

Accuracy in Classification of Liver Lesion Ultrasound Using Artificial Intelligence



A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in Medical Physics

Department of Radiology

FACULTY OF MEDICINE

Chulalongkorn University

Academic Year 2022

Copyright of Chulalongkorn University

ความแม่นยำของการจำแนกรอยโรคตับในภาพอัลตราซาวด์ด้วยการใช้ปัญญาประดิษฐ์



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต

สาขาวิชาฟิสิกส์การแพทย์ ภาควิชารังสีวิทยา

คณะแพทยศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2565

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

สุทธิรักษ์ ตั้งเรื่องเกียรติ : ความแม่นยำของการจำแนกรอยโรคตับในภาพอัลตราซาวด์ด้วยการใช้ปัญญาประดิษฐ์. (Accuracy in Classification of Liver Lesion Ultrasound Using Artificial Intelligence) อ.ที่ปรึกษาหลัก : รศ. ดร.อัญชลี ฤกษ์จินดา, อ.ที่ปรึกษาร่วม : รศ.ดร.สุพัฒนา เอื้อทวีเกียรติ

การตรวจอัลตราซาวด์แบบ B-mode เป็นมาตรฐานในการตรวจคัดกรองมะเร็งเซลล์ตับ (hepatocellular carcinoma, HCC) เมื่อรอยโรคตับมีลักษณะเด่นชัดแตกต่างกัน ตัวตรวจจับแบบดั้งเดิมก็สามารถนำมาใช้ได้ หากรอยโรคตับมีลักษณะที่คล้ายกันกับรอยโรคอื่นในตับ จำเป็นต้องใช้ฐานข้อมูลภาพที่มากขึ้น ซึ่งบางครั้งอาจมีไม่เพียงพอ ดังนั้นวิธีการแบบ 2 ขั้นตอนจึงถูกนำมาใช้ โดยในขั้นตอนแรก จะทำการตรวจจับโรคตับโดยไม่ต้องแยกประเภทให้ชัดเจน แล้วทำการจัดประเภทโรคตับที่ตรวจจับได้ในขั้นตอนที่ 2 การศึกษานี้ใช้ Convolutional Neural Network, CNN ในการจำแนกลักษณะของรอยโรคตับ

วัตถุประสงค์ของวิทยานิพนธ์นี้ เพื่อออกแบบโมเดลด้วยปัญญาประดิษฐ์ และสำรวจความถูกต้องแม่นยำของโมเดล โดยพิจารณาจากความมีเอกลักษณ์ของรอยโรคด้วยวิธี CNN ในการออกแบบ 2 ขั้นตอน

แม้ว่าลักษณะปรากฏของถุงน้ำและหลอดเลือดตับในภาพอัลตราซาวด์ จะเป็นแบบไม่มีคลื่นเสียงสะท้อน มีลักษณะเป็นวงกลมสีดำเหมือนกัน แต่ถุงน้ำในตับจะมีลักษณะพิเศษคือมีความเข้มเสียงมากขึ้นได้ต่อถุงน้ำนั้น การรวมลักษณะพิเศษในการตรวจจับ จะช่วยให้สามารถใช้ตัวตรวจจับแบบดั้งเดิมได้ การศึกษานี้ได้ใช้ Region-based convolutional neural networks, R-CNN พร้อมกับ Residual Network-50, ResNet-50 เป็นโมเดลในการตรวจจับถุงน้ำในตับจำนวน 615 ภาพ ระบบได้ประเมินผลโดยใช้วิธีไขว้ทบ 5 ส่วน ผลการศึกษาแสดงว่าการเพิ่มลักษณะพิเศษ จะเพิ่มความแม่นยำในการตรวจจับ ทั้งยังช่วยลดค่าผลบวกหลงและผลลบหลงได้

