Usage of Artificial Intelligence in Gallbladder Segmentation to Diagnose Acute Cholecystitis

Benjamin Wu[1], Yucheng Liu[2], Meng Jou Wu[2], Hiram Shaish[3] and Hong Yun Ma[3]*

[1]NYU Stern School of Business, New York University, USA

[2]Department of Radiology, Division of physics, Columbia University Irving Medical Center, USA

[3]Department of Radiology, Division of Body, Columbia University Irving Medical Center, USA

*Corresponding author: Hong Yun Ma, Department of Radiology, Division of Body, Columbia University Irving Medical Center, New York, NY, USA

ARTICLE INFO

ABSTRACT

Acute Cholecystitis is a sudden inflammation of the gallbladder that affects hundreds of thousands of people per year. Though a common condition, methods of diagnosis still underperform modern standards of medicine. As such, there has been a demand to incorporate new innovations in the diagnostic process. In this study, a modified U-Net was trained to automatically segment gallbladder ultrasound images. Then SVM, KNN, and other machine learning methods were used to develop models to diagnose acute cholecystic. KNN had the highest performance, with an accuracy and sensitivity of 84.09%, and precision of 84.74%. This work incorporated novel 3D volume measurements of the gallbladder with comparable accuracy, precision, and sensitivity to similar 2D work, as well as statistics gathered from traditional diagnosis through manual ultrasound examination. Based on the current work, future improvements could entail more physical features as variables in the machine learning process and building more robust models of KNN and SVM for potential applications in a clinical setting.

Keywords: Acute Cholecystitis; Gallbladder Volume; Machine Learning; U-Net; Support Vector Machine; KNN; Naïve Bayes

Introduction

Cholelithiasis is currently the most common gallbladder condition with around 10% of patients with Cholelithiasis developing acute cholecystitis. In the United States, on average, acute cholecystitis affects 200,000 people annually [1]. While the pathogenesis and treatment of acute cholecystitis is well understood, methods of diagnosis are still complicated and require multiple physical symptoms and laboratory tests. Radiological scanning methods include Ultrasonography, CAT scan, and Hepatobiliary Iminodiacetic Acid scan. In addition to radiological scanning, physical and pathological symptoms such as fever, Murphy's sign, and elevated WBC count are also important for the diagnosis of acute cholecystitis. The Tokyo Guidelines, initially introduced in 2007, were presented as a method to streamline and standardize the diagnosis of acute cholecystitis. The guidelines create three different criteria, with two needed to be met to diagnose a patient with acute cholecystitis. These criteria include findings from sonographic imaging, localized inflammation of the upper right quadrant of the abdomen, and pathological findings. However, even with two revisions of the Tokyo Guidelines, success rates, specifically specificity still fluctuates drastically in various validation studies.

As such, there has been a demand to show the improved accuracy of ultrasonography in diagnosing acute cholecystitis. Common internal conditions often examined for include thickening of the gallbladder wall, presence of pericholecystic fluid within the gallbladder, and gallbladder distension. In particular, the measuring gallbladder distension, if given the correct parameters, can have high accuracy of diagnosis. A review, authored by Kiewet, et al. pooled 57 studies (5859 patients) and found that ultrasonography when used to diagnosis acute cholecystitis has a sensitivity of 81% and a specificity of 83% [2]. Previous novel research conducted to demonstrate the diagnostic power of ultrasound in acute cholecystitis includes the work done by Shaish, et al. who quantifiably defined the width of 3.5 cm (with 83%

sensitivity and 88% specificity) as a reasonable cutoff for considering the diagnosis of acute cholecystitis [3]. In the past decade, the usage of artificial intelligence (AI) in radiology has increased tremendously and been applied to the diagnosis of diseases in many different organs [4,5]. Chang, et al. were able to use various machine learning methods to detect differences between biomarkers of patients with and without Alzheimer's disease [6]. Meanwhile, Yoo et al. used deep learning to diagnose bladder tumors with results ≥ 98% for low grade benign tumors [7].

