

Article

Using AI Segmentation Models to Improve Foreign Body Detection and Triage from Ultrasound Images

Lawrence Holland, Sofia I. Hernandez Torres and Eric J. Snider * 

Organ Support and Automation Technologies Group, U.S. Army Institute of Surgical Research, JBSA Fort Sam Houston, San Antonio, TX 78234, USA; lawrence.a.holland7.ctr@health.mil (L.H.); sofia.i.hernandeztorres.ctr@health.mil (S.I.H.T.)

* Correspondence: eric.j.snider3.civ@health.mil; Tel.: +1-210-539-8721

Abstract: Medical imaging can be a critical tool for triaging casualties in trauma situations. In remote or military medicine scenarios, triage is essential for identifying how to use limited resources or prioritize evacuation for the most serious cases. Ultrasound imaging, while portable and often available near the point of injury, can only be used for triage if images are properly acquired, interpreted, and objectively triage scored. Here, we detail how AI segmentation models can be used for improving image interpretation and objective triage evaluation for a medical application focused on foreign bodies embedded in tissues at variable distances from critical neurovascular features. Ultrasound images previously collected in a tissue phantom with or without neurovascular features were labeled with ground truth masks. These image sets were used to train two different segmentation AI frameworks: YOLOv7 and U-Net segmentation models. Overall, both approaches were successful in identifying shrapnel in the image set, with U-Net outperforming YOLOv7 for single-class segmentation. Both segmentation models were also evaluated with a more complex image set containing shrapnel, artery, vein, and nerve features. YOLOv7 obtained higher precision scores across multiple classes whereas U-Net achieved higher recall scores. Using each AI model, a triage distance metric was adapted to measure the proximity of shrapnel to the nearest neurovascular feature, with U-Net more closely mirroring the triage distances measured from ground truth labels. Overall, the segmentation AI models were successful in detecting shrapnel in ultrasound images and could allow for improved injury triage in emergency medicine scenarios.

Keywords: artificial intelligence; segmentation; triage; ultrasound imaging; shrapnel; foreign body; emergency medicine



Citation: Holland, L.; Hernandez Torres, S.I.; Snider, E.J. Using AI Segmentation Models to Improve Foreign Body Detection and Triage from Ultrasound Images. *Bioengineering* **2024**, *11*, 128. <https://doi.org/10.3390/bioengineering11020128>

Academic Editor: Andrea Cataldo

Received: 12 December 2023

Revised: 20 January 2024

Accepted: 26 January 2024

Published: 29 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

It is critical in emergency medicine situations to adequately triage casualties to determine how to prioritize limited resources. This is amplified in remote or military medicine, where resources and medical evacuation resources are scarce. Future battlefields will likely involve more contested airspace, further limiting the opportunity for casualty evacuation [1,2]. Medical imaging-based triage is often limited to ultrasound imaging in remote austere environments due to its small size and portability compared to other medical imaging approaches [3]. Ultrasound imaging can allow for the detection of injuries or illnesses such as shrapnel or foreign bodies embedded in tissue [4], COVID-19 lung pathologies [5], pneumothorax, hemothorax, or abdominal hemorrhage [6], which all may be critical to detect during casualty triage.

There are, however, a few challenges with ultrasound-based triage. First, proper image acquisition requires extensive training for the proper positioning of the ultrasound transducer so that a suitable image is acquired. Multiple approaches have been taken to improve ultrasound image acquisition, which include improved standards and protocols, augmented reality overlays while scanning [7], and using robotics to automate image

acquisition [8]. Second, image interpretation requires a skilled radiographer to interpret the ultrasound image and make a diagnosis. This type of professional may often not be available in remote or military environments. Third, even with proper image interpretation, there is still a need for objective metrics on which to base triage decisions. Oftentimes these decisions are subjective, but they may be less consistent if taken by less skilled operators in remote emergency scenarios.

Artificial intelligence (AI) models for image interpretation and the subsequent analysis of these models can serve as a basis for simplifying image interpretation and developing objective triage metrics. Deep learning approaches have been extensively applied to ultrasound imaging to automate the detection of various injuries and abnormalities. However, deep learning image interpretation techniques can take different forms. Simple image classification approaches provide a categorical description of the ultrasound image; however, information about regions of interest in the image is not provided, which may be critical for adding granularity to the triage process (Figure 1A). Object detection models provide a higher level of explainability to model predictions by generating bounding boxes around the objects of interest present in an image, such as shrapnel (Figure 1B). We were recently able to use various object detection architectures to identify shrapnel and neurovascular features and use the predictions to measure distances between objects as a preliminary triage metric [9,10]. However, bounding boxes are rectangular shapes that may not properly delineate features at certain angles and may overlook irregular shapes, lessening the accuracy of making triaging predictions from bounding boxes. Instead, different deep learning architectures have been developed for creating segmentation masks for objects of any shape or size (Figure 1C). This allows for higher pixel-wise accuracy when recognizing objects from which triage decisions can be made.