เนื้องอกหลอดเลือดในตับ (Hemangioma, HEM) และมะเร็งเซลล์ตับ มีลักษณะคล้ายกันในภาพอัลตราซาวด์ ไม่มีลักษณะพิเศษแบบเฉพาะเจาะจงที่จะช่วยแยกแยะระหว่างสองรอยโรคนี้ได้ ดังนั้นจึงประยุกต์วิธีแบบสองขั้นตอนโดยขั้นตอนแรกจะสอนตัวตรวจจับให้สามารถจับรอยโรคเนื้องอกที่มีลักษณะคล้ายคลึงกับ HEM และ HCC ขั้นตอนที่สอง ใช้ตัวจำแนกเพื่อแยกเป็น 3 กลุ่ม คือ typical HEM, atypical HEM และ HCC ซึ่ง atypical HEM และ HCC จะมีลักษณะภาพคล้ายกัน ดังนั้น การใช้เพียงภาพอัลตราซาวด์แบบ B-mode ไม่สามารถแยกแยะสองรอยโรคได้ จึงต้องเพิ่มการตรวจวินิจฉัยด้วยเครื่องเอกซเรย์คอมพิวเตอร์ และ / หรือ เครื่องสร้างภาพด้วยสนามแม่เหล็กไฟฟ้า ซึ่งไม่จำเป็นต้องแยกแยะสองรอยโรคนี้อย่างแม่นยำในขณะคัดกรอง เมื่อเปรียบเทียบผลการศึกษาการรวมกลุ่มเนื้องอก 2 ชนิดจะเพิ่มความไวในการตรวจจับ HCC จาก 0.64 เป็น 0.68 และจะเพิ่มความไวขึ้นอีกถ้าใช้วิธีสองขั้นตอน โดยเพิ่มขึ้นจาก 0.68 เป็น 0.72 อัตราความไวนี้เทียบเคียงได้กับโมเดลที่ถูกฝึกโดยใช้ฐานข้อมูลขนาดใหญ่กว่าที่มีเขตการยอมรับ (IoU) ว่าโมเดลทำนายถูกต้องเพียง 0.2 ขณะที่วิธีสองขั้นตอนตั้งไว้ที่ 0.5

สาขาวิชา ฟิสิกส์การแพทย์
ปีการศึกษา 2565

ลายมือชื่อนิสิต
ลายมือชื่อ อ.ที่ปรึกษาหลัก
ลายมือชื่อ อ.ที่ปรึกษาร่วม

6174756830 : MAJOR MEDICAL PHYSICS

KEYWORD: hepatocellular carcinoma, screening model, two-stage method, detector, classifier

Sutthirak Tangruangkiat : Accuracy in Classification of Liver Lesion Ultrasound Using Artificial Intelligence.

Advisor: Assoc. Prof. ANCHALI KRISANACHINDA, Ph.D. Co-advisor: Assoc. Prof. SUPATANA AUETHAVEKIAT, Ph.D.

B-mode ultrasound imaging is the standard method for hepatocellular carcinoma (HCC) screening. If a lesion has a unique appearance, a conventional detector can be applied. However, if the lesion shares its appearance with other lesion, a large training dataset is required, which may be unavailable. Therefore, a two-stage method is proposed. In the first stage, lesions are detected but not differentiated into any particular class. The detected lesions are then classified in the second stage using a conventional convolutional neural network (CNN).

The aims of the dissertation are to design an artificial intelligent system model and to investigate the most accurate deep learning structure for hepatic lesion classification using two-stage model.

Even though the cysts and hepatic vessels are both anechoic pattern and present black oval in ultrasound images, cysts have unique artifacts that present posterior acoustic enhancement. By including the artifacts in the detection, conventional detectors can be applied. In the study, Region-based convolutional neural networks, R-CNN, with Residual Network-50, ResNet-50, as the backbone was applied for the detection of 615 hepatic cysts. The system was evaluated by five-fold cross validation. The result indicated that the addition of artifacts led to better detection in term of accuracy, reduction of false positives and false negatives.

Hemangioma (HEM) and HCC share lots of sonographic appearance. There is no unique artifact to differentiate these two lesions. A two-stage method is applied. In the first stage, the detector is trained to identify HEM and HCC like lesions. In the second stage, the classifier is applied to differentiate lesions into typical HEM, atypical HEM and HCC. Since atypical HEM and HCC display the same sonographic appearance, both lesions cannot be differentiated solely by B-mode ultrasound images. They require further CT or MR investigation, so it is unnecessary to accurately differentiate these two lesions during screening. The study showed that grouping HCC and atypical HEM into the same one led to the increase of HCC recall of the detector from 0.64 to 0.68. The application of the two-stage method in place of detector only method improved the recall from 0.68 to 0.72. The recall rate was comparable to the detector only method that was trained by a much larger database and used more relax criterion (0.5 and 0.2 intersection over union (IoU) for the correct detection in the proposed two-stage and the detector only methods, respectively).

Field of Study: Medical Physics

Academic Year: 2022

Student's Signature

Advisor's Signature

Co-advisor's Signature

ACKNOWLEDGEMENTS

I am deeply grateful to the following individuals who have contributed significantly to the completion of my Ph.D. Dissertation:

First and foremost, I would like to express my heartfelt gratitude to my advisor, Associate Professor Anchali Krisanachinda, Ph.D. Division of Nuclear Medicine, Department of Radiology, Faculty of Medicine, Chulalongkorn University. Her unwavering support, care, and guidance throughout the research work have been invaluable to me. I am grateful for her tireless efforts in polishing my work and guiding me towards the right direction. I am fortunate to have her as my mentor and advisor.

