISCAS): IEEE; 2018. p. 1–5.

55. Cheong H, Devalla SK, Pham TH, et al. Deshadowgan: a deep learning approach to remove shadows from optical coherence tomography images. *Translational Vision Science & Technology* 2020;9(2):23-23.

56. *Tablet-based Ultrasound Trial Shows Lifesaving Potential in Emergency Services.* Journal of mHealth, 2014. **1**(6): p. 17.

57. Kolbe, N., et al., *Point of care ultrasound (POCUS) telemedicine project in rural Nicaragua and its impact on patient management.* J Ultrasound, 2015. **18**(2): p. 179-85.

58. Ferreira, A.C., et al., *Teleultrasound: historical perspective and clinical application.* Int J Telemed Appl, 2015. **2015**: p. 306259.
59. Levine, A.R., et al., *Tele-intensivists can instruct non-physicians to acquire high-quality ultrasound images.* J Crit Care, 2015.
60. Georgescu, M., et al., *Remote Sonography in Routine Clinical Practice Between Two Isolated Medical Centers and the University Hospital Using a Robotic Arm: A 1-Year Study.* Telemed J E Health, 2015.
61. Heegaard, W., et al., *Prehospital ultrasound by paramedics: results of field trial.* Acad Emerg Med, 2010. **17**(6): p. 624-30.
62. Shwe S, Witchey L, Lahham S, Kunstadt E, Shniter I, Fox JC. Retrospective analysis of eFAST ultrasounds performed on trauma activations at an academic level-1 trauma center. World J Emerg Med. 2020;11(1):12-17. https://doi.org/10.5847/wjem.j.1920-8642.2020.01.002.