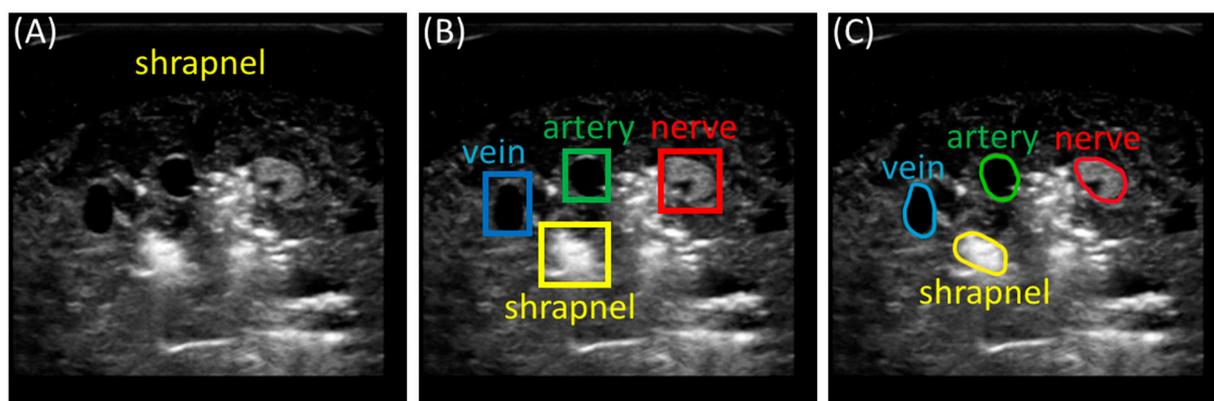


Figure 1. Comparison of AI model types for shrapnel and neurovascular feature identification. Representative ultrasound image showing output differences between a (A) binary classification model, (B) a multi-class object detection model, and (C) a multi-class segmentation model. The classification model labels images as either positive or negative for shrapnel. The object detection model makes predictions and places a bounding box around the predicted objects. The segmentation model creates an outline around the detected object.

In this work, we first evaluate the feasibility of developing deep learning segmentation models for identifying shrapnel in tissue phantoms as well as performing multi-class segmentation of shrapnel and neurovascular features. Shrapnel or foreign body detection in tissues is especially relevant for improved triages, as data from the Soviet–Afghan war highlighted how 71% of casualties were able to be treated minimally, non-operatively minimizing the need for medical evacuations when these casualties could be properly classified at or near the point of injury [11]. Using the multi-class segmentation model, we develop and improve on a foreign body distance triage metric originally used with object detection AI models.

To accomplish these goals, this paper first reviews how segmentation models have been used for other medical applications. Next, the methods section details how the images were collected and processed as well as the AI model setup for this application. The results section first describes how single-class segmentation models can be used for shrapnel detection followed by expanding these models for multi-class detection of additional neurovascular features. An application is lastly shown to highlight how a multi-class segmentation model can track the distance of shrapnel to neurovascular features as a potential triage tool.

Overview of Ultrasound Imaging Artificial Intelligent Segmentation Models

Various edge detection or intensity thresholding methods have been used for medical image segmentation [12]. However, these methods are not fully automated or as robust as more recent AI approaches. Focusing on AI modalities for ultrasound imaging, different architectures have been used, with encoder–decoder neural network approaches being the most widely used, followed by fully convolutional network model architectures [13]. Encoder–decoder segmentation models are similar in structure to more widely used object detection architectures, relying on convolution models to encode the image and extract features, followed by a deconvolution model or decoder which restores the pixel locations for the extracted features.

The most widely used segmentation model for medical imaging is the U-Net encoder–decoder model [14]. U-Net’s name is based on its U-shaped network architecture, which comprises a series of encoder steps to extract features, followed by a series of decoder blocks connected to each encoder by a bridge. The decoder blocks are used to align the extracted features with image pixels and generate segmentation masks. The bridge connections between each encoder and decoder help provide clarity for features identified at each step that may be lost at deeper convolutional steps to improve the accuracy of the segmentation masks. U-Net was developed specifically for biomedical image segmentation, initially for the segmentation of neurons and cell structures in microscope images [14]. It has since been utilized for a wide array of biomedical segmentation tasks including blood vessel, tumor, and organ segmentation [12,13,15].

Another widely used architecture for image segmentation is the YOLO (You Only Look Once) series of architectures [16]. YOLO models are single-stage object detectors made up of a convolutional backbone, where features are extracted, a neck, where features are mixed and combined, and a head, where locations and classes of objects are predicted. One of the latest models in this series is YOLOv7 [17]. YOLOv7 improves upon previous architectures by updating the backbone to include Extended Efficient Layer Aggregation (E-ELAN). This and several other improvements have served to make the model both more accurate and faster than previous YOLO releases. While YOLO is traditionally an object detection framework, YOLOv7 has been modified for instance segmentation use cases [17]. This architecture has been used to develop segmentation models for self-driving cars [18] as well as some medical image segmentation applications [19,20].

Image segmentation AI models can have many benefits for medical image interpretation, aiding with the recognition of regions of interest which could impact treatment and diagnosis [21,22]. As mentioned previously, U-Net was purpose-built for medical imaging. In certain emergent conditions the prompt interpretation of medical imaging can be critical, particularly when highly specialized personnel are not available. For this reason, U-Net architecture has been used to identify pleural effusion [23], measure the optic nerve sheath [24], and diagnose ascites [25]. The common use of U-NET for the segmentation of biomedical imaging has led to architecture modification that better fit specific areas and imaging modalities [26,27]. Along this line, segmentation architectures such as YOLOv7 have been used to achieve a reliable performance for the diagnostic segmentation of fetal cardiac features [28], kidney stones [29], and cervical lymph nodes [30].

2. Materials and Methods

2.1. Phantom Preparation and Imaging

The shrapnel image dataset was collected from an ultrasound tissue phantom that has been extensively used to train multiple AI models [31]. Briefly, a double-layer gelatin phantom was made by dissolving gelatin 10.0% *w/v* (Thermo-Fisher, Waltham, MA, USA) and flour (0.25% *w/v* inner layer and 0.10% *w/v* outer layer, unbleached wheat flour, H-E-B, San Antonio, TX, USA) in an evaporated milk (Costco, Seattle, WA, USA) and water mixture at a 1:2 ratio, respectively. Fragments made of agarose (Fisher Scientific, Fair Lawn, NJ, USA) at 2.0% *w/v*, with varying flour concentrations of 0.10% and 1.0% *w/v*, were added to the inner layer of the phantom in order to add training noise between different phantoms. A total of six phantom subjects were created to generate the training data: half of the subjects were a simple phantom with no additional features, and the other half were complex phantoms with an additional neurovascular bundle. The neurovascular bundle consisted of three channels created with a biopsy punch, which were left hollow to simulate the vein and artery and filled with the gelatin mixture with flour at 0.5% *w/v* for the nerve channel. Shrapnel insertion and imaging of the phantom were performed under water in 10 s clips. A 2.0 mm diameter brass rod (McMaster Carr, Elmhurst, IL, USA) was cut into pieces of five different lengths (1.0 cm, 0.8 cm, 0.6 cm, 0.4 cm, and 0.2 cm), that were used as shrapnel. All the ultrasound images were collected using a Sonosite Edge II System (Fujifilm, Bothell, WA, USA) with the HL50x (Fujifilm, Bothell, WA, USA) transducer.