I would also like to thank my co-advisor, Associate Professor Dr. Supatana Auethavekiat, Department of Electrical Engineering, Faculty of Engineering, Chulalongkorn University. Her expertise in deep learning and patient guidance has been instrumental in my research journey. Her helpful advice and suggestions have been invaluable to me, and I am grateful for her support.

I would like to extend my gratitude to Professor Franco Milano, the Chairman of my thesis defense, for his help, kind suggestions, and comments on this research. His insights and feedback have been crucial to the success of my research work.

I would like to express my sincere appreciation to Assistant Professor Yothin Rakvongthai and Picha Shunhavanich, Ph.D., Department of Radiology, Faculty of Medicine, Chulalongkorn University, for their contribution as committee members. Their insightful comments and feedback have been crucial in shaping my research work.

Last but not least, I would like to thank Professor Dr. Kosuke Matsubara, Department of Quantum Medical Technology, Faculty of Health Sciences, Kanazawa University, Japan, the external examiner of the thesis, for his kind suggestions, feedback, and comments on this research. His expert opinion has been invaluable to me, and I am grateful for his help.

Once again, I would like to express my heartfelt appreciation to all those who have contributed to my research work. Without their guidance, support, and encouragement, this dissertation would not have been success.

Sutthirak Tangruangkiat

TABLE OF CONTENTS

	Page
.....	iii
ABSTRACT (THAI).....	iii
.....	iv
ABSTRACT (ENGLISH).....	iv
ACKNOWLEDGEMENTS.....	v
TABLE OF CONTENTS.....	vi
LIST OF TABLES.....	x
LIST OF FIGURES.....	xi
LIST OF ABBREVIATIONS.....	xiv
CHAPTER 1 INTRODUCTION.....	1
1.1 Background and Rationale.....	1
1.2 Hypothesis.....	2
1.3 Research Questions.....	2
1.4 Research Objectives.....	3
1.4.1 Primary Objective.....	3
1.4.2 Secondary Objective.....	3
1.5 Flow Chart of Thesis.....	3
1.6 Ultimate Goal.....	2
CHAPTER 2 THEORY AND RELATED LITERATURE.....	3
2.1 Theory.....	3
2.1.1 Liver Anatomy and Histology.....	3

2.1.2 Liver Tumors	4
2.1.3 Liver Cancers.....	5
2.1.4 Imaging Guidelines of HCC	6
2.1.5 Diagnosis of Liver Diseases.....	7
2.1.6 Sonographic Appearance of Hepatic Tumor	10
2.1.7 Machine Intelligence	11
2.1.8 Convolutional Neural Network (CNN).....	14
2.1.9 ResNet	15
2.1.10 Object Detection Techniques.....	16
2.1.11 Object Detectors.....	21
2.1.12 Statistical Analysis and Model Evaluation	22
2.2 Review of related literature	25
2.3 Using YOLOv4 for Classifying HCC and HEM in Ultrasound Images with Limited Dataset	28
2.4 Application of Modified Object Detectors on ResNet-50 Architectures for Liver Lesion Detection and Classification in Ultrasound Images	29
2.5 Normalization Techniques for Liver Ultrasound Images	35
2.6 Intensity Normalization	36
CHAPTER 3 MATERIALS	37
3.1 Computer System.....	37
3.2 MATLAB Software	37
3.3 RadiAnt DICOM Viewer Software.....	38
3.4 Dataset.....	39
CHAPTER 4 METHODS.....	40

4.1 Cyst Detection.....	40
4.1.1 Data Preparation	40
4.1.2 Observation Method	41
4.1.3 Selection of Detector Models	44
4.2 Hepatic Tumor Detection and Classification	45
4.2.1 Data Preparation and Marker Analysis in HEM and HCC	45
4.2.2 Two-Stage Method for Hepatic Tumor Detection and Classification	46
4.2.3 Training Phase for Detector and Classifier	47
4.2.4 Selection of Detector and Classifier Models.....	48
4.2.5 Observation Methods.....	49
CHAPTER 5 RESULTS	51
5.1 Cyst Detection.....	51
5.2 Hepatic Tumor Detection and Classification	52
5.2.1 Performance Analysis of Detectors	52
5.2.2 Performance Analysis of Classifier	56
5.2.3 Performance Analysis of Two-Stage Model.....	59
CHAPTER 6 DISCUSSION	61
6.1 Cyst Detection.....	61
6.2 Hepatic Tumor Detection and Classification	63
6.2.1 Analysis and Interpretation of Results in The First Stage Model or The Tumor Detector.....	63
6.2.2 Analysis and Interpretation of Results in The Second Stage Model or The Tumor Classifier.....	64
6.2.3 Analysis and Interpretation of Results of the Two-Stage Model	65

