HHS Public Access

Author manuscript

Med Biol Eng Comput. Author manuscript; available in PMC 2024 June 24.

Published in final edited form as:

Med Biol Eng Comput. 2023 August ; 61(8): 1947–1959. doi:10.1007/s11517-023-02855-6.

Automated Detection and Localization of Pericardial Effusion from Point-of-Care Cardiac Ultrasound Examination[1]

İlkay Yıldız Potter, PhD[1], **Megan M. Leo, MD**[2,3], **Ashkan Vaziri, PhD**[1], **James A. Feldman, MD MPH**[2,3]

[1] BioSensics LLC, Newton, MA, USA

[2] Boston University Chobanian & Avedisian School of Medicine (BU), Boston, MA, USA

[3] Department of Emergency Medicine, Boston Medical Center (BMC), Boston, MA, USA

Abstract

Focused Assessment with Sonography in Trauma (FAST) exam is the standard of care for pericardial and abdominal free fluid detection in emergency medicine. Despite its life saving potential, FAST is underutilized due to requiring clinicians with appropriate training and practice. To aid ultrasound interpretation, the role of artificial intelligence has been studied, while leaving room for improvement in localization information and computation time. The purpose of this study was to develop and test a deep learning approach to rapidly and accurately identify both the presence and location of pericardial effusion on point-of-care ultrasound (POCUS) exams. Each cardiac POCUS exam is analyzed image-by-image via the state-of-the-art YoloV3 algorithm and pericardial effusion presence is determined from the most confident detection. We evaluate our approach over a dataset of POCUS exams (cardiac component of FAST and ultrasound), comprising 37 cases with pericardial effusion and 39 negative controls. Our algorithm attains 92% Specificity and 89% Sensitivity in pericardial effusion identification, outperforming existing deep learning approaches, and localizes pericardial effusion by 51% Intersection Over Union with ground-truth annotations. Moreover, image processing demonstrates only 57 msec latency. Experimental results demonstrate the feasibility of rapid and accurate pericardial effusion detection from POCUS exams for physician overread.

[1]This version of the article has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature's AM terms of use, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: https://doi.org/10.1007/s11517-023-02855-6

Corresponding author: İlkay Yıldız Potter, ilkay.yildiz@biosensics.com.

Author Contributions

All of the listed authors have participated actively in the entire study project, including study design, data acquisition, analysis, and manuscript preparation. ML, AV, JF developed the design and conduct of the study. IYP, ML, AV, JF participated in the data analysis, interpretation and manuscript preparation. IYP drafted the original manuscript. All authors participated in and approved the final submission. IYP assumes responsibility for the paper as a whole.

Competing Interests

Dr. Megan M. Leo and Dr. İlkay Yıldız Potter declare that they have no financial interests. Dr. Ashkan Vaziri and Dr. James Feldman received the research grants funding this work as investigators.

Ethics Approval

This retrospective study was deemed Exempt from review by the Institutional Review Board of Boston Medical Center/ Boston University Medical Campus.

Keywords

Focused Assessment with Sonography for Trauma; Point-of-care Ultrasound; Deep Learning; Artificial Intelligence; Free Fluid Detection

1. Introduction

The Focused Assessment with Sonography in Trauma (FAST) is a rapid point-of-care ultrasound (POCUS) assessment to screen for pericardial and abdominal free fluid. The exam incorporates four views of the abdomen as illustrated in Figure 1, including the subxiphoid view of the pericardium, i.e., the cardiac view that we study. Figure 2 displays the anatomy of this view as revealed by ultrasound. A positive FAST exam has been shown to be a strong independent predictor of the need for therapeutic surgical intervention and critical decision-making in trauma [1–4]. Timely detection of free fluid with the FAST exam has also been shown to decrease time to surgical intervention and improve resource utilization [5]. The FAST exam has been accepted as an element of the primary survey by Advanced Trauma Life Support since 1997 [6]. Studies have shown that the FAST exam has high specificity for abdominal and pericardial free fluid in both blunt and penetrating trauma [7–22], including cardiac injury [7].

In the setting of trauma, free fluid is assumed to be blood, until proven otherwise. The presence of free fluid on ultrasound, in conjunction with clinical acumen, can aid in decision making regarding the need for surgical intervention or decisions to transfer to another facility for tertiary care [23–25]. For patients in cardiac arrest, who have critically low blood pressures or loss of pulses requiring chest compressions, the use of cardiac POCUS has been shown to rapidly identify reversible causes. These causes include pericardial effusions, in which free fluid collects between the heart and pericardium, which lead to cardiac tamponade [26]. Cardiac POCUS thus allows clinicians to make rapid decisions and perform life-saving procedures once pericardial effusion is identified. In particular, identifying a pericardial effusion can indicate the presence of cardiac injury and the need for a thoracotomy, aiding in the operative cavitary triage decision [27–29]. Figure 3 is an example of a pericardial effusion identified on POCUS.

Despite its life saving potential, POCUS is underutilized in many clinical settings, as performing and interpreting ultrasound requires personnel with appropriate training and practice. The FAST exam is part of emergency medicine and many surgical residency training programs [30–33]. Emergency medicine physicians and critical care physicians may further elect to complete an Advanced Emergency Medicine Ultrasonography fellowship or take the Critical Care Echocardiography exam, respectively, to achieve advanced certification. However, in rural and community emergency departments, low-resource international settings, and pre-hospital settings, clinicians with ultrasound training are often not available. To this end, the role of automated identification via artificial intelligence has been studied by several works to aid untrained personnel in the interpretation of ultrasound applications such as pregnancy evaluation and specific organ diagnoses [34–40]. Particularly in cardiovascular ultrasound, applications of deep learning comprise imaging

view classification [41], ventricular segmentation [42], electronic medical report generation [43], and disease classification [44], with the most commonly employed architecture being U-Net [45]. Two recent studies tackled free fluid detection from FAST exams via ResNet for detection [46] and U-Net [47] for localization, albeit focusing on right upper quadrant imaging. Similar to our study, Nayak et al. [48] applied Inception-V3 for pericardial effusion detection on echocardiography exams, attaining 91% accuracy. Despite high prediction performance, this approach did not provide location of identified effusion. Later, Wu et al. [49] employed MaskRCNN for detection and segmentation of pericardial effusion on echocardiography images, attaining 84% area under the receiving operating characteristic curve (AUC).

For computer aided interpretation to be useful for providers in making clinical decisions on trauma patients, high accuracy in detection must be paired with real-time inference to guide critical and time-dependent decisions in diverse clinical settings. Providing the location of the finding also allows clinicians to confirm the results of computer aided interpretation. This is similar conceptually to an automated electrocardiogram interpretation, which is ultimately overread by clinicians. Motivated by these observations, we develop a deep learning approach by extending YoloV3 to rapidly and accurately identify the presence and location of pericardial effusion in cardiac view of POCUS exams to aid in ultrasound interpretation and analysis by clinical personnel.

2. Methods

2.1 Study Design and Data Collection

A retrospective analysis over archived POCUS images obtained in 2010–2018 from adult patients who had FAST or POCUS examinations performed by emergency physicians was conducted. The study was deemed exempt from review by the Institutional Review Board of the study site. In order to obtain an adequate sample size of cases with pericardial effusions, both traumatic and non-traumatic causes of pericardial effusion were included in the analysis. A balanced number of confirmed cases with pericardial effusion and negative controls were identified. Trauma Registry was used to match images to patient medical records; matched electronic medical records were reviewed to confirm pericardial effusion by additional imaging such as computerized tomography or operative report. Non-trauma cases with pericardial effusion were identified by reviewing the Emergency Department Ultrasound image archive (Telexy, Qpath).