2.2. Preprocessing and Labeling Images

After the ultrasound videos had been recorded, individual frames were extracted as images, totaling 12,144 collected ultrasound images. Next, the images were manually annotated with the locations of neurovascular features and shrapnel, with 6385 receiving annotations only for shrapnel and the remaining 5759 being annotated with vein, artery, nerve, and shrapnel features. From these, 15 and 59 images were split to be used as holdout test images after the models were trained for the single-class and multi-class models, respectively. The annotations were generated manually using the COCO annotator software to label pixels corresponding to the desired features in each image. The annotations were then saved to a JSON file containing file and annotation information for each image.

Before the training of the models could start, Python scripts were developed to read the COCO JSON file and use it to produce images and masks compatible with the needs of the segmentation models. In the case of U-Net, this involved resizing all the images to 256×256 and generating image masks from the original COCO data. These masks consist of one-channel 256×256 images with each desired class represented by a unique integer value (such as vein = 1, artery = 2, etc.). For YOLOv7 training, the images were kept in their original 512×512 size, and the COCO annotation data were used to generate text files for each image that contain segmentation information for each class.

2.3. Training Segmentation Algorithms

All training was conducted using the same computer hardware setup: an NVIDIA GeForce RTX 3090 Ti 24 GB VRAM system with Intel i9-12900k and 128 GB RAM with a Windows 10 Pro operating system. YOLOv7 training was completed using an Anaconda distribution running Python 3.8.8. U-Net training was conducted using MATLAB (R2023a, Mathworks, Natick, MA, USA) with deep learning, computer vision, and image processing packages installed. Before training, the images were randomly shuffled and separated into training and validation sets at an 80:20 ratio and paired with their respective masks.

A publicly available version of the YOLOv7 segmentation model was used [17]. Minor changes were made to the training and prediction scripts of the original code to be able to extract generated prediction masks into text files for each image per YOLOv7 requirements. The ultrasound images were paired with their respective segmentation text file containing ground truth information prior to training. The model was trained for a total of ten epochs, recording precision, recall, loss, and mean average precision for each epoch.

The U-Net model was implemented in MATLAB, which provides a ready-made 4-depth U-Net for computer vision tasks. The customizations for our purposes included removing the final layer of the model and replacing it with a pixel classification layer; batch normalization layers were added after each convolution. Batch normalization works by normalizing the output of each convolutional layer, allowing for a deeper and faster running model and helping to prevent overfitting by inserting a small amount of noise into the data. The original U-Net publication included max pooling layers after each convolution layer. However, these were not present in the MATLAB implementation of U-Net. The U-Net model was trained for a total of twenty epochs, with accuracy and loss values being recorded after processing each batch of images. Similar training approaches were taken with both architectures when training a single-class model for shrapnel and the multi-class models further including neurovascular classes.

2.4. Performance Metrics for Evaluating Trained Models

For each of the trained models, prediction segmentation masks were generated from a separate set of blind images. From these, every pixel in the prediction mask was compared to the ground truth masks, and labeled as either True Positive (TP), False Positive (FP), True Negative (TN), or False Negative (FN). The total number of each label was compiled for each class as well as for the overall image and used to calculate performance metrics including accuracy, precision, recall, the intersection-over-union (IOU) score, and the Dice coefficient. Formulas for all the metrics used are shown in Table 1. Accuracy is a measure of how frequently a model can be expected to make correct predictions. Precision measures how often positive predictions are correct, while recall shows the efficiency of a model to accurately predict positive features. The IOU score and Dice coefficient are common metrics in object detection and segmentation models that determine how well the model prediction masks correspond to the ground truth masks.

Table 1. Formulas for performance metrics. Each performance metric is calculated using True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN) predictions.

| Metric | Formula |
|------------------|---------------------------------------|
| Accuracy | $(TP + TN)/(TP + FN + TN + FP)$ |
| Precision | $TP/(FP + TP)$ |
| Recall | $TP/(FN + TP)$ |
| IOU | $TP/(TP + FP + FN)$ |
| Dice Coefficient | $2 \times TP/(2 \times TP + FP + FN)$ |

2.5. Distance Triage

The distance between shrapnel and neurovascular bundle anatomical features was calculated as it could be a useful metric for triaging injury severity. This analysis was performed using segmentation masks generated from the predictive models. The perimeter of each object (shrapnel, nerve, artery, and vein) was found after isolating each object.

Distance triage is calculated from the predicted shrapnel to the predicted vein, artery, or nerve, and then the distance from the shrapnel to the closest structure is provided. This was performed using MATLAB's regionprops functionality to find the closest distance of shrapnel to each neurovascular feature, as well as combinations of multiple classes if multiple shrapnel or nerves were identified. This was achieved by comparing each pixel in the boundary of each anatomical feature to every pixel value in the boundary of the shrapnel segmentation mask to find the two closest points between the two (given as a pair of X,Y coordinates). After these were identified, the minimum of all the options was selected, which served as the triage score for this application (Figure 2). Multi-class segmentation masks ($N = 46$) generated by the AI models were compared to their respective ground truth masks to measure the accuracy of each model for triage. The outlier exclusion

ROUT (robust regression followed by outlier identification) methodology was applied to the measured distance error to remove data points with a false discovery rate set at 1%. Statistical analysis and figures were generated using GraphPad Prism 10.1.2 (La Jolla, CA, USA).