6.3 Overall Discussion	66
CHAPTER 7 CONCLUSIONS.....	68
7.1 Implications of Cyst Detection Study.....	68
7.2 Implications in Detecting and Classifying Hepatic Tumor	68
7.3 Overall Conclusions.....	69
7.4 Suggestions for Future Research	69
REFERENCES	72
VITA.....	78



LIST OF TABLES

Table 1 Architectures for ResNet-50 (Adapted from [24]).	16
Table 2 Summary of works using AI in medical imaging analysis.	26
Table 3 The distribution of available lesions in Hassan et al. [39] and Schmauch et.al. [4].	27
Table 4 Results of the different ground truth labelling for detections.	51
Table 5 Detection results of four YOLOv4 models.	53
Table 6 Recall Rate and Variation Range of Four YOLOv4 Models.	53
Table 7 The detection results of Model 4 as grouped by lesion type.	55
Table 8 Results of YOLOv4 Detector and ResNet-50 Classifier on 3 x 3 confusion matrix.	57
Table 9 Results of modified 3 x 3 confusion matrix as 2 x 2 confusion matrix.	57
Table 10 The results of HCC detection using both single-stage detector and two-stage methods.	60

LIST OF FIGURES

Figure 1 Flow Chart of Thesis Chapter 3 - 7.	4
Figure 2 Structural organization of the liver [12].....	4
Figure 3 The HCC surveillance guideline of the American Association for the Study of Liver Diseases (AASLD) in 2010 [19].	8
Figure 4 (A) An ultrasound image with a hepatic cyst (B) The arrows present the alignment of ultrasound beams and the dashed arrow point to the cyst (black oval) and its artefact in the dashed line area align with the ultrasound beam.....	9
Figure 5 Echogenic patterns of hepatocellular carcinoma (HCC) and hepatic hemangioma (HEM) lesions. The top and the bottom rows show the echogenic patterns of HCC and HEM lesions, respectively. (a) hypoechoic, (b) hyperechoic, (c) isoechoic, and (d) mixed echoic.....	11
Figure 6 AI encompasses the development of intelligent machines, ML enables computers to learn and make predictions from data, and DL leverages deep neural networks for complex pattern recognition tasks.	Error! Bookmark not defined.
Figure 7 ML works as a simple neural network (A) only a few hidden layers (B) DL works as multiple hidden layers.	13
Figure 8 The image shows IoU calculation.....	19
Figure 9 (a) represents a precision – recall curve (b) shows mAP could be calculated by summarization area A to area D.	20
Figure 10 (a) ResNet, a convolutional neural network and (b) feature pyramid net (FPN) is a key component of RetinaNet and plays an important role in the model's ability to detect objects at different scales and resolutions. [45].	28
Figure 11 Architecture R-CNN on ResNet-50 for cyst detection. The architecture consists of region proposals with feature extraction and object classification with bounding box regression.	31

Figure 12 Two-stage model for hepatic tumor detection and classification.....	34
Figure 13 Architecture YOLOv4 and ResNet-50 as backbone for hepatic detection and classification in the first stage.	34
Figure 14 Architecture Resnet-50 classifier for classify hepatic tumor into typical hemangioma (T_HEM), atypical hemangioma (A_HEM), and hepatocellular carcinoma (HCC) in the second stage.....	34
Figure 15 Top row: Liver ultrasound images - (a) Homogeneous liver, (b) Hepatic cyst, (c) Hepatocellular carcinoma (HCC), and (d) Hepatic hemangioma (HEM). Bottom row: Abdominal wall indicated by yellow box.	36
Figure 16 The liver image was cropped to be a square.	41
Figure 17 The normalization of the image according to the intensity median inside the abdominal tissues (shown in the red windows of the left figure). The right image is normalized left image to the intensity median of 70.	43
Figure 18 The samples from two training sets. (A) Displays a sample from the Cyst Ground Truth set, which contains an ROI with only one cyst. (B) A sample from the Cyst with Acoustic Enhancement Ground Truth set, which includes an ROI with both a cyst and a bright artifact located below the cyst.	43
Figure 19 Examples of the lack of markers and the marker for lesions/tissues other than HEM and HCC. HEMs and HCCs were shown inside the dashed circle. (a) HEM without marker, (b) HCC without marker, (c) marker for measuring hepatic cyst in the HEM image and (d) marker for measuring the vessel near the HCC.....	46
Figure 20 The proposed two-stage method and dataset.	47
Figure 21 Comparison of detection results using YOLOv2 and YOLOv4. The top and the bottom rows display the detection by YOLOv2 and YOLOv4, respectively. YOLOv4 produced higher prediction scores (the number in the yellow box).	49
Figure 22 Training YOLOv4 to detect HEM and HCC like lesions as one class resulted in an improved HCC detection, as depicted in the bottom row of the figures. The top	