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. Analysis was performed on de-identified data, with demographics, injury mechanism, confirmatory test results, ultrasound device manufacturer and probe stored on a password protected secure computer. The de-identified video clips (mp4) were used for analysis using a deep learning algorithm.

Two emergency physicians, one fellowship-trained expert in emergency ultrasonography and one senior attending credentialled in FAST and cardiac POCUS examination confirmed the final coding of each cardiac study as positive or negative according to American College of

Emergency Physician (ACEP) guidelines [50]. The resulting dataset comprised 37 cardiac ultrasound videos collected from patients with pericardial effusion, and 39 ultrasound videos collected from patients without free fluid in their pericardium. Each video was associated with a different patient, with 16 pericardial effusion positive female cases, 21 positive male cases, 17 negative female cases, and 22 negative male cases.

Two emergency physicians in collaboration manually segmented each region in each video frame of positive cases that indicate a clearly visible area of pericardial effusion in white color and heart in red color as demonstrated in Figure 3a, using PhotoPad Image Editor for Windows (NCH Software) [51]. To devise ground-truth annotations for the proposed detection algorithm, segmented pericardial effusion regions were then mapped to bounding boxes by fitting a tight rectangular bounding box to fully contain each pericardial effusion segment, as demonstrated in Figure 3b. Each case contained on average 123 (+/− 70) ground-truth pericardial effusion boxes, reaching up to 352 boxes per case. Ground-truths also exhibited a wide range of sizes: a pericardial effusion box occupied on average 5% (+/− 4.3) of a video frame, with the smallest box occupying only 0.5% and the largest box occupying up to 21% of their corresponding video frames.

The resulting dataset of 76 cases were partitioned via 5-fold cross validation in a stratified manner, keeping a uniform ratio of positive and negative cases in each set. For each fold, 8% of the cases were held-out for validation and 20% of the cases were held-out for testing. As YoloV3 is a detection method and requires ground-truth pericardial effusion boxes, we used only positive cases from the remaining 72% for training YoloV3. We then evaluated the performance of the trained model on the 8% set of positive and negative cases for validation and on the held-out 20% set of positive and negative cases for final 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.

2.2 Automated Pericardial Effusion Detection

For automatic detection and localization of pericardial effusion on ultrasound videos, we employed and extended the YoloV3 algorithm [52]. Yolo is one of the most prevalently employed object detection approaches in real-life deployment due to its simple architectural design, low complexity, and easy implementation [53]. Moreover, Yolo has been extensively tested with high prediction performance in object detection over various domains [53], detecting not only people, vehicles and common objects, but also clinical conditions from medical images [54]. While the first Yolo version demonstrated errors in localizing smaller sized objects, YoloV3 version that we use was designed to overcome these drawbacks and improved computational efficiency.

YoloV3 comprises a convolutional neural network, termed as Darknet-53, that divides the input image into regions and predicts bounding boxes and probabilities for each region. The Darknet-53 architecture is summarized in Figure 4, with detailed breakdown of layers presented in Table 2. Bounding boxes are weighted by the predicted probabilities to form object detections. As predictions are informed by the whole input image, global context is considered in object detection, aiding detection performance. Moreover, unlike former

object detection methods that apply a neural network at multiple locations and scales, the single pass inference approach of YoloV3 improves computational efficiency.

In our setting, each POCUS exam was analyzed frame by frame using YoloV3 and pericardial effusion presence for the exam was determined based on detections from all frames. In particular, YoloV3 received each 2D video frame from each exam and predicted (i) rectangular bounding boxes that contain potential pericardial effusion regions, and (ii) a confidence score in the range 0–1 that is associated with each pericardial effusion detection. Pericardial effusion confidence scores represented probabilities of pericardial effusion existence and were thresholded in inference to make a binary decision on pericardial effusion presence for each exam.

To combat overfitting over our small training dataset, transfer learning [55] was employed by initializing Darknet-53 with weights pre-trained over the benchmark Common Objects in Context (COCO) object detection dataset [56]. Following initialization, YoloV3 was trained on the pairs of video frames and corresponding ground-truth pericardial effusion boxes that belong to positive training cases only. Following Redmon et al. [52], binary cross-entropy loss between confidence score predictions and ground-truth (positive or negative) labels was minimized to train for pericardial effusion detection, while mean-squared error loss between center coordinates, widths and heights of predicted and ground-truth boxes was minimized to train for localization. Training lasted for a maximum of 100 epochs and was stopped when the loss function evaluated over the validation set did not decrease below the lowest validation loss so far for 5 epochs consistently. Adam optimization with a learning rate of 10^{-3} was employed to optimize neural network weights [57].

To further combat overfitting and aid performance generalization, input frames in training were perturbed via Gaussian distributed additive noise [58] and neural network weights were regularized via weight decay with regularization level 0.02 [59]. To apply Gaussian noise, we sampled a value from a Gaussian distribution with 0 mean and standard deviation 0.35 independently for each pixel of each training image and added the sampled value to the pixel value. We followed the literature [60, 61, 62] in employing additive independent Gaussian noise augmentation, which has been shown to substantially reduce errors due to data quality and improve robustness in predictions. Data quality challenges are well-known in ultrasound imaging due to ultrasound wave aberration and reverberation [63]. These image degradation effects are typically observed as shadows over images that can lead to errors in free fluid detection [44]. These observations motivated us to employ Gaussian noise augmentation for the trained model to combat prediction errors due to image quality.

To perform inference over each exam in the test set, the trained model was first applied on each video frame to detect pericardial effusion boxes and corresponding confidence scores. If there are multiple detections in one frame, the highest confidence detection was kept. Finally, each exam was represented with the highest confidence score across all frames, which was used to classify the case as positive or negative.

2.3 Data Preparation

In addition to the region of interest, each ultrasound video frame includes image acquisition details such as frequency in the margins. Thus, we began preprocessing with first cropping each video frame by 5% from left and right sides to zoom into the region of interest. In doing so, image data and pericardial effusion regions were not discarded. Following the common practice in object detection literature [52], we resized each cropped video frame to obtain a uniform size of 256 × 256 pixels. Finally, we normalized pixel values between 0–1, so as to accelerate convergence of neural network training [64] and mitigate the common lighting and contrast variations across different ultrasound videos. To do so, we computed the maximum pixel value in each video and divided each pixel value within the video by this value.

2.4 Performance Metrics

To perform inference over each exam in the test set, the trained model was applied over all images in the exam to obtain the highest confidence score, which was thresholded at 0.1 to classify the case as positive or negative. The threshold of 0.1 was inspired by Eykholt et al. [65] for better performance generalization across unseen cases.

Pericardial effusion detection performance was assessed via several performance metrics. Using the confidence scores prior to thresholding, AUC was computed. After thresholding for binary classification of each case, sensitivity, specificity, likelihood ratios and predictive values for positive and negative classes, as well as accuracy were computed via:

$$\text{Sensitivity} = \frac{\#\ of\ positive\ predictions}{\#\ of\ ground-truth\ positive\ cases} = \frac{\#\ of\ true\ positives\ (TP)}{TP\ +\ \#\ of\ false\ negatives\ (FN)} \tag{1}$$

$$\text{Specificity} = \frac{\#\ of\ negative\ predictions}{\#\ of\ ground-truth\ negative\ cases} = \frac{\#\ of\ true\ negatives\ (TN)}{TN\ +\ \#\ of\ false\ positives\ (FP)} \tag{2}$$

$$\text{Positive Likelihood Ratio} = \frac{Sensitivity}{1 - Specificity} \tag{3}$$

$$\text{Negative Likelihood Ratio} = \frac{1 - Sensitivity}{Specificity} \tag{4}$$

$$\text{Positive Predictive Value} = \frac{TP}{TP\ +\ FP}$$

(5)

$$\text{Negative Predictive Value} = \frac{TN}{TN + FN}$$

(6)

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

(7)

The average of each metric, along with its 95% confidence interval (CI), were reported 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, the trained model that leads to the least false positive detections was applied on the negative validation and test cases, as this model has the best potential to be deployed. Intersection over Union (IOU) metric [66] was used to assess pericardial effusion localization over positive cases, computed as the overlap percentage between a ground-truth pericardial effusion box and the corresponding detected box. For each positive test case, the IOU on the video frame in which the pericardial effusion box was localized the best was computed via:

$$\text{IOU} = max_{i \in [1, \#\, of\, frames]} \frac{Ground-truth\ box_i \cap Predicted\ box_i}{Ground-truth\ box_i \cup Predicted\ box_i}$$

(8)

To assess computational efficiency, the amount of time it takes for each video frame to be analyzed was also captured, i.e., latency, and reported as averaged over all cases.