One of the most influential machine learning technologies is U-Net. U-Net was developed in 2015 by researchers at the University of Freiburg and is a convolutional neural network that is frequently used in biological image segmentation. This method proved to be accurate and efficient, with segmentation of a 512x512 image only taking around one second with a graphics processing unit [8]. However, there has still yet to be an application of AI in. Given the prevalence of acute cholecystitis and the high potential to inaccurately diagnosis this condition, acute cholecystitis serves as an important target for innovation. The goal of the current study is to use segmentation of 2D ultrasound to estimate total volume of the gallbladder to diagnose acute cholecystitis. U-Net segmentation, a novel machine learning method, was used to auto-segment ultrasound images of gallbladders and subsequently automatically measure the gallbladder length, width, and other variables. Support Vector Machines were applied to determine which features have highest predictive power and test different kernels to assess the accuracy of volume as a predictor of acute cholecystitis. Results in the U-Net segmentation showed up to 95.71% mean intersection over union (IOU) and 80.49% accuracy and sensitivity of KNN and SVM with radial basis function kernel, on the test cohort to diagnose acute cholecystitis.

Methods

Ultrasound Image Segmentation and Annotation by U-Net

The dataset used in this study contain a total patient size of 127. For each patient, one longitudinal and one transverse scan of the gallbladder was chosen at random and manually segmented to show the outline of the gallbladder excluding the gallbladder wall. The annotated images were set, and ground truths used to train a neural network to automatically annotate gallbladder ultrasound scans. Specifically, the research utilized a modified version of U-Net, a very commonly used convolutional neural network function designed for biological image segmentation. The modified U-Net was designed with one input layer, 42 hidden convolution, max-pooling, and concatenate layers, and one output layer with both the input and output being size [512, 512, 1]. Data was split into a 75% testing, 20% validation, to 5% testing ratio. The hyperparameter used in this experiment was learning rate ranging from 0.0003 to 0.001. The associated U-Net epoch numbers for each learning rate (number of iterations through the test data before model has reached sufficient predictive power) was based on the early stop method to reduce overfitting in the validation set. Throughout the training process, various optimizers were tested

to refine weights of nodes and epoch number. These optimizers include Adam, RMSprop, SGM, Nadam, and Adadelta, with only Adam, RMSprop, and Nadam functioning properly as optimizers. Given these optimizers, dice score and mean IOU were calculated to assess accuracy of the U-Net and the different applied optimizers.

Extraction of Gallbladder Features

Automatically segmented gallbladder ultrasound scans were used to calculate the volume of the gallbladder. A script was created to measure the longest axis of both the longitudinal and transverse scans of a patient using the annotation's coordinates. In total, nine features were recorded for each pair of scans. These were: longitudinal long axis length, longitudinal short axis length, longitudinal 2D area, longitudinal eccentricity, transverse long axis length, transverse short axis length, transverse 2D area, and transverse eccentricity.

Machine Learning Based Diagnosis

A random forest test was performed in R to determine the predicative powers of each of the nine variables with mean decrease in Gini coefficients. In the final step, multiple tests were performed using different machine learning algorithms on the 127-patient data with all nine variables. These include SVM (linear, RBF, polynomial, and sigmoid), KNN, logistic regression, and naïve bayes. Data was split into a standard 80% training to 20% testing proportion. Accuracy, precision, and sensitivity were recorded for both training and testing data.

Results

Ultrasound Image Segmentation and Annotation by U-Net

The modified U-Net model developed produced automatically segmented gallbladders with high accuracy compared to corresponding ground truth images. A representative original ultrasound image, the auto segmented image by modified U-Net, and the manually segmented ground truth image are shown in Figure 1. The auto segmented image resembled the manually segmented ground truth image with high similarity (Figure 1). Mean IOU scores, which represents how much overlap occurs between the auto segmented and ground truth image, ranged between 0.8803-0.9571 with all but one score ranging between 0.9435-0.9571 with that score specifically associated with the Nadam optimizer (Table 1).

Table 1: Summary of the Performance of U-Net Auto Segmentation.