row displays the results of the same region when the models were trained to detect HCC or HEM as separate classes, which failed to detect the lesions.....	55
Figure 23 Examples of multiple detection of the same lesion by YOLOv4.....	56
Figure 24 Average area under the ROC Curve as 92% for HEM and HCC-like Lesions.	57
Figure 25 Examples of errors from the classifier trained with Cyst Ground Truth set. Actual cysts were indicated by arrows. The solid box was the cyst detected by the classifier. The classifier detected a branch of right hepatic vein as a cyst in (A) and a part of subcapsular as a cyst in (B). The classifier failed to detect a cyst in an image with low contrast in (C) and a cyst with incomplete border in (D).....	62
Figure 26 Examples of errors from the classifier trained with Cyst with Acoustic Enhancement Ground Truth set. Actual cysts are indicated by arrows. The windows are the cysts detected by the classifier. The classifier detected a branch of hepatic vein as a cyst in (A) and a part of diaphragm as a cyst in (B). The classifier failed to detect a cyst in an image with low contrast in (C) and a cyst with incomplete border in (D).....	62
Figure 27 Small oval-shaped hyperechoic lesions resembling HCC were misclassified as typical HEM by the Classifier.....	65

LIST OF ABBREVIATIONS

AASLD	American Association for the Study of Liver Disease
ADN	Adaptive Deconvolutional Networks
AI	Artificial Intelligence
ANN	Artificial Neural Network
APASL	Asia-Pacific Association for the Study of the Liver
AUC	Area under the curve
CAD	Computer-Assisted Diagnosis
CCA	CholangioCarcinoma
CEUS	Contrast-enhance ultrasound
CI	Confidence interval
CNN	Convolutional Neural Network
CT	Computed Tomography
DBN	Deep Belief Network
DICOM	Digital imaging and communications in medicine
DL	Deep learning
EASL	European Association for the Study of the Liver
eCCA	Extrahepatic cholangiocarcinoma
Fig.	Figure
FLC	Fibrolamellar carcinoma
FNH	Focal Nodular Hyperplasia

FN, FNR	False negative, False negative rate
FP, FPR	False positive, False positive rate
GB	Gigabyte
HA	Hepatic artery.
HBV	Hepatitis B virus
HCC	Hepatocellular carcinoma
HCV	Hepatitis C virus
HEM	Hepatic hemangioma
HV	Hepatic vein
iCCA	Intrahepatic cholangiocarcinoma
ILSVRC	ImageNet large scale visual recognition challenge
IoU	Intersection of union
KLCSG	Korean liver cancer study group
mAP	Mean average precision
ML	Machine learning
MR, MRI	Magnetic resonance, Magnetic resonance imaging
MS COCO	Microsoft Common Objects in Context
NPV	Negative predictive value
PACs	Picture archiving and communication system
PPV	Positive predictive value
PV	Portal vein

R-CNN	Region-based convolutional neural networks
ReLU	Rectified linear unit
ROC	Receiver operating characteristic
ROI	Region of interest
SSAE	Stacked sparse auto-encoder
SSD	Single shot multibox detector
SVM	Support vector machine
TN, TNR	True negative, True negative rate
TP, TPR	True positive, True positive rate
US	Ultrasound
v.	Version
χ^2	Pearson's chi squared
XAI	Explainable Artificial Intelligence
YOLO	You only look once

CHAPTER 1

INTRODUCTION

1.1 Background and Rationale

Global cancer statistics in 2018 reported that liver cancer is the sixth most common cancer and the third most common cause of cancer mortality [1]. Approximately 90% of primary liver cancer is hepatocellular carcinoma (HCC) [2]. The detection of early-stage HCC leads to more curative treatment, a decrease in morbidity and the goal of HCC surveillance.

Patients with hepatitis B or C have a higher risk of HCC due to the development of liver fibrosis in these diseases. They should be enrolled in the HCC screening and surveillance program and routinely have their liver screened for lesions by ultrasound (US) imaging. If the suspected lesion is detected, they are further investigated by computed tomography (CT) or magnetic resonance (MR) imaging. Since CT and MR imaging have not been available in primary healthcare systems, the longtime waiting caused the rapid progress of the HCC.