3. Results

3.1 YoloV3 Prediction Performance

Our analysis involved 37 ultrasound videos collected from patients with pericardial effusion and 39 ultrasound videos collected from patients without free fluid in the pericardium on their POCUS exams. Each video was associated with a different patient, with 16 pericardial effusion positive female cases, 21 positive male cases, 17 negative female cases, and 22 negative male cases.

Automated pericardial effusion detection based on YoloV3 resulted in 92% Specificity, 89% Sensitivity, 91% Accuracy, and 94% AUC in distinguishing positive and negative cases (Table 3), missing only 4 cases out of the 37 positive cases and leading to only 3 false negative predictions. All classification metrics attained near 90% or above average prediction performance, following the typical rule-of-thumb of diagnostic testing and demonstrating the discrimination capability of the confidence probability scores of

detected boxes [67]. As data was stratified across genders, we found no differences in the performance of the algorithm based upon sex.

In addition to detection, pericardial effusion was localized on the video frames of positive cases by YoloV3. As we also demonstrate visually below, our algorithm exhibited strength in localizing the pericardial effusion region of interest, while the detected box sizes and aspect ratios varied, resulting in 51% IOU averaged over all positive cases. Crucially, our approach demonstrated only 57 milliseconds latency in image processing on a Macbook Pro laptop (2021 model Macbook Pro with Apple M1 Max chip), equivalent to processing 18 ultrasound images per second. This example assessment of processing speed exhibited potential for real-time feedback to aid clinicians.

3.2 Comparison to State-of-the-art

Following the state-of-the-art on deep learning for joint free fluid detection and localization from ultrasound exams, we implemented the 2D U-Net [47] and MaskRCNN [49] as competing methods, using the same data preparation and performance analysis steps as YoloV3. Furthermore, as an object detection baseline, we implemented a Single Shot Detector (SSD) network following Liu et al [68]. Similar to our object detection approach via YoloV3, SSD eliminates multiple bounding box generations per location and scale, and encapsulates all computations in a single network for computational efficiency. MaskRCNN also performs object detection as in YoloV3, with the difference of an additional fully-connected branch that predicts a segmentation mask inside each detected bounding box. For fair comparison with YoloV3, we initialized MaskRCNN and SSD weights via pre-training over the COCO dataset and used bounding box confidence scores to detect pericardial effusion. U-Net comprised 4 encoder-decoder blocks with 64, 128, 256, 512 channels and 4 residual units, respectively [47]. In training U-Net and MaskRCNN, we used the original manual ground-truth segmentation masks for pericardial effusion and trained segmentation via binary cross-entropy loss. To compare localization performances against YoloV3, we mapped predicted pericardial effusion segmentations to bounding boxes by fitting a tight rectangular bounding box around each pericardial effusion segment. As in YoloV3, we used the highest segmentation probability over each exam and thresholded this value to classify the exam.

Table 4 compares the performances of YoloV3, U-Net, MaskRCNN and SSD. YoloV3 consistently outperformed competing methods against *all detection metrics,* by up to 57% AUC against MaskRCNN. Our method performed rapid detection and localization of pericardial effusion by drawing a box around each detection, rather than pixel-by-pixel segmentation as in, e.g., MaskRCNN. This design choice demonstrated significantly higher accuracy in pericardial effusion detection than all competing methods, with a trade-off in estimating the exact shape and size of pericardial effusion. Thus, while detected pericardial effusion regions of interests were correct with respect to ground-truth boxes, detected box sizes and aspect ratios varied and lowered the average IOU against MaskRCNN and SSD. As efficient and accurate detection is the priority in POCUS applications, our approach precisely accounted for this trade-off. We further demonstrated the advantage of YoloV3 in this trade-off via computational efficiency: MaskRCNN exhibited 195 ms average latency,

running 3 times slower than YoloV3 that require only 57 ms. YoloV3 also fared better than the objection detection counterpart SSD that exhibited 63 ms average latency. This improvement was naturally by a smaller margin compared to MaskRCNN that performs pixel-by-pixel segmentation in addition to object detection.

3.3 Visual Results

Figure 5 demonstrates example ground-truth pericardial effusion boxes compared to the detections from our algorithm. Each ground-truth image demonstrates the box around the annotated pericardial effusion region by clinical experts. The corresponding predicted image demonstrates the box around the detected pericardial effusion region by YoloV3. The confidence score that varies between 0 and 1 is shown next to each box. Naturally, confidence score is the highest value of 1 for ground-truth boxes. Visual results validate that our algorithm exhibited strength in good localization, including very small pericardial effusion regions as in Figure 5b. While detected locations and regions of interests were correct with respect to ground-truth boxes, detected box sizes and aspect ratios varied, as also discussed quantitatively above.

Figure 6 visualizes example images from the positive cases that were missed by our algorithm. We included the original segmentation annotations in Figure 6 to discuss the potential reasons why YoloV3 was not able to correctly classify these cases as positive. Pericardial effusion regions in Figures 6a–6b are not only very small, but also exhibit much lower contrast compared to the heart regions. The image in Figure 6c was acquired from a different device than the typical Philips models used in our dataset; we believe that the atypical acquisition setting made this case particularly challenging. Finally, our algorithm could in fact detect the top pericardial effusion region in Figure 6d, but with a low confidence score of 0.01 that did not pass the detection threshold. While the detection threshold can be further tuned to tackle such rare cases, we observed that 0.1 established a good balance between average specificity and sensitivity as shown in Table 3.

4. Discussion

We proposed an automated deep learning approach for detecting the existence of pericardial effusion, as well as localizing pericardial effusion in positive cardiac POCUS studies by drawing a box around the detected pericardial effusion region. Our algorithm attains 92% Specificity and 89% Sensitivity, outperforming existing deep learning approaches, while demonstrating *only* 57 msec latency and confirming our contribution in rapid detection.

Figure 5c demonstrates a case in which there are multiple ground-truth pericardial effusion boxes, while our algorithm focuses on its most confident detection per frame, prioritizing the goal of not missing positive cases. Moreover, for some positive cases, detected boxes contain both the pericardial effusion region, as well as a section of the heart, as in Figure 5d. This behavior is natural, as the algorithm typically observes pericardial effusion appearing close to the heart. These design choices support the long-term goal of aiding a clinical care environment via rapid detections, while clinical personnel are expected to review the positive-flagged cases and make the final decision on pericardial effusion presence and location.

Our study is related to recent works on free fluid detection from right upper quadrant (RUQ) FAST exams using ResNet [46] and pericardial effusion detection from echocardiography exams using Inception-v3 [48]. These are classification methods that receive a video frame to predict which category the frame belongs to, such as the existence of free fluid. Unlike our approach via YoloV3, classification architectures cannot unravel the location of free fluid, while location of the finding is a significant additional information for clinicians to confirm the results of computer aided interpretation. Regarding computational differences, YoloV3 has been shown to attain similar prediction performance to ResNet with two times faster processing time [52]. Moreover, Inception-V3 with similar computation time to YoloV3 has been reported to attain 20% lower accuracy in medical imaging applications [69]. Crucially, our approach has the unique capability of classifying POCUS cardiac exams based on pericardial effusion existence, as well as visualizing the location of pericardial effusion for positive cases, providing a further step towards such informative point-of-care applications.