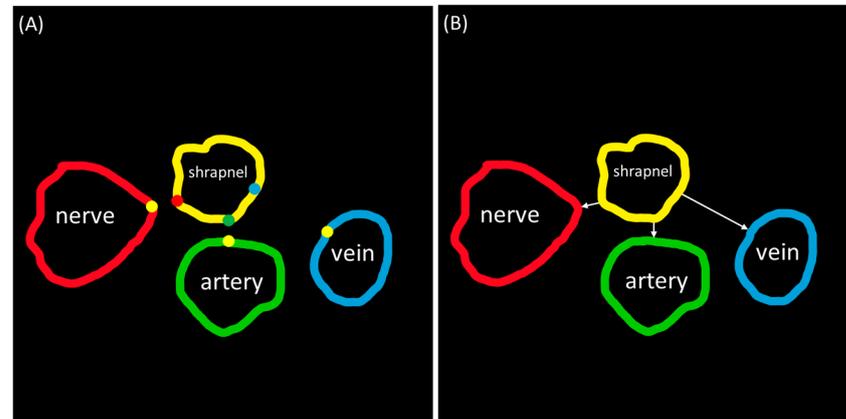


Figure 2. Diagram of triage metric calculation. (A) Each pixel from the boundary of each vein (blue), artery (green), and nerve (red) are compared to boundary pixels of the shrapnel (yellow) to find the closest two points (given as a pair of X,Y coordinates), with the closest distance pairs represented by the different dots. (B) The distance between the closest point on each feature and the shrapnel are calculated using the distance formula, shown as white arrows from shrapnel mask.

3. Results

3.1. Single-Class Segmentation Models for Shrapnel

For shrapnel segmentation, the YOLOv7 model achieved an 82.2% validation precision in nine training epochs, whereas U-Net reached a validation accuracy of over 90% in the first epoch and continued to increase steadily throughout training, reaching a final validation accuracy of 99%. From the test set for the single-class shrapnel segmentation models, the accuracy, precision, recall, IOU, and Dice coefficient scores were calculated by comparing the predicted shrapnel masks to the ground truth masks (Table 2). Both models performed alike in identifying shrapnel in the test images, with both receiving an accuracy of 99.1%. Both models also received similar scores in all the other recorded metrics, with YOLOv7 slightly outscoring or equaling U-Net in all the metrics. Representative results of both models on the same ultrasound image are shown in Figure 3.

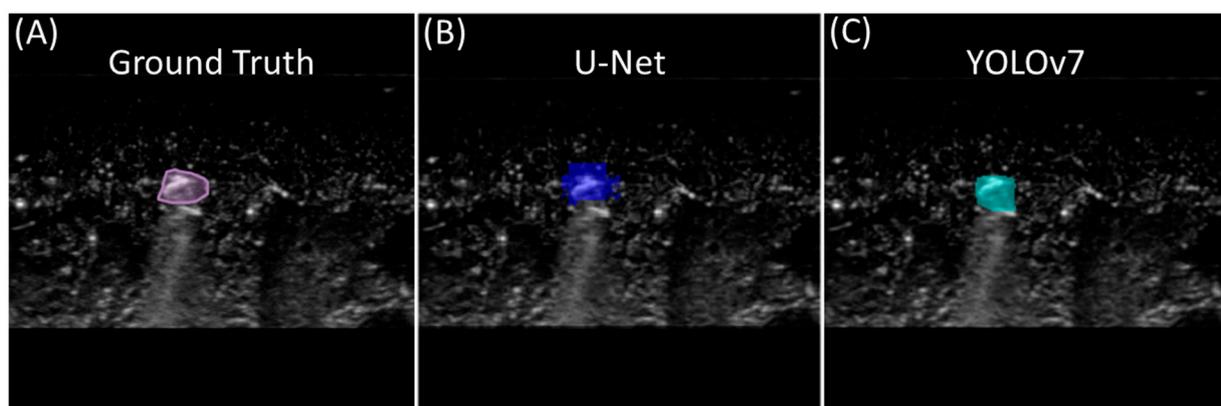


Figure 3. Single-class segmentation for shrapnel. Representative image showing (A) ground truth mask, (B) U-Net-model-generated mask, and (C) YOLOv7-model-generated mask for the same shrapnel image.

Table 2. Summary of performance metrics for single-class shrapnel segmentation.

| | Accuracy | Precision | Recall | IOU | Dice Coefficient |
|---------------|----------|-----------|--------|-------|------------------|
| YOLOv7 | 99.1% | 76.0% | 74.3% | 55.1% | 69.7% |
| U-Net | 99.1% | 63.8% | 67.7% | 51.5% | 64.7% |

3.2. Multi-Class Segmentation Models

The YOLO model was trained for ten epochs, reaching a best precision of 84.8% and recall of 84.2%. The U-net model reached an accuracy value over 90% in the fourth epoch, steadily increasing after reaching a final accuracy of 96.5% at the end of twenty training epochs. By visual comparison of the predicted masks for both models to the ground truth masks, the U-Net model produced generally larger segmentation masks for each feature than the YOLOv7 model, with the size of the ground truth masks falling in-between the two (Figure 4). In some instances, the U-Net model would produce segmentation masks with two or more class prediction masks in contact with one another or even overlapping.