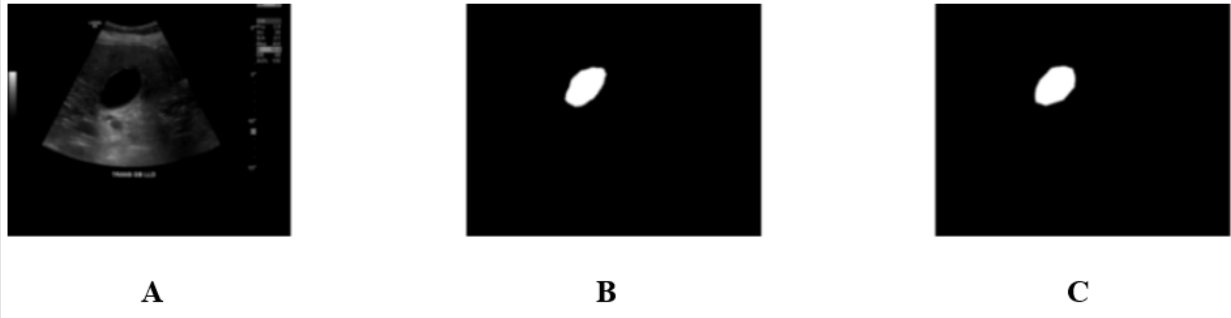| Optimizer	Learning Rate	Epoch	Dice Score	Mean IOU
Adam	0.001	362	0.9464	0.9462
	0.0003	501	0.9482	0.9479
	0.0005	568	0.9437	0.9435
RMSprep	0.001	528	0.9568	0.9562
	0.0003	590	0.9495	0.9492
	0.0005	597	0.9577	0.9571
Nadam	0.0005	150	0.8726	0.8803

Figure 1: Image A depicts original ultrasound scan. Image B represents modified U-Net auto segmented image. Image C is ground truth image.

Extraction of Gallbladder Parameters

The gallbladder features from longitudinal scan (Long Axis, Short Axis, 2D Area, Eccentricity) and transverse scan (Long Axis, Short Axis, 2D Area, Eccentricity), and 3D Volume were extracted by a Python algorithm. Results of relationships between different variables are displayed in a pair plot below (Figure 2).
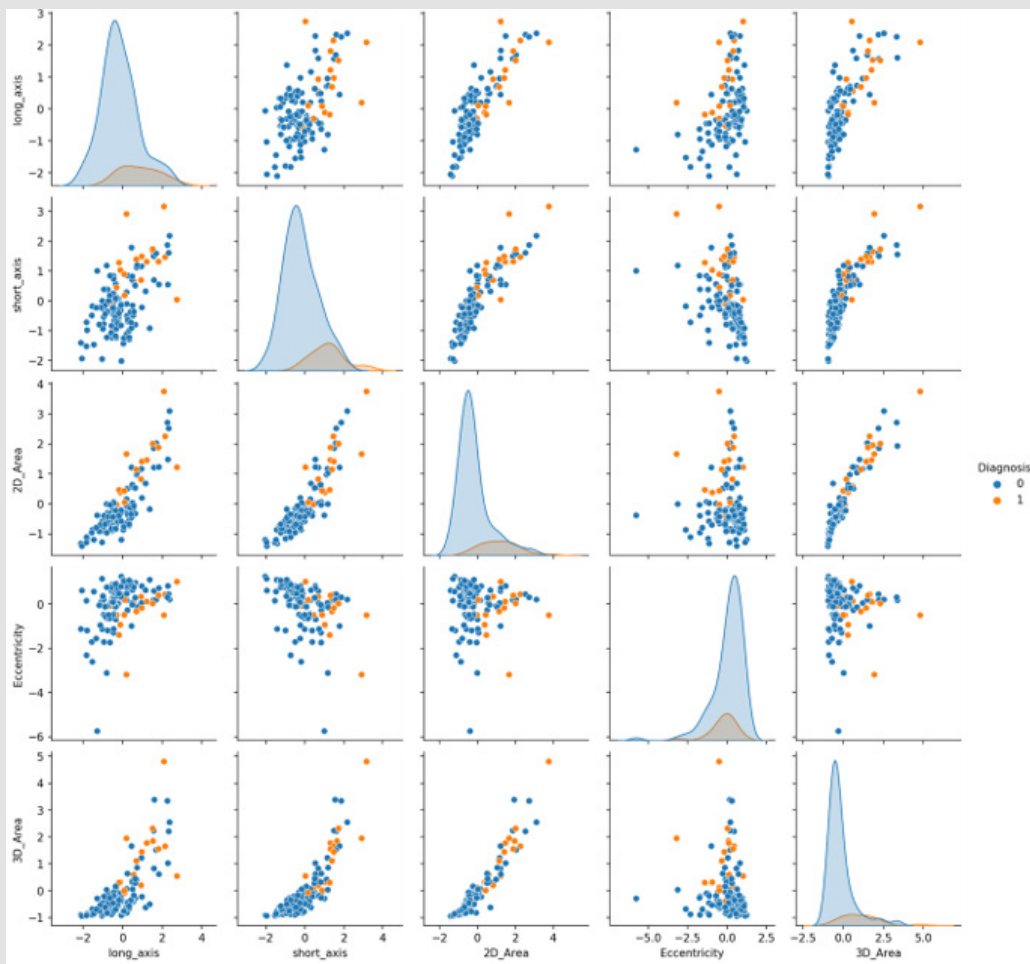