More recently, Lin et al. [47] used the 2D U-Net architecture for pixel-by-pixel segmentation of free fluid from right upper quadrant FAST exams and diagnosed each exam based on the segmentation result. More similar to our study, Wu et al. [49] employed the MaskRCNN deep learning architecture for segmentation and detection of pericardial effusion on echocardiography images. As existing deep learning models for pericardial effusion detection in cardiac POCUS exams require segmentation for localization, our main contribution is to take the first steps in rapid and single shot detection and localization via YoloV3. Classification from pixel-by-pixel segmentation is less efficient than localizing free fluid by drawing a box [70], as we demonstrated in our results that MaskRCNN exhibited 3 times more latency than YoloV3. Furthermore, YoloV3 consistently outperformed both U-Net and MaskRCNN against all pericardial effusion detection metrics (Table 4), with a trade-off in finding the exact shape of localized pericardial effusion.

Pericardial effusion identification via deep learning has also been explored via other imaging modalities, including Computed Tomography. Liu et al. [71] applied holistically-nested convolutional networks (HNN) and U-Net on a dataset of 25 CT scans (1206 images). They focused on pixel-by-pixel segmentation of free fluid, attaining 76% Dice score for HNN and 77% Dice score for U-Net. Similarly, Wilder-Smith et al. [72] employed an extension of U-Net (nnU-Net), which was trained on 316 cases. The sensitivity and specificity were 97% and 100% for pericardial effusion detection, and 90% and 84% for diagnosing hemopericardium. Diagnostic accuracy of ultrasound has been reported to be affected by the presence of clots, complex loculations and post-surgical changes [73]; these observations agree with the higher performance metrics of CT-based deep learning methods compared to our results on ultrasound exams. Meanwhile, our focus is on aiding timely detection in emergency medicine and trauma critical care applications, for which ultrasound is the primary tool for diagnosis [74].

Our study was limited by the sample size of 76 cases and by including both traumatic and non-traumatic cases of pericardial effusion in our analysis. The small sample size resulted in large confidence intervals over the performance metrics reported in Table 3, while attaining high sensitivity and specificity metrics on average. Regarding the inclusion of non-traumatic

cases, free fluid should appear similar on ultrasound in non-traumatic causes of pericardial effusion, while traumatic pericardial ultrasound findings could include clot. Moreover, even in a large urban Level 1 trauma center, traumatic cardiac injury is very rare and emergent circumstances often lead to many cases having an absence of documented POCUS images. In particular, traumatic cardiac injury has been documented to have occurred in 426 of the 107,549 trauma patients (0.4%), of whom 160 patients (37.6%) suffered penetrating trauma [75]. This phenomenon limits the number of cases with traumatic cardiac injury available for retrospective analysis of FAST examinations and make prospective validation difficult.

5. Conclusion

We established the first steps in a deep learning approach to rapidly and accurately identify the presence and location of pericardial effusion in cardiac POCUS exams (including the cardiac component of the FAST exam), extending the state-of-the-art YoloV3 algorithm. Our algorithm attained 92% Specificity and 89% Sensitivity in pericardial effusion identification, outperforming existing deep learning approaches, and localized pericardial effusion with 51% Intersection over Union with ground-truth annotations. Moreover, image processing demonstrated only 57 msec latency, equivalent to 18 frames per second. Experimental 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 diverse clinical settings for physician overread.

The proposed method and analysis were part of an ongoing project to employ deep learning for interpreting FAST examination including all of its component images. Future steps of the proposed method will involve including other quadrants of the FAST exam for free fluid detection such as pelvic quadrant or including extended FAST with the additional lung views to detect pleural effusions (hemothorax in trauma), as well as lung sliding while accessing for pneumothorax. The pelvic quadrant may introduce an additional challenge due to the anatomical differences between genders. On the technical approach side, the literature on Yolo and other deep learning-based object detection models develops fast, with newer versions that are even more rapid and accurate in detections than YoloV3. While our work took the first steps in a rapid object detection approach for identifying the presence and location of pericardial effusion in cardiac POCUS exams, extensions via newer models can bring the proposed approach even closer to real-life application, with additional testing on computing devices deployed in hospital settings.

Acknowledgements

This version of the article has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature's AM terms of use, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: https://doi.org/10.1007/s11517-023-02855-6

We thank the Department of Surgery Section of Trauma and Acute Care Surgery and Ms. Heidi A. Wing, Trauma Registry Supervisor at Boston Medical Center as well as research assistants Samantha Roberts, MPH, Tyler Pina, Shinelle Kirk, Haley Connelly and all the research staff who contributed countless hours to this study. Ms. Ijeoma Okafor MPH assisted in the data analysis.

Funding

Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number R44GM123821. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. National Center for Advancing Translational Sciences, National Institutes of Health, through BU-CTSI Grant Number 1UL1TR001430 provided support for this study through the REDCap electronic data capture tools hosted at Boston University. Dr. Feldman is supported in part by UL1TR001430.

Biographies

Dr. İlkay Yıldız Potter is a Senior Research Scientist, leading several medical imaging projects with machine learning analysis. She has extensive experience in machine learning research across many domains, with 21+ peer-reviewed published articles.

Dr. Megan M. Leo is Director of Emergency Ultrasound at BMC and an Assistant Professor of Emergency Medicine at BU. She has extensive experience in emergency ultrasound training and served as co-investigator or PI on several relevant research studies.

Dr. Ashkan Vaziri is the CEO of BioSensics, overseeing digital healthcare projects on imaging and wearable sensors. He served as co-investigator or PI on several relevant research studies, and contributed to the commercialization of 4 medical devices.

Dr. James A. Feldman is a Professor of Emergency Medicine at BU with extensive experience in the development of novel diagnostics. He served as co-investigator or PI for several research studies including new technologies to aid physician decision making.

References

1. Rose JS, et al. , The fast is positive, now what? Derivation of a clinical decision rule to determine the need for therapeutic laparotomy in adults with blunt torso trauma and a positive trauma ultrasound. J Emerg Med, 2005. 29(1): p. 15–21. [PubMed: 15961002]
2. Moylan M, et al. , Association between a positive ED FAST examination and therapeutic laparotomy in normotensive blunt trauma patients. J Emerg Med, 2007. 33(3): p. 265–71. [PubMed: 17976554]
3. Helling TS, Wilson J, and Augustosky K, The utility of focused abdominal ultrasound in blunt abdominal trauma: a reappraisal. Am J Surg, 2007. 194(6): p. 728–32; discussion 732–3. [PubMed: 18005762]
4. Quinn AC and Sinert R, What is the utility of the Focused Assessment with Sonography in Trauma (FAST) exam in penetrating torso trauma? Injury, 2011. 42(5): p. 482–7. [PubMed: 20701908]
5. Melniker LA, et al. , 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. 48(3): p. 227–35. [PubMed: 16934640]
6. American College of Surgeons Committee on Trauma (1997) Advanced Trauma Life Support Course for Physicians. American College of Surgeons: Chicago.
7. Tayal VS, et al. , FAST (focused assessment with sonography in trauma) accurate for cardiac and intraperitoneal injury in penetrating anterior chest trauma. J Ultrasound Med, 2004. 23(4): p. 467–72. [PubMed: 15098863]
8. Ma OJ, et al. , Prospective analysis of a rapid trauma ultrasound examination performed by emergency physicians. J Trauma, 1995. 38(6): p. 879–85. [PubMed: 7602628]
9. Brooks A, et al. , Prospective evaluation of non-radiologist performed emergency abdominal ultrasound for haemoperitoneum. Emerg Med J, 2004. 21(5): p. 580–1. [PubMed: 15333536]