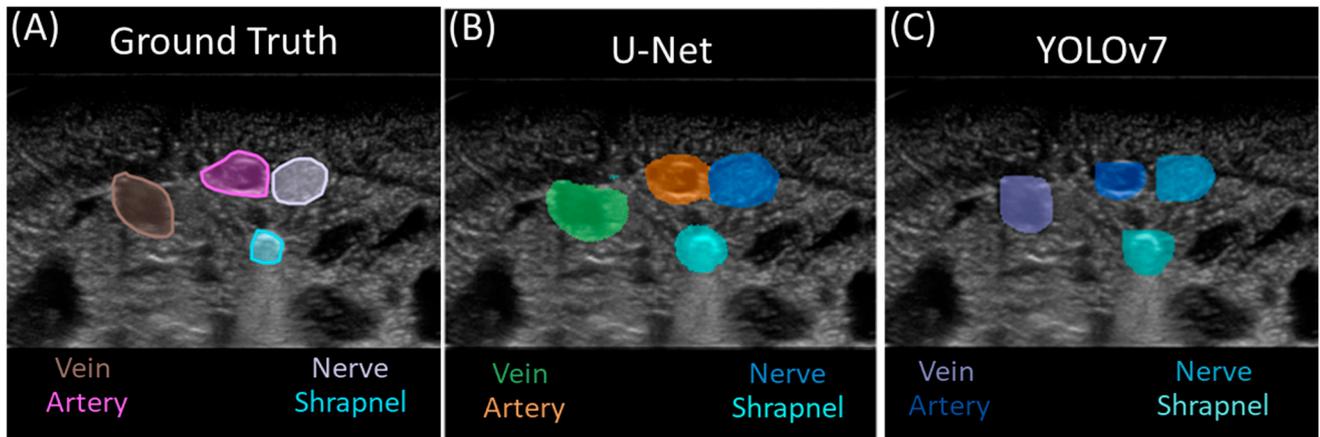


Figure 4. Multi-class segmentation for shrapnel. Representative image showing (A) ground truth mask, (B) U-Net-model-generated mask, and (C) YOLOv7-model-generated mask for the same ultrasound image. All classes are identified in each image.

To assess the performance of both models, metrics were calculated for each prediction class and image background classification, with artery features having the highest test accuracy at more than 99.0% for each model (Table 3). Averaging across all the features, both models achieved accuracy scores above 98.0% for all classes (Figure 5). However, recall was much higher for U-Net than YOLOv7 at 92.6% vs. 82.7%, in contrast to precision, which was higher for YOLOv7 than U-Net, at 77.7% and 66.2%, respectively.

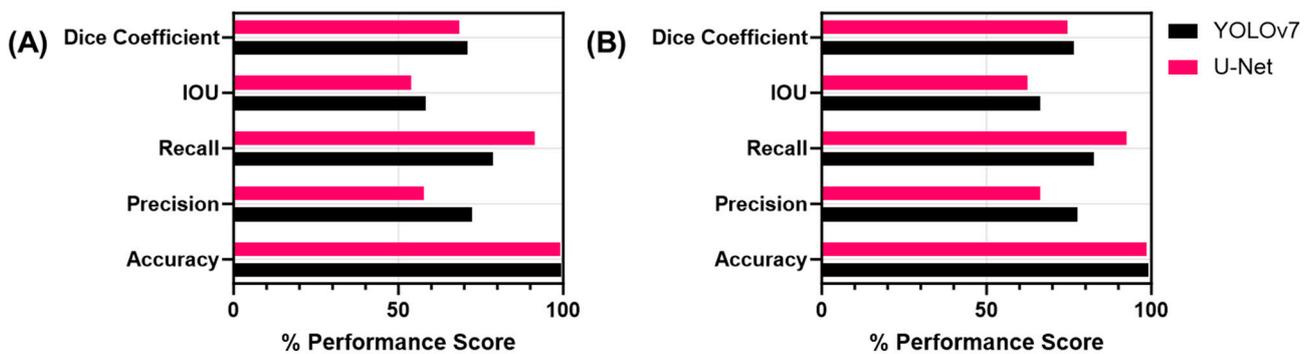


Figure 5. Summary of performance metrics for multi-class segmentation. (A) Average metrics for vein, artery, nerve, and shrapnel features and (B) performance metrics including the background.

Table 3. Summary of performance metrics by class for the segmentation models.

| YOLOv7 | | | | | |
|-------------------|----------|-----------|--------|-------|------------------|
| | Accuracy | Precision | Recall | IOU | Dice Coefficient |
| Vein | 99.4% | 75.9% | 76.2% | 61.6% | 74.0% |
| Artery | 99.5% | 79.2% | 78.3% | 64.0% | 76.8% |
| Nerve | 99.4% | 76.9% | 81.4% | 61.9% | 74.9% |
| Shrapnel | 99.2% | 57.5% | 78.8% | 45.7% | 58.6% |
| Background | 97.8% | 99.0% | 98.7% | 97.7% | 98.8% |
| U-Net | | | | | |
| | Accuracy | Precision | Recall | IOU | Dice Coefficient |
| Vein | 99.2% | 60.5% | 85.2% | 54.7% | 69.9% |
| Artery | 99.4% | 72.1% | 84.1% | 63.7% | 77.1% |
| Nerve | 98.9% | 55.2% | 97.4% | 54.3% | 69.5% |
| Shrapnel | 98.9% | 43.7% | 84.1% | 43.5% | 58.2% |
| Background | 97.1% | 99.8% | 97.1% | 96.9% | 98.5% |

In addition, we evaluated performance excluding the image background class as it was the most prevalent feature in all the images. Focusing on only the features, YOLOv7 outscored U-Net for test precision, with a difference of 14.6% between the two models. Conversely, U-Net outperformed YOLOv7 with regard to test recall, with a difference of 12.8%. Both models received similar scores for the IOU and Dice coefficient, with YOLOv7 slightly outscoring U-Net for the IOU by 4.3% and the Dice coefficient by 2.5%.

In contrast to the single-class shrapnel model, the U-Net multi-class model precision fell by a value of 0.201 while recall increased by 16.4% for the shrapnel predictions. The IOU and Dice coefficient metrics decreased slightly between shrapnel-only and multi-class U-Net by 8.0% and 6.5%, respectively. YOLOv7 saw similar trends, with precision falling by 18.5% and recall increasing by 4.5%. The accuracy and IOU saw almost no change, while the Dice coefficient fell by 11.1% for the YOLOv7 multi-class model.