The current standard protocol for liver screening is B-mode ultrasound imaging. However, lesions may not have a distinct appearance, which can make it difficult to differentiate hepatic lesions. To address this, the proposed approach suggests designing the architecture of artificial intelligent systems based on the uniqueness of the lesions. For lesions with a distinct appearance, a conventional detector can be applied. However, when lesions share an appearance with other lesions, a large training dataset is necessary, which may not always be available. To address this challenge, a two-stage method is proposed. In the first stage, lesions are detected without being classified into a particular group, and in the second stage, a conventional convolutional neural network is used to classify the detected lesions.

A hepatic cyst on an ultrasound image is one of the lesions that has a distinct appearance. However, interpreting 2-dimensional ultrasound images can be

challenging, particularly in differentiating cystic lesions and hepatic vessels, as both present as anechoic patterns [3]. Some authors used deep learning models to detect hepatic lesions and found that the models failed to distinguish hepatic vessels from hepatic cysts [4, 5].

Hepatic Hemangioma (HEM) and HCC are both hepatic tumors that can be difficult to differentiate from each other when both present similar features in ultrasonography [6, 7]. HEM is a common benign tumor mostly found in clinical practice [6, 8, 9]. However, in many patients, the lesion may appear atypical, and further CT/MR examination is necessary. Conventional ultrasonography has limitations, which lead to the need for more advanced diagnostic tools.

Artificial intelligence (AI), specifically deep learning, has been successful in pattern recognition, including computer-assisted diagnosis (CAD) [10, 11]. In recent years, there have been promising results in the application of deep learning algorithms to assist radiologists and sonographers in diagnosing hepatic lesions in ultrasonography. Therefore, the aim of this research is to develop and evaluate a deep learning model to aid in differentiating cysts and hepatic vessels, as well as HEM and HCC lesions.

1.2 Hypothesis

The implementation of AI in liver ultrasonography can enhance the accuracy of liver lesion detection and classification.

1.3 Research Questions

What is the design of the feature extraction algorithm for the accurate detection and classification of liver lesions in ultrasonography?

1.4 Research Objectives

1.4.1 Primary Objective

To design the feature extraction algorithm using AI for focal liver lesion detection and classification in liver ultrasound images.

1.4.2 Secondary Objective

To investigate the most accurate deep learning structure for hepatic lesion classification when used with the proposed two-stage model.

1.5 Flow Chart of Thesis

The thesis titled "Accuracy in Classification of Liver Lesion Ultrasound Using Artificial Intelligence" employs a flowchart to enhance the understanding of the interrelation between each chapter as shown in Figure 1. The study considers a small dataset and different lesion characteristics as key factors.

The flowchart initially focuses on the unique appearance of the lesion, specifically investigating hepatic cysts. In cases where the lesion exhibits a distinctive appearance, the "Conventional Detector 'R-CNN'" is employed for precise cyst detection and classification.

Conversely, if the lesion lacks a unique appearance, as seen in hepatic tumors such as HCC or HEM, the study utilizes the "Two-Stage Model" for detection and classification of these tumor types. The flowchart helps guide the study's progression and clarifies the decision-making process based on the characteristics of the lesions.

The flowchart of the thesis studies is designed to present the overall thesis and establish the relationship between each chapter. Here is a breakdown of the chapters:

Chapter 3: This chapter provides a description of the materials used in the two studies.