10. Rozycki GS, et al. , A prospective study of surgeon-performed ultrasound as the primary adjuvant modality for injured patient assessment. J Trauma, 1995. 39(3): p. 492–8; discussion 498–500. [PubMed: 7473914]
11. Soundappan SV, et al. , Diagnostic accuracy of surgeon-performed focused abdominal sonography (FAST) in blunt paediatric trauma. Injury, 2005. 36(8): p. 970–5. [PubMed: 15982655]
12. Kimura A and Otsuka T, Emergency center ultrasonography in the evaluation of hemoperitoneum: a prospective study. J Trauma, 1991. 31(1): p. 20–3. [PubMed: 1986127]
13. Biffl WL, et al. , Management of patients with anterior abdominal stab wounds: a Western Trauma Association multicenter trial. J Trauma, 2009. 66(5): p. 1294–301. [PubMed: 19430229]
14. O'Connor G, et al. , Looking beyond Morison's pouch in focused assessment with sonography for trauma: penetrating hepatobiliary trauma and a new sign for emergency physicians. Emerg Med J, 2013. 30(9): p. 778–9. [PubMed: 22492124]
15. Hoffmann B, Nguyen H, and Hill HF, Diaphragmatic laceration after penetrating trauma: direct visualization and indirect findings on focused assessment with sonography for trauma in the emergency department. J Ultrasound Med, 2009. 28(9): p. 1259–63. [PubMed: 19710226]
16. Boulanger BR, et al. , The routine use of sonography in penetrating torso injury is beneficial. J Trauma, 2001. 51(2): p. 320–5. [PubMed: 11493792]
17. Kirkpatrick AW, et al. , The hand-held ultrasound examination for penetrating abdominal trauma. Am J Surg, 2004. 187(5): p. 660–5. [PubMed: 15135687]
18. Soffer D, et al. , A prospective evaluation of ultrasonography for the diagnosis of penetrating torso injury. J Trauma, 2004. 56(5): p. 953–7; discussion 957–9. [PubMed: 15179232]
19. Branney SW, et al. , Quantitative sensitivity of ultrasound in detecting free intraperitoneal fluid. J Trauma, 1995. 39(2): p. 375–80. [PubMed: 7674411]
20. Bode PJ, et al. , Abdominal ultrasound as a reliable indicator for conclusive laparotomy in blunt abdominal trauma. J Trauma, 1993. 34(1): p. 27–31. [PubMed: 8437192]
21. Jehle D, Guarino J, and Karamanoukian H, Emergency department ultrasound in the evaluation of blunt abdominal trauma. Am J Emerg Med, 1993. 11(4): p. 342–6. [PubMed: 8216513]
22. McGahan JP, et al. , Use of ultrasonography in the patient with acute abdominal trauma. J Ultrasound Med, 1997. 16(10): p. 653–62; quiz 663–4. [PubMed: 9323670]
23. Maitra S, et al. , When FAST is a FAFF: is FAST Scanning Useful in Non-Trauma Patients? Ultrasound, 2008. 16(3): p. 165–168.
24. Moore C, et al. , Free fluid in Morison's pouch on bedside ultrasound predicts need for operative intervention in suspected ectopic pregnancy. Acad Emerg Med, 2007. 14(8): p. 755–8. [PubMed: 17554008]
25. Volpicelli G, et al. , Point-of-care multiorgan ultrasonography for the evaluation of undifferentiated hypotension in the emergency department. Intensive Care Med, 2013. 39(7): p. 1290–8. [PubMed: 23584471]
26. Gaspari R, Weekes A, Adhikari S, et al. Emergency department point-of-care ultrasound in out-of-hospital and in-ED cardiac arrest. Resuscitation. 2016;109:33–39. [PubMed: 27693280]
27. Tayal VS, Beatty MA, Marx JA, Tomaszewski CA, Thomason MH. FAST (focused assessment with sonography in trauma) accurate for cardiac and intraperitoneal injury in penetrating anterior chest trauma. J Ultrasound Med. 2004 Apr;23(4):467–72. doi: 10.7863/jum.2004.23.4.467. [PubMed: 15098863]
28. Netherton S, Milenkovic V, Taylor M, & Davis P (2019). Diagnostic accuracy of eFAST in the trauma patient: A systematic review and meta-analysis. Canadian Journal of Emergency Medicine, 21(6), 727–738. doi:10.1017/cem.2019.381 [PubMed: 31317856]
29. Hall MK, Omer T, Moore CL, Taylor RA. Cost-effectiveness of the Cardiac Component of the Focused Assessment of Sonography in Trauma Examination in Blunt Trauma. Acad Emerg Med 2016 Apr;23(4):415–23. doi: 10.1111/acem.12936. Epub 2016 Mar 25. [PubMed: 26857839]
30. Moore CL, Molina AA, and Lin H, Ultrasonography in community emergency departments in the United States: access to ultrasonography performed by consultants and status of emergency physician-performed ultrasonography. Ann Emerg Med, 2006. 47(2): p. 147–53. [PubMed: 16431225]

31. Akhtar S, et al., Resident training in emergency ultrasound: consensus recommendations from the 2008 Council of Emergency Medicine Residency Directors Conference. Acad Emerg Med, 2009. 16 Suppl 2: p. S32–6.
32. Counselman FL, et al. , The status of bedside ultrasonography training in emergency medicine residency programs. Acad Emerg Med, 2003. 10(1): p. 37–42. [PubMed: 12511313]
33. Freitas ML, Frangos SG, and Frankel HL, The status of ultrasonography training and use in general surgery residency programs. J Am Coll Surg, 2006. 202(3): p. 453–8. [PubMed: 16500250]
34. Wehbe RM, Thomas JD. Validating Deep Learning to Distinguish Takotsubo Syndrome From Acute Myocardial Infarction-Beware of Shortcuts, Human Supervision Required. JAMA Cardiol. 2022 May 1;7(5):477–479. doi: 10.1001/jamacardio.2022.0193. [PubMed: 35353132]
35. 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). doi: 10.1056/EVIDoa2100058
36. Wilson M, Chopra R, Wilson MZ, et al. Validation and Clinical Applicability of Whole-Volume Automated Segmentation of Optical Coherence Tomography in Retinal Disease Using Deep Learning. JAMA Ophthalmol. 2021 Sep 1;139(9):964–973. doi: 10.1001/jamaophthalmol.2021.2273. [PubMed: 34236406]
37. Van Sloun RJG, Cohen R, Eldar YC., Deep Learning in Ultrasound Imaging. Proc IEEE 2020;108(1). doi: 10.1109/JPROC.2019.2932116.
38. 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. [PubMed: 33123376]
39. Liu S, Wang Y, Yang X, et al. Deep Learning in Medical Ultrasound Analysis: A Review. Engineering 2019;5(Generic):261–275. 10.1016/j.eng.2018.11.020.
40. 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. doi: 10.1016/j.jacr.2019.06.004. [PubMed: 31492410]
41. Madani A, Arnaout R, Mofrad M and Arnaout R, 2018. Fast and accurate view classification of echocardiograms using deep learning. NPJ digital medicine, 1(1), p.6. [PubMed: 30828647]
42. Dong S, Luo G, Wang K, Cao S, Li Q, and Zhang H. A combined fully convolutional networks and deformable model for automatic left ventricle segmentation based on 3d echocardiography. BioMed research international, 2018: 5682365, 2018. [PubMed: 30276211]
43. Moradi M, Guo Y, Gur Y, Negahdar M, and Syeda-Mahmood T. A cross-modality neural network transform for semi-automatic medical image annotation. In Med Image Comput Comput Assist Interv, volume 9901 of Lect Notes Comput Sci, pages 300–307, 2016. doi: 10.1007/978-3-319-46723-8_35.
44. 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. doi: 10.1001/jamacardio.2022.0183. [PubMed: 35353118]
45. Litjens G, Ciompi F, Wolterink JM, de Vos BD, Leiner T, Teuwen J and Išgum I, 2019. State-of-the-art deep learning in cardiovascular image analysis. JACC: Cardiovascular imaging, 12(8 Part 1), pp.1549–1565. [PubMed: 31395244]
46. 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. doi: 10.3389/fmed.2021.707437. [PubMed: 34631730]
47. 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. doi: 10.1002/acm2.13695. 2022 Jun. [PubMed: 35723875]
48. Nayak A, Ouyang D and Ashley EA, 2020. A DEEP LEARNING ALGORITHM ACCURATELY DETECTS PERICARDIAL EFFUSION ON ECHOCARDIOGRAPHY. Journal of the American College of Cardiology, 75(11_Supplement_1), pp.1563–1563.