Figure 2: Pair plots between every variable. The two diagnoses are interspersed. Trends of more positive diagnoses moving from positive to negative on the x axis.
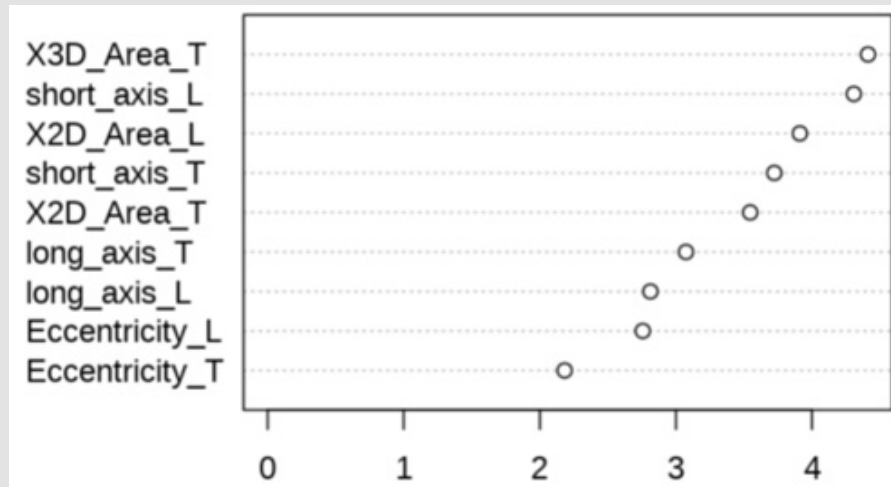
Figure 3: Graphical interpretation of mean Gini coefficient where predictive power is ranked increasing order from lower to upper.

Machine Learning Based Diagnosis

Mean Gini coefficient was calculated to determine the predictive power of each variable. 3D volume and the short axis of the longitudinal scan were found to have the highest power (4.411284 and 4.307068) (Table 2 & Figure 3). Out of the different machine learning algorithms tested, KNN and SVM with an RBF kernel performed the best. Training data produced accuracy of 0.907 and 0.843 for KNN and SVM with RBF respectively (Table 3). Meanwhile testing data produced very similar results with accuracy of 0.841 for both tests. Accuracy of around 84.1%-90.7% is higher than that of the pooled study on standard ultrasound diagnosis which specifically recorded sensitivity and specificity of 81% and 83% respectively (Table 4).

Table 2: Mean decrease in Gini coefficients for every variable.

Features	Value
Long Axis (longitudinal)	2.812422
Short Axis (longitudinal)	4.307068
2D Area (longitudinal)	3.910424
Eccentricity (longitudinal)	2.756172
Long Axis (transverse)	3.073955
Short Axis (transverse)	3.722548
2D Area (transverse)	3.545584
Eccentricity (transverse)	2.182415
3D Volume	4.411284

Table 3: Results for various machine learning methods in diagnosis of acute cholecystitis for training data.

	Accuracy	Precision	Sensitivity
SVM with Linear kernel	0.7965	0.7975	0.7965
SVM with RBF kernel	0.843	0.8538	0.843
SVM with polynomial kernel	0.7849	0.7995	0.7849
SVM with Sigmoid kernel	0.7209	0.721	0.7209
Logistic Regression	0.7965	0.7975	0.7965
KNN	0.907	0.9151	0.907
Naïve Bayes	0.8023	0.803	0.8023

Table 4: Results for various machine learning methods in diagnosis of acute cholecystitis for testing data.