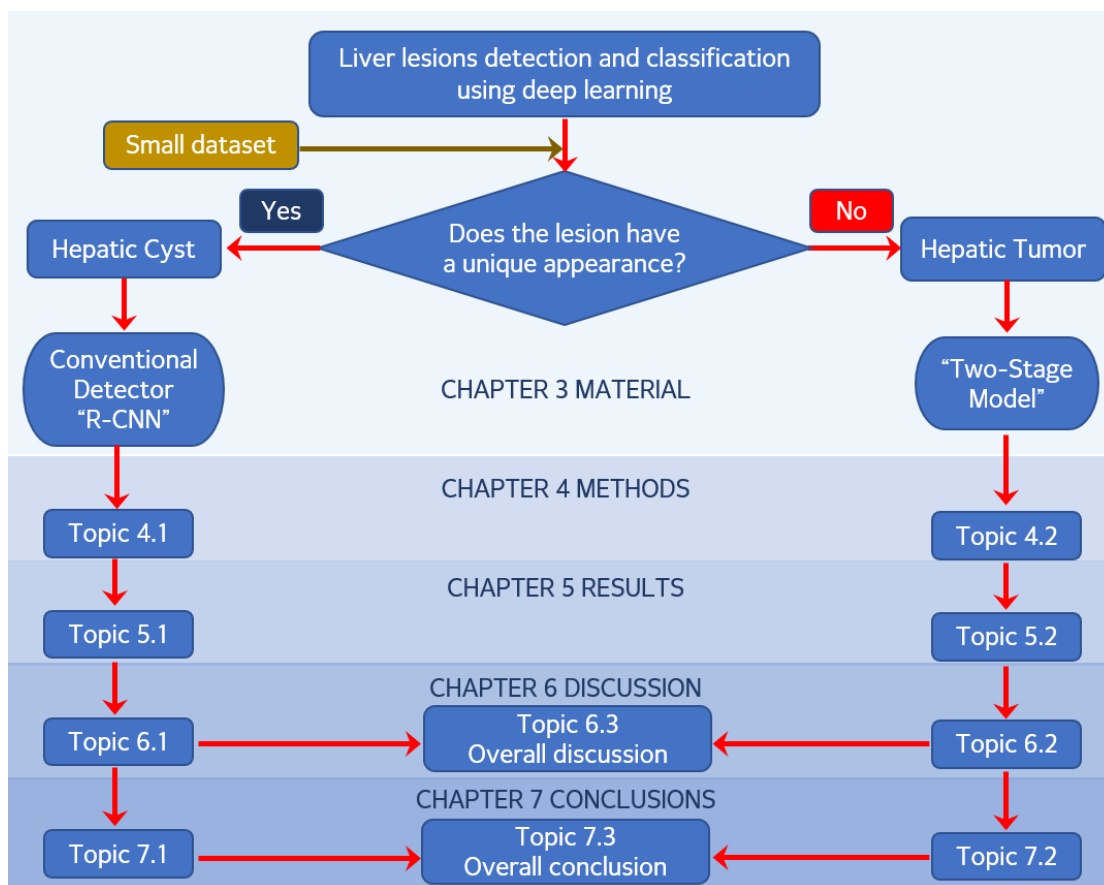


Figure 1 Flow Chart of Thesis Chapter 3 - 7.

Chapter 4: The methods of two studies are in detail. Topic 4.1 covers cyst detection, and topic 4.2 examines hepatic tumor detection and classification.

Chapter 5: This chapter presents the results obtained from the two studies. Topic 5.1 refers to the results of cyst detection, while topic 5.2 focuses on hepatic tumor detection and classification.

Chapter 6: The discussion refers to the findings of two studies, topic 6.1 refers to cyst detection, while topic 6.2 examines hepatic tumor detection and classification.

Chapter 7: The conclusions include two studies, topic 7.1 presents the conclusions related to cyst detection, and topic 7.2 concludes the hepatic tumor detection and classification.

This chapter enables a comprehensive understanding of the research, allowing readers to navigate through the different aspects of the study with clarity.

1.6 Ultimate Goal

The ultimate goal of this thesis is to improve the current practice of HCC screening by leveraging artificial intelligence (AI) in ultrasound imaging. By developing and evaluating a deep learning model specifically designed for liver ultrasound images, in order to enhance the accuracy and efficiency of HCC diagnosis and differentiation between hepatic cysts, hepatic vessels, HEM, and HCC lesions.

With higher performance achieved through the integration of AI in ultrasound, the purpose is to transform ultrasound image from being solely a modality for detecting abnormal findings in accordance with HCC guidelines to a powerful tool capable of effectively ruling out typical HEM cases. This would result in a reduction in the number of patients requiring further investigations by CT/MRI, thereby streamlining the diagnostic process. Moreover, by improving the efficiency of lesion detection and classification, the waiting time for patients with suspicious liver tumors can be significantly reduced.

Ultimately, by enhancing the accuracy and capabilities of ultrasound through AI, the goal is to optimize the HCC screening process, enabling timely and accurate diagnoses, reducing unnecessary investigations, and improving patient outcomes.

It is important to note that this thesis represents an initial step towards improving the current practice of HCC screening. While the integration of AI in ultrasound imaging shows promising potential, further research and development are required to fully realize its benefits in clinical settings. This thesis serves as a foundation for future advancements in AI-assisted diagnosis and classification of hepatic lesions in ultrasound, aiming to contribute to the ongoing efforts in enhancing HCC screening programs and ultimately improving patient health care.

CHAPTER 2

THEORY AND RELATED LITERATURE

2.1 Theory

2.1.1 Liver Anatomy and Histology

The liver is the largest solid organ in the human body, located in the upper right of the abdominal cavity, beneath the diaphragm. The liver is divided into right and left lobes, the left lobe is smaller than the right lobe. There are three blood vessels connected to the liver as follows.

- The hepatic artery (HA) brings oxygenated blood to the liver.
- The portal vein (PV) or hepatic portal vein brings nutrient-rich blood from the digestive system to the liver.
- The hepatic vein (HV) carries oxygen-poor blood out of the liver.