49. Wu CC, Cheng CY, Chen HC, Hung CH, Chen TY, Lin CHR and Chiu IM, 2022. Development and validation of an end-to-end deep learning pipeline to measure pericardial effusion in echocardiography. medRxiv, pp.2022–08.
50. Ultrasound Guidelines: Emergency, Point-of-Care and Clinical Ultrasound Guidelines in Medicine. Ann Emerg Med. 2017 May;69(5):e27–e54. doi: 10.1016/j.annemergmed.2016.08.457. [PubMed: 28442101]
51. NCH Software, Inc. (2019). PhotoPad Image Editor [Computer software]
52. Redmon J & Farhadi A "YOLOv3: An Incremental Improvement." arXiv preprint arxiv:1804.02767 (2018).
53. Diwan T, Anirudh G and Tembhurne JV, 2022. Object detection using YOLO: Challenges, architectural successors, datasets and applications. Multimedia Tools and Applications, pp.1–33.
54. 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).
55. Bengio Yoshua, Courville Aaron, and Vincent Pascal. "Representation learning: A review and new perspectives." IEEE Transactions on Pattern Analysis and Machine Intelligence 35, no. 8 (2013): 1798–1828.
56. Lin Tsung-Yi, Maire Michael, Belongie Serge, Hays James, Perona Pietro, Ramanan Deva, Dollár Piotr, and Zitnick C. Lawrence. "Microsoft coco: Common objects in context." In European Conference on Computer Vision, pp. 740–755. Springer, Cham, 2014.
57. Kingma Diederik P., and Ba Jimmy. "Adam: A method for stochastic optimization." arXiv preprint arXiv:1412.6980 (2014).
58. Dey Prasenjit, Gopal Madhumita, Pradhan Payal, and Pal Tandra. "On robustness of radial basis function network with input perturbation." Neural Computing and Applications 31, no. 2 (2019): 523–537.
59. Krogh Anders, and Hertz John. "A simple weight decay can improve generalization." Advances in Neural Information Processing Systems 4 (1991).
60. Geirhos R, Rubisch P, Michaelis C, Bethge M, Wichmann FA and Brendel W, 2018. ImageNet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness. arXiv preprint arXiv:1811.12231.
61. Ford N, Gilmer J, Carlini N and Cubuk D, 2019. Adversarial examples are a natural consequence of test error in noise. arXiv preprint arXiv:1901.10513.
62. Rusak E, Schott L, Zimmermann RS, Bitterwolf J, Bringmann O, Bethge M and Brendel W, 2020. A simple way to make neural networks robust against diverse image corruptions. In Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part III 16 (pp. 53–69). Springer International Publishing.
63. Pinton GF, Trahey GE and Dahl JJ, 2011. Sources of image degradation in fundamental and harmonic ultrasound imaging using nonlinear, full-wave simulations. IEEE transactions on ultrasonics, ferroelectrics, and frequency control, 58(4), pp.754–765. [PubMed: 21507753]
64. Ioffe Sergey, and Szegedy Christian. "Batch normalization: Accelerating deep network training by reducing internal covariate shift." In International Conference on Machine Learning, pp. 448–456. PMLR, 2015.
65. Eykholt K, Evtimov I, Fernandes E, Li B, Rahmati A, Tramer F, Prakash A, Kohno T and Song D, 2018. Physical adversarial examples for object detectors. arXiv preprint arXiv:1807.07769, 1(3), p.4.
66. Redmon Joseph, and Farhadi Ali. "YOLO9000: better, faster, stronger." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7263–7271. 2017.
67. Power Michael, Fell Greg, and Wright Michael. "Principles for high-quality, high-value testing." BMJ Evidence-Based Medicine 18, no. 1 (2013): 5–10.
68. Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu CY and Berg AC, 2016. SSD: Single shot multibox detector. In Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14 (pp. 21–37). Springer International Publishing.

69. Elgendi M, Nasir MU, Tang Q, Fletcher RR, Howard N, Menon C, Ward R, Parker W and Nicolaou S, 2020. The performance of deep neural networks in differentiating chest X-rays of COVID-19 patients from other bacterial and viral pneumonias. Frontiers in medicine, 7, p.550. [PubMed: 33015100]

70. Sumit SS, Watada J, Roy A and Rambli DRA, 2020, April. In object detection deep learning methods, YOLO shows supremum to Mask R-CNN. In Journal of Physics: Conference Series (Vol. 1529, No. 4, p. 042086). IOP Publishing.

71. Liu J, Cai J, Chellamuthu K, Bagheri M, Lu L and Summers RM, 2018, April. Cascaded coarse-to-fine convolutional neural networks for pericardial effusion localization and segmentation on CT scans. In 2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018) (pp. 1092–1095). IEEE.

72. Wilder-Smith AJ, Yang S, Weikert T, Bremerich J, Haaf P, Segeroth M, Ebert LC, Sauter A and Sexauer R, 2022. Automated Detection, Segmentation, and Classification of Pericardial Effusions on Chest CT Using a Deep Convolutional Neural Network. Diagnostics, 12(5), p.1045. [PubMed: 35626201]

73. Azarbal A, LeWinter MM Pericardial Effusion. Cardiol. Clin. 2017;35:515–524. doi: 10.1016/j.ccl.2017.07.005. [PubMed: 29025543]

74. Vakamudi S, Ho N, Cremer PC Pericardial Effusions: Causes, Diagnosis, and Management. Prog. Cardiovasc. Dis. 2017;59:380–388. doi: 10.1016/j.pcad.2016.12.009 [PubMed: 28062268]

75. Tran HV, Charles M, Garrett RC, Kempe PW, Howard CA, Khorgami Z. Ten-Year Trends in Traumatic Cardiac Injury and Outcomes: A Trauma Registry Analysis. Ann Thorac Surg. 2020 Sep;110(3):844–848. doi: 10.1016/j.athoracsur.2019.12.038. Epub 2020 Feb 4.