3.3. Triage Metric for Segmentation

For the triage metric, minimum distances between each feature and shrapnel were calculated for all the test images, using both the ground truth and prediction masks (Figure 6). For the ground truth labels, the average minimum distance between the shrapnel and the closest anatomical feature was found to be 10.2 pixels, with a standard deviation of 10.8 pixels (Figure 7). The average minimum distances for both U-Net and YOLOv7 were 3.8 and 9.7 pixels, respectively. For each predicted image, the minimum distance triage metric was compared to its corresponding ground truth label, with the difference between the two distances (difference) and the percentage error (% error) being calculated for each image. These values were then averaged across all the test images to obtain a broad view of the distance triage performance. The average difference for the minimum distances between the U-Net predictions and the ground truth for all the images was 7.7 pixels and an average error of 72.5%. For the YOLOv7 predictions, a 5.1-pixel difference was calculated at an average error of 45.5% (Figure 7).

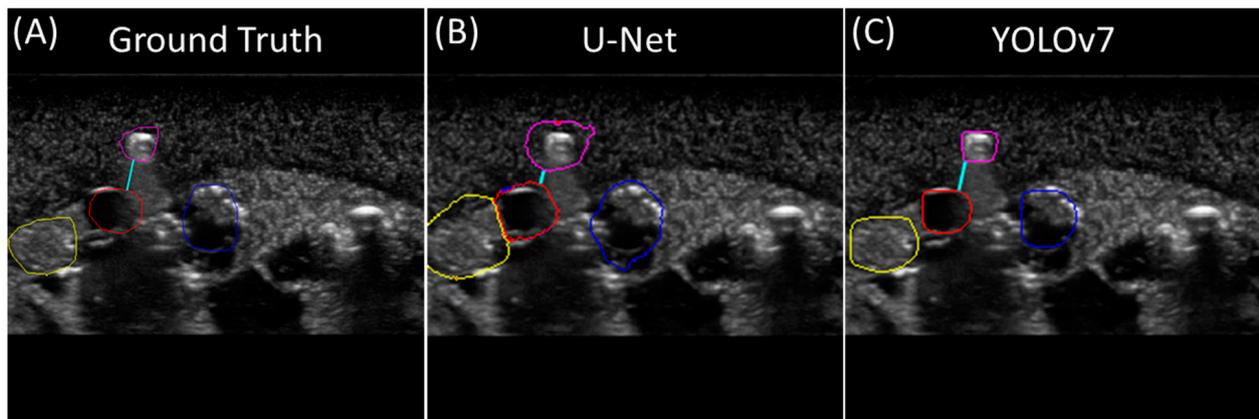


Figure 6. Distance triage metric results. Representative image showing (A) ground truth mask, (B) U-Net-model-generated mask, and (C) YOLOv7-model-generated mask for the same ultrasound image. The light blue line shown indicates the generated minimum distance between the shrapnel (shown in pink) and the closest anatomical feature: nerve (yellow), artery (red), or vein (dark blue).

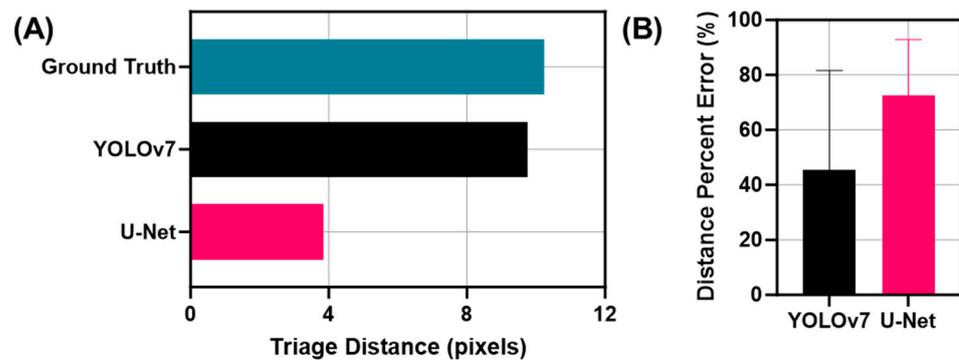


Figure 7. Summary of triage metrics for multi-class segmentation. (A) Average minimum triage distance measurements for ground truth, YOLOv7, and U-Net ($N = 46$ images). (B) Percent error calculations relative to the ground truth distances for YOLOv7 ($N = 41$ after outliers removed) and U-Net ($N = 38$ after outliers removed).

4. Discussion

Given the small form factor, low power requirements, and portability of ultrasound technology, it can be a valuable tool for providing more objective criteria for medical triage at or near the point of injury. This is especially critical in remote medical scenarios where resources are limited and medical evacuation can only be made available to the most severe injuries, such as during combat casualty care. However, ultrasound imaging is only valuable if the images can be easily and objectively interpreted for making medical triage decisions. The artificial intelligent segmentation models proposed here for identifying shrapnel and neurovascular features highlight how AI technology can help standardize medical decision making.

For this effort, we evaluated the use of two segmentation models—YOLOv7 segmentation and U-Net—for tracking shrapnel only or multi-class applications including vein, artery, and nerve features. Both models were successful in identifying shrapnel and neurovascular features using the thousands of training images collected in a custom-made ultrasound tissue phantom. However, there were performance differences between the two models, notably with precision and recall. Precision, a measurement of the accuracy of positive predictions, was significantly higher with YOLOv7, while recall, a measurement of accurately identifying positive features, was higher with U-Net. This was mostly due to the prediction masks around the identified features in the YOLOv7 model were fitted around the object perimeter, resulting in more false-negative pixels but few false-positive

pixels, while U-Net generally had larger prediction masks resulting in an opposite false bias. Which bias is preferred for medical triage tasks such as this one is unknown. The smaller YOLOv7 prediction masks would improve accuracy in identifying the exact shape of the shrapnel object, which could be critically important for aiding surgical removal, while U-Net would provide a more cautious triage outlook by overestimating the closeness of the shrapnel to neurovascular features (triage distances of 3.8 and 9.7 pixels for U-Net and YOLOv7, respectively). However, either segmentation model likely provides a much higher fidelity for the determination of distances between shrapnel and neurovascular features than the object detection models we previously used for this application. Bounding boxes may be inferior for labeling the irregularly shaped geometries that can often be expected with foreign-body objects embedded in tissues.