	Accuracy	Precision	Sensitivity
SVM with Linear kernel	0.7954	0.7961	0.7955
SVM with RBF kernel	0.8409	0.8416	0.8409
SVM with polynomial kernel	0.75	0.7547	0.75
SVM with Sigmoid kernel	0.7273	0.7292	0.7273
Logistic Regression	0.7955	0.7961	0.7955
KNN	0.8409	0.8474	0.8409
Naïve Bayes	0.7727	0.775	0.7727

Discussion

The modified U-Net model developed produced automatically segmented gallbladders with high accuracy compared to corresponding ground truth images. Mean IOU scores ranged between 0.8803-0.9571 with all but one score ranging between 0.9435-0.9571 with that score specifically associated with the Nadam optimizer. Nadam also had a low epoch value of 150 which may have affected Mean IOU and there may be a case to exclude this optimizer for future experiments. In general, the segmentations of the U-Net were able to overlap in area with the majority of the segmented ground truth images. Strong and accurate performance of the U-Net segmentation is critical to the overall effectiveness of the machine learning method to predict acute cholecystitis. Mean Gini coefficient was calculated to determine the predictive power of each variable. 3D volume and the short axis of the longitudinal scan were found to have the highest power (4.411284 and 4.307068). These two variables were planned to be used to develop our models to predict acute cholecystitis, as it was believed that excessive variables would cause overfitting. However, I decided to test the model first with all nine variables and found that accuracy, sensitivity, and precision between training and test datasets were very similar, with 0%-6.7% decrease in accuracy of the model between the training and test data.

Out of the different machine learning models tested, KNN and SVM with an RBF kernel performed the best. Training data produced accuracy of 0.907 and 0.843 for KNN and SVM with RBF respectively. Meanwhile testing data produced very similar results with accuracy of 0.841 for both tests. Accuracy of around 84.1%-90.7% is higher than that of the pooled study on standard ultrasound diagnosis which specifically recorded sensitivity and specificity of 81% and 83% respectively. The current dataset only included 127 patients which was limited by the number of manual annotations. As such, once split into multiple different cohorts of testing, training, and validation, there were very few cases with positive diagnosis of acute cholecystitis. For future work, the segmentation will be fully automated by U-Net to speed up the data extraction. Subsequently, the dataset could be expanded to include more patients and generate results that reflect a larger population.

Conclusions and Future Work

This research has effectively developed methods to both segment gallbladder ultrasound images and diagnose acute cholecystitis with multiple machine learning methods. This work incorporated novel 3D volume measurements of the gallbladder with comparable accuracy, precision, and sensitivity to similar 2D work, as well as statistics gathered from traditional diagnosis through manual ultrasound examination. More robust models may be achievable using an expanded dataset with more patients. Furthermore, given the success of the KNN and SVM with an RBF kernel, further research may include constructing more complex versions of these specific machine learning models to raise accuracy. Other variables besides gallbladder size may be considered such as thickening of the gallbladder wall and the presence of pericholecystic fluid. With these advancements, using machine learning methods to diagnosis of acute cholecystitis can potentially be applied to clinical settings.

References

1. Gallaher JR, Charles A (2022) Acute Cholecystitis: A Review. JAMA 327(10): 965-975.

2. Kiewiet JJ, Leeuwenburgh MM, Bipat S, Bossuyt PM, Stoker J, et al. (2012) A systematic review and meta-analysis of diagnostic performance of imaging in acute cholecystitis. Radiology 264(3): 708-720.

3. Shaish H, Ma HY, Ahmed FS (2021) The utility of an under-distended gallbladder on ultrasound in ruling out acute cholecystitis. Abdominal radiology (New York) 46(6): 2498-2504.

4. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts H (2018) Artificial intelligence in radiology. Nature reviews. Cancer 18(8): 500-510.

5. Pianykh OS, Langs G, Dewey M, Enzmann DR, Herold CJ, et al. (2020) Continuous Learning AI in Radiology: Implementation Principles and Early Applications. Radiology 297(1): 6-14.

6. Chang CH, Lin CH, Lane HY (2021) Machine learning and Novel Biomarkers for the Diagnosis of Alzheimer's Disease. International journal of molecular sciences 22(5): 2761.

7. Yoo JW, Koo KC, Chung BH, Baek SY, Lee SJ, et al. (2022) Deep learning diagnostics for bladder tumor identification and grade prediction using RGB method. Scientific reports 12(1): 17699.

8. Ronneberger O, Fischer P, Brox T (2015) U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention Springer, Cham, pp. 234-241.

BIOMEDICAL RESEARCHES

ISSN: 2574-1241

Assets of Publishing with us

- Global archiving of articles
- Immediate, unrestricted online access
- Rigorous Peer Review Process
- Authors Retain Copyrights
- Unique DOI for all articles

https://biomedres.us/