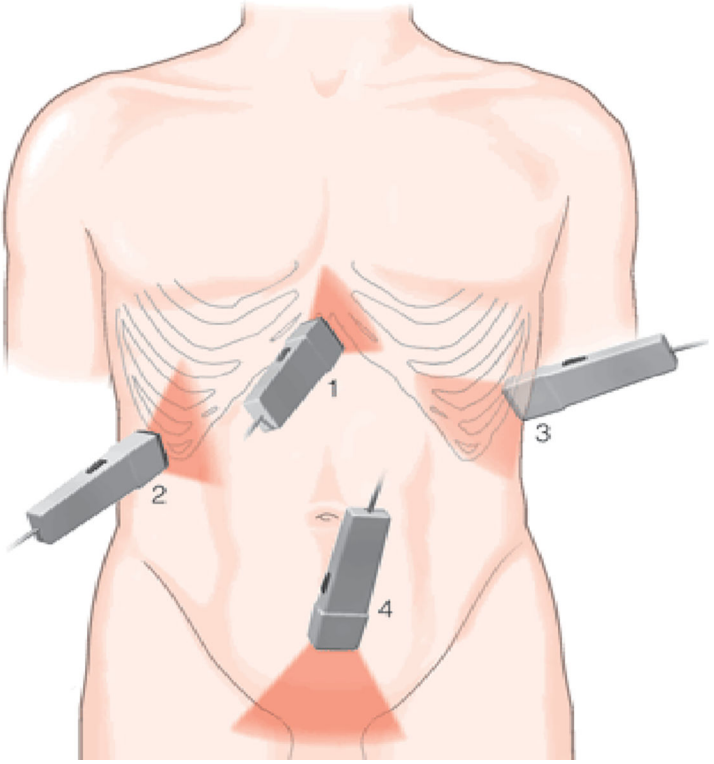

Fig. 1.

FAST is a rapid bedside ultrasound examination to screen for abdominal free fluid. The exam consists of four views: 1) subxiphoid pericardium, 2) right upper quadrant abdominal, 3) left upper quadrant abdominal, and 4) the suprapubic pelvic. This study focuses on pericardial effusion detection, in which free fluid collects between the heart and pericardium.

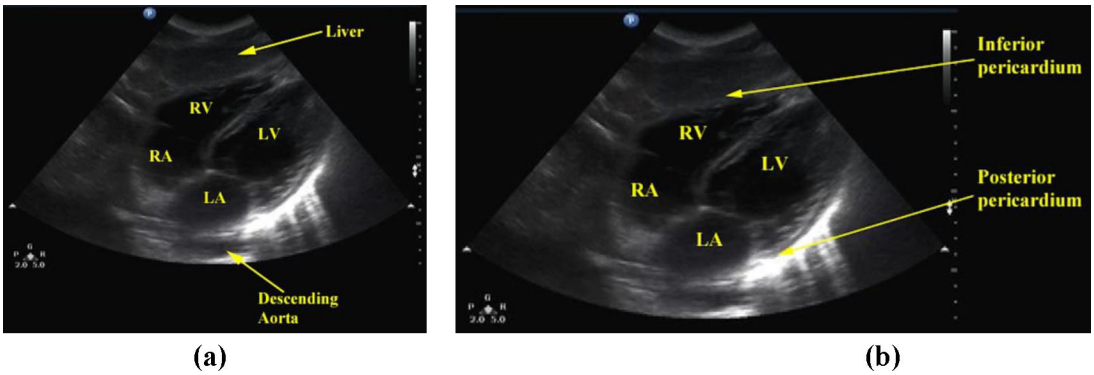

Fig. 2.

Anatomy of the subxiphoid pericardial view, i.e., the cardiac view that we study, as revealed by ultrasound. Four chambers of the heart comprising left atrium (LA), right atrium (RA), left ventricle (LV) and right ventricle (RV), as well as the surrounding organs and anatomical parts are annotated.

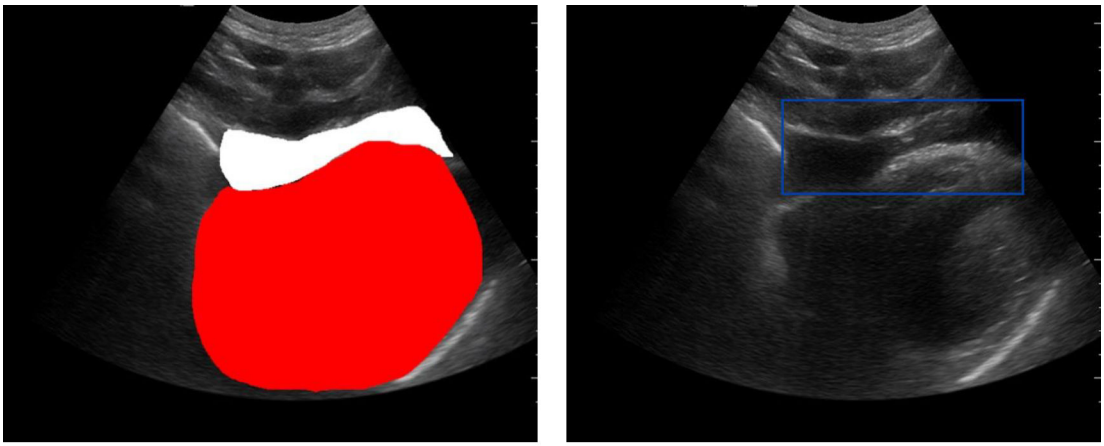

Fig. 3.

An example of a pericardial effusion identified on POCUS, in which free fluid collects between the heart and pericardium. (a) demonstrates manual segmentation annotations of pericardial effusion in white color and heart in red color. To prepare the ground-truth pericardial effusion annotations for the detection algorithm, we mapped pericardial effusion regions to bounding boxes by fitting a tight rectangular bounding box to fully contain each pericardial effusion region, as in (b).

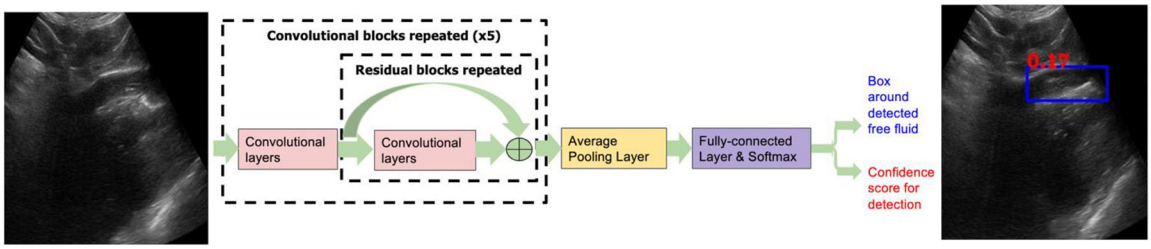

Fig. 4.

YoloV3 deep learning architecture for automated pericardial effusion detection. YoloV3 receives a 2D video frame and makes two sets of predictions: (i) rectangular bounding boxes that circumscribe potential pericardial effusion regions, and (ii) a confidence score in the range 0–1. Each cardiac POCUS exam is analyzed image-by-image and pericardial effusion presence in the exam is determined from the most confident detection.