It is worth noting that the determination of model performance for both AI models is highly dependent on the training dataset. While attempts were made to add subject variability to the phantom creation process, the tissue phantom used to collect the shrapnel image dataset lacks tissue-level organization, as seen in animal tissue. In addition, the ground truth labels are biased by the image labeler, which can be prone to error. Given the quantity of images being labeled (over 20k images), only a single label was drawn for each US image. For a more robust training set, multiple labelers should be pooled to remove labeling subjectivity from the process. This is especially critical in the case of YOLOv7's performance, which oftentimes focused on accurately defining the edge of labeled features compared to the ground truth mask. Additional US images will be needed to expand on the training of these segmentation models prior to real-time deployment for use with animal or human triage processing. The translation of this effort will require engagement with the FDA to ensure that the regulatory concerns associated with AI technologies are properly addressed.

5. Conclusions

In summary, medical imaging-based triage is a critical tool for accurate decision making regarding injury severity and evacuation priorities if the images can be properly interpreted. Image interpretation challenges are compounded in remote emergency medical situations, such as combat casualty care, due to limited medical personnel and resource availability. The segmentation AI models developed in this work were successful in identifying both foreign-body objects and neurovascular features. This allowed for the measurement of the proximity of the shrapnel to artery, vein, or nerve features as a means of objectifying the triage criteria for this application. Further work is needed to train the AI models using *in-vivo* medical images rather than *in-vitro*, and the triage framework can be expanded to other applications of relevance to emergency medicine. This will improve the prioritization of limited resources at or near the point of injury and improve evacuation priorities during remote emergency medical situations.

Author Contributions: Conceptualization, E.J.S. and S.I.H.T.; methodology, E.J.S., S.I.H.T. and L.H.; data curation, S.I.H.T. and L.H.; writing—original draft preparation, E.J.S., S.I.H.T. and L.H.; writing—review and editing, E.J.S., S.I.H.T. and L.H.; visualization, S.I.H.T. and L.H.; supervision, E.J.S.; funding acquisition, E.J.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the U.S. Army Medical Research and Development Command (IS220007). This project was supported in part by an appointment to the Science Education Programs at National Institutes of Health (NIH), administered by ORAU through the U.S. Department of Energy Oak Ridge Institute for Science and Education (L.H.).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Acknowledgments: The authors would like to acknowledge Catherine Thorpe and Jeanette Rocha who assisted with the ground truth labelling of the ultrasound images that were used for training the segmentation AI models.

Conflicts of Interest: The authors declare no conflicts of interest.

DoD Disclaimer: The views expressed in this article are those of the authors and do not reflect the official policy or position of the U.S. Army Medical Department, Department of the Army, DoD, or the U.S. Government.