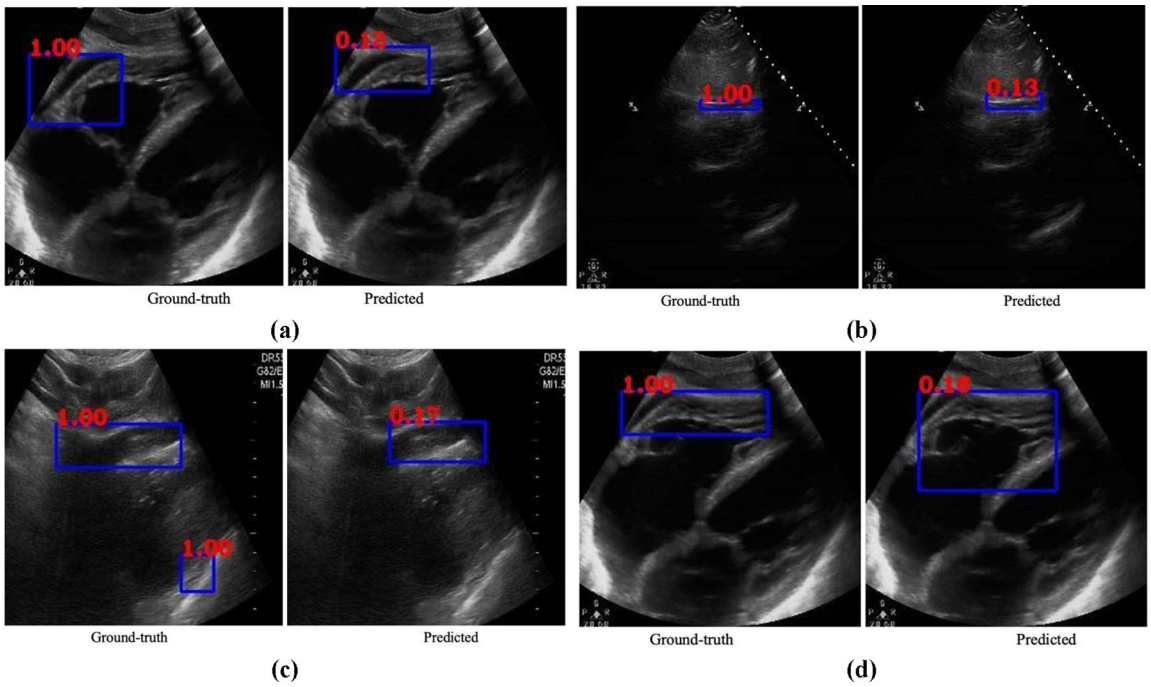

Fig. 5.

Example ground-truth vs. detected pericardial effusion boxes. The ground-truth image displays a box around the coded pericardial effusion region by experts. The corresponding predicted image displays a box provided by YoloV3 that indicates the predicted pericardial effusion region. Next to each box, a score that varies between 0 and 1 is shown, indicating the confidence on the coded pericardial effusion.

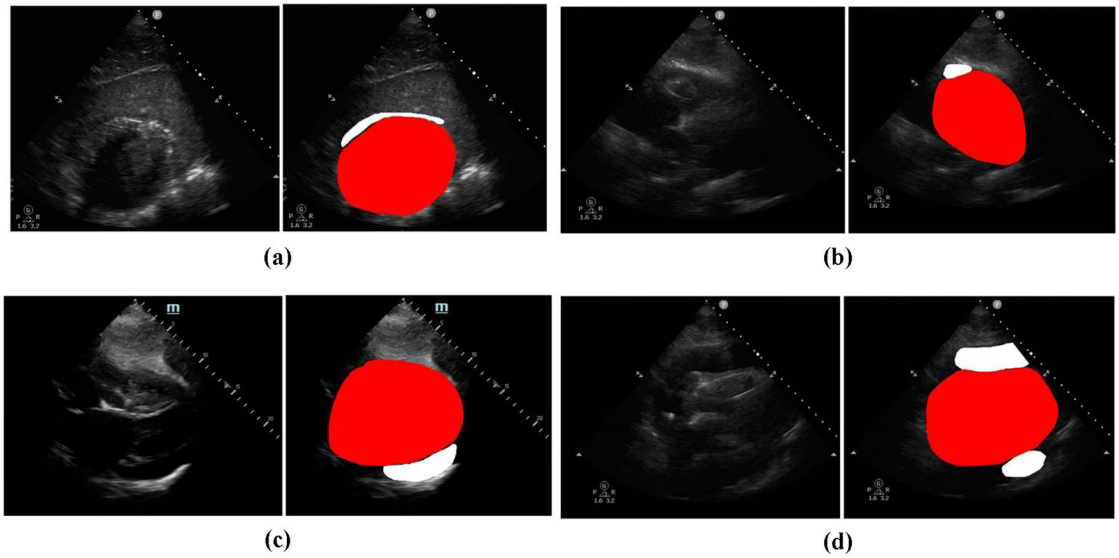

Fig. 6.

Example images from 4 ground-truth positive cases that were missed by the algorithm. Each case has an example image without annotation on the left and the corresponding annotations on the right, with pericardial effusion masked with white and heart masked with red.

Table 1:

Demographics of the cases in our dataset

	Negative ECHO (n=39)	Positive ECHO (n=37)
Age, median (IQR)	44 (34–61)	59 (53–72)
Age, mean (SD)	47 (18.96)	60.94 (15.99)
Gender, n (%)		
Male	22 (57)	22 (61)
Female	17 (43)	15 (39)
N/A	0 (0)	0 (0)
Race, n (%)		
White	17 (42.5)	11 (30.56)
Black or African American	13 (32.5)	17 (47.22)
Asian	1 (2.5)	1 (2.78)
Other	0 (0)	0 (0)
Unknown	8 (22.5)	8 (19.44)
Ethnicity, n (%)		
Hispanic or Latino	8 (20)	6 (16.67)
Not Hispanic or Latino	30 (75)	29 (80.56)
N/A	1 (5)	2 (2.78)
Trauma, n (%)		
Yes	18 (47)	6 (14)
No	20 (50)	30 (83)
N/A	1 (3)	1 (3)
Device, n (%)		
C5–2	0 (0)	0 (0)
C6–2	27 (70)	7 (19.44)
S4–2	12 (30)	15 (41.67)
SP5–1	0 (0)	13 (36.11)
N/A	0 (0)	2 (2.78)
Probe Type, n (%)		
Mindray	0 (0)	14 (38.89)
Philips	39 (100)	22 (58.33)
Zonaire	0 (0)	1 (2.78)
N/A	0 (0)	0 (0)

Table 2:

Detailed breakdown of the convolutional network architecture within YoloV3

Block ID	Number of Repetitions of Block	Layer Types in Order	Number of Channels	Size	Output Size
I	1	Convolutional	32	3 × 3 filter with stride 1	256 × 256
II	1	Convolutional	64	3 × 3 filter with stride 2	128 × 128
III	1	Convolutional	32	1 × 1 filter with stride 1	
		Convolutional	64	3 × 3 filter with stride 1	
		Residual			128 × 128
IV	1	Convolutional	128	3 × 3 filter with stride 2	64 × 64
V	2	Convolutional	64	1 × 1 filter with stride 1	
		Convolutional	128	3 × 3 filter with stride 1	
		Residual			64 × 64
VI	1	Convolutional	256	3 × 3 filter with stride 2	32 × 32
VII	8	Convolutional	128	1 × 1 filter with stride 1	
		Convolutional	256	3 × 3 filter with stride 1	
		Residual			32 × 32
VIII	1	Convolutional	512	3 × 3 filter with stride 2	16 × 16
IV	8	Convolutional	256	1 × 1 filter with stride 1	
		Convolutional	512	3 × 3 filter with stride 1	
		Residual			16 × 16
X	1	Convolutional	1024	3 × 3 filter with stride 2	8 × 8
XI	4	Convolutional	512	1 × 1 filter with stride 1	
		Convolutional	1024	3 × 3 filter with stride 1	
		Residual			8 × 8
XII	1	Average Pooling			
		Fully-connected		1000 neurons	
		Softmax			

Table 3:

Pericardial effusion detection and localization performances via YoloV3

Metric	Value	95% CI
Sensitivity	89%	56% to 98%
Specificity	92%	77% to 99%
Positive Likelihood Ratio	11.13	2.4 to 98
Negative Likelihood Ratio	0.12	0.02 to 0.57
Positive Predictive Value	92%	59% to 98%
Negative Predictive Value	90%	75% to 99%
Accuracy	91%	73% to 99%
Area Under the Curve (AUC)	94%	75% to 99%
Intersection over Union (IOU)	51%	

Table 4:

Classification and localization performances of YoloV3 vs. U-Net, MaskRCNN and SSD

	Specificity	Sensitivity	Accuracy	AUC	IOU
YoloV3	92%	89%	91%	94%	51%
U-Net	79%	62%	74%	74%	41%
MaskRCNN	56%	38%	47%	37%	83%
SSD	86%	79%	83%	82%	82%