References

1. Townsend, S.; Lasher, W. *The U.S. Army in Multi-Domain Operations 2028*; U.S. Army: Arlington County, VA, USA, 2018.
2. Epstein, A.; Lim, R.; Johannigman, J.; Fox, C.J.; Inaba, K.; Vercruyssen, G.A.; Thomas, R.W.; Martin, M.J.; Konstantyn, G.; Schwaartzberg, S.D.; et al. Putting Medical Boots on the Ground: Lessons from the War in Ukraine and Applications for Future Conflict with Near-Peer Adversaries. *J. Am. Coll. Surg.* **2023**, *237*, 364–373. [[CrossRef](#)]
3. Harper, H.; Myers, M. Military and Tactical Ultrasound. In *Emergency Ultrasound*; ACEP: Irving, TX, USA, 2008.
4. Snider, E.J.; Hernandez-Torres, S.I.; Boice, E.N. An Image Classification Deep-Learning Algorithm for Shrapnel Detection from Ultrasound Images. *Sci. Rep.* **2022**, *12*, 8427. [[CrossRef](#)]
5. Diaz-Escobar, J.; Ordóñez-Guillén, N.E.; Villarreal-Reyes, S.; Galaviz-Mosqueda, A.; Kober, V.; Rivera-Rodriguez, R.; Rizk, J.E.L. Deep-Learning Based Detection of COVID-19 Using Lung Ultrasound Imagery. *PLoS ONE* **2021**, *16*, e0255886. [[CrossRef](#)] [[PubMed](#)]
6. Canelli, R.; Leo, M.; Mizelle, J.; Shrestha, G.S.; Patel, N.; Ortega, R. Use of eFAST in Patients with Injury to the Thorax or Abdomen. *N. Engl. J. Med.* **2022**, *386*, e23. [[CrossRef](#)]
7. Farshad-Amacker, N.A.; Bay, T.; Rosskopf, A.B.; Spirig, J.M.; Wanivenhaus, F.; Pfirmann, C.W.A.; Farshad, M. Ultrasound-Guided Interventions with Augmented Reality in Situ Visualisation: A Proof-of-Mechanism Phantom Study. *Eur. Radiol. Exp.* **2020**, *4*, 7. [[CrossRef](#)] [[PubMed](#)]
8. Abolmaesumi, P.; Salcudean, S.E.; Zhu, W.H.; DiMaio, S.P.; Siroospour, M.R. A User Interface for Robot-Assisted Diagnostic Ultrasound. In Proceedings of the Proceedings 2001 ICRA, IEEE International Conference on Robotics and Automation (Cat. No.01CH37164), Seoul, Republic of Korea, 21–26 May 2001; Volume 2, pp. 1549–1554.
9. Snider, E.J.; Hernandez-Torres, S.I.; Avital, G.; Boice, E.N. Evaluation of an Object Detection Algorithm for Shrapnel and Development of a Triage Tool to Determine Injury Severity. *J. Imaging* **2022**, *8*, 252. [[CrossRef](#)]
10. Hernandez-Torres, S.I.; Hennessey, R.P.; Snider, E.J. Performance Comparison of Object Detection Networks for Shrapnel Identification in Ultrasound Images. *Bioengineering* **2023**, *10*, 807. [[CrossRef](#)] [[PubMed](#)]
11. Bowyer, G.W. Management of Small Fragment Wounds: Experience from the Afghan Border. *J. Trauma* **1996**, *40*, S170–S172. [[CrossRef](#)]
12. Patil, D.D.; Deore, S.G. Medical Image Segmentation: A Review. *Int. J. Comput. Sci. Mob. Comput.* **2013**, *2*, 22–27.
13. Wang, Z. Deep Learning in Medical Ultrasound Image Segmentation: A Review. *arXiv* **2020**, arXiv:2002.07703.
14. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; Springer: Cham, Switzerland, 2015; pp. 234–241.
15. Ho, T.-W.; Qi, H.; Lai, F.; Xiao, F.-R.; Wu, J.-M. Brain Tumor Segmentation Using U-Net and Edge Contour Enhancement. In Proceedings of the 2019 3rd International Conference on Digital Signal Processing, Jeju Island, Republic of Korea, 24–26 February 2019; pp. 75–79.
16. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. Available online: <https://arxiv.org/abs/1506.02640v5> (accessed on 20 March 2023).
17. Wang, C.-Y.; Bochkovskiy, A.; Liao, H.-Y.M. YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors. Available online: <https://arxiv.org/abs/2207.02696v1> (accessed on 20 March 2023).
18. Yildirim, E.; Sefercik, U.G.; Kavzoglu, T. Automated Vehicle Detection and Instance Segmentation from High-Resolution UAV Imagery Using YOLOv7 Model. *Intercont. Geoinf. Days* **2022**, *5*, 116–119.
19. Ahmad, S.; Kim, J.-S.; Park, D.K.; Whangbo, T. Automated Detection of Gastric Lesions in Endoscopic Images by Leveraging Attention-Based YOLOv7. *IEEE Access* **2023**, *11*, 87166–87177. [[CrossRef](#)]
20. Liu, Z.; Zheng, L.; Gu, L.; Yang, S.; Zhong, Z.; Zhang, G. InstrumentNet: An Integrated Model for Real-Time Segmentation of Intracranial Surgical Instruments. *Comput. Biol. Med.* **2023**, *166*, 107565. [[CrossRef](#)]
21. Azad, R.; Aghdam, E.K.; Rauland, A.; Jia, Y.; Avval, A.H.; Bozorgpour, A.; Karimjafarbigloo, S.; Cohen, J.P.; Adeli, E.; Merhof, D. Medical Image Segmentation Review: The Success of U-Net. *arXiv* **2022**, arXiv:2211.14830.
22. Siddique, N.; Paheding, S.; Elkin, C.P.; Devabhaktuni, V. U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications. *IEEE Access* **2021**, *9*, 82031–82057. [[CrossRef](#)]
23. Huang, L.; Lin, Y.; Cao, P.; Zou, X.; Qin, Q.; Lin, Z.; Liang, F.; Li, Z. Automated Detection and Segmentation of Pleural Effusion on Ultrasound Images Using an Attention U-Net. *J. Appl. Clin. Med. Phys.* **2024**, *25*, e14231. [[CrossRef](#)] [[PubMed](#)]

24. Pang, M.; Liu, S.; Lin, F.; Liu, S.; Tian, B.; Yang, W.; Chen, X. Measurement of Optic Nerve Sheath on Ocular Ultrasound Image Based on Segmentation by CNN. In Proceedings of the 2019 IEEE International Conference on Signal, Information and Data Processing (ICSIDP), Chongqing, China, 11–13 December 2019; pp. 1–5.
25. Lin, Z.; Li, Z.; Cao, P.; Lin, Y.; Liang, F.; He, J.; Huang, L. Deep Learning for Emergency Ascites Diagnosis Using Ultrasonography Images. *J. Appl. Clin. Med. Phys.* **2022**, *23*, e13695. [[CrossRef](#)] [[PubMed](#)]
26. Li, H.; Fang, J.; Liu, S.; Liang, X.; Yang, X.; Mai, Z.; Van, M.T.; Wang, T.; Chen, Z.; Ni, D. CR-Unet: A Composite Network for Ovary and Follicle Segmentation in Ultrasound Images. *IEEE J. Biomed. Health Inform.* **2020**, *24*, 974–983. [[CrossRef](#)]
27. NAS-Unet: Neural Architecture Search for Medical Image Segmentation | IEEE Journals & Magazine | IEEE Xplore. Available online: <https://ieeexplore.ieee.org/abstract/document/8681706> (accessed on 19 January 2024).
28. Iriani Sapitri, A.; Nurmaini, S.; Naufal Rachmatullah, M.; Tutuko, B.; Darmawahyuni, A.; Firdaus, F.; Rini, D.P.; Islami, A. Deep Learning-Based Real Time Detection for Cardiac Objects with Fetal Ultrasound Video. *Inform. Med. Unlocked* **2023**, *36*, 101150. [[CrossRef](#)]
29. Effective Kidney Stone Prediction Based on Optimized Yolov7 Segmentation and Deep Learning Classification | International Journal of Intelligent Systems and Applications in Engineering. Available online: <https://ijisae.org/index.php/IJISAE/article/view/3776> (accessed on 19 January 2024).
30. Ariji, Y.; Kise, Y.; Fukuda, M.; Kuwada, C.; Ariji, E. Segmentation of Metastatic Cervical Lymph Nodes from CT Images of Oral Cancers Using Deep-Learning Technology. *Dentomaxillofac. Radiol.* **2022**, *51*, 20210515. [[CrossRef](#)]
31. Hernandez-Torres, S.I.; Boice, E.N.; Snider, E.J. Using an Ultrasound Tissue Phantom Model for Hybrid Training of Deep Learning Models for Shrapnel Detection. *J. Imaging* **2022**, *8*, 270. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.