The hepatic artery and portal vein subdivide into small capillaries known as liver sinusoids, which end in a lobule. A liver sinusoid serves as the location for mixing the oxygen-rich blood from the hepatic artery and the nutrient-rich blood from the portal vein.

The hepatic lobule is the functional unit of a liver. Each lobule is made of millions of hepatocytes (hepatic cells) radiating from a central vein. The central vein joins the hepatic vein to carry blood out of the liver. A distinctive component of a lobule is the portal triad, which runs along each of the lobule's corners [12] as shown in Figure 2. The Portal triad, misleadingly named, consists of five structures: the hepatic artery, the portal vein, hepatic duct, lymphatic vessels and vagus nerve [13]. Between the hepatocyte plates are liver sinusoids, which are enlarged capillaries through which the blood from the hepatic portal vein and the hepatic artery enters via the portal triads, and then drains to the central vein. The main function of the liver is to filter the blood coming from the digestive system, before passing it to the rest of the body. Moreover, the liver also metabolizes drugs, detoxifies chemicals, and produces bile which is passed to the gallbladder. Therefore, the liver is a vital organ and supports other organs and is also prone to many diseases.

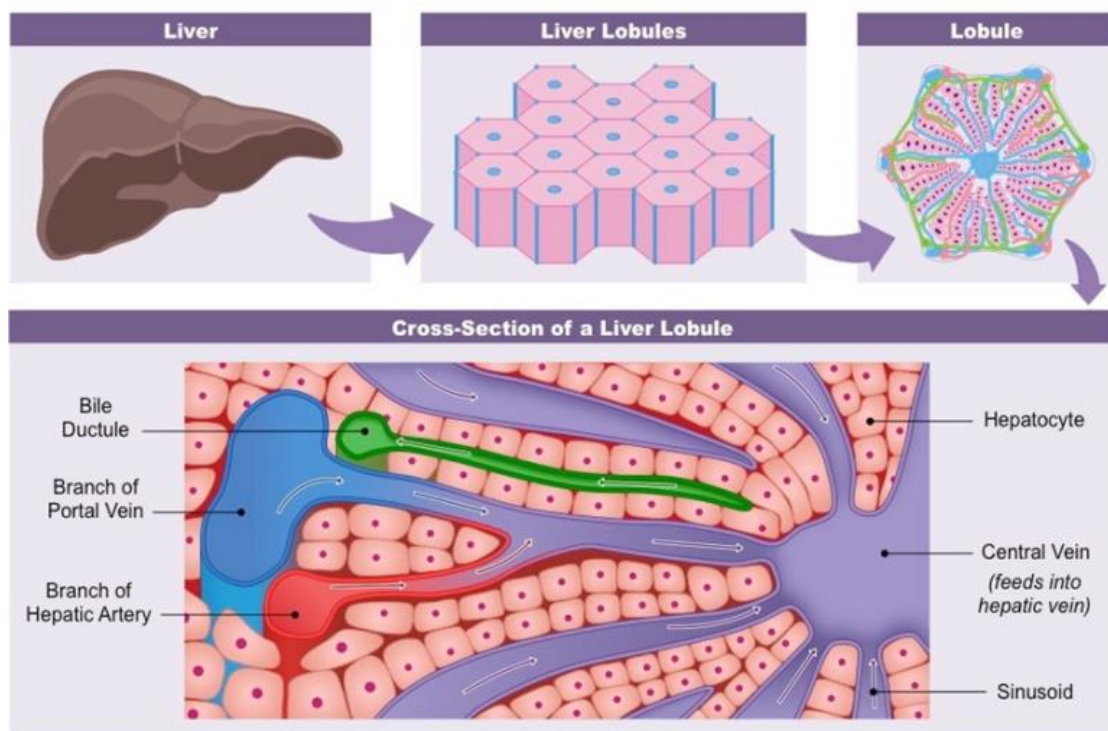


Figure 2 Structural organization of the liver [12].

Liver disease can be hereditary or acquired, such as a virus or bacterial infection, alcohol abuse, obesity, etc. Chronic liver diseases may result in cirrhosis, and some may cause liver cancer.

2.1.2 Liver Tumors

The liver is made up of various cell types so several types of tumors can be formed. Some are benign and some are cancerous. These tumors, having different causes, are treated differently. The most common types of benign liver lesions are as follows.

A. Hepatic Hemangioma — is also called a cavernous hemangioma or ‘HEM’. The tumor is the most common benign tumor of the liver [8, 9], which is made up of abnormal blood vessels. Usually, the tumors do not cause symptoms and need no treatment. In some cases, larger HEMs can cause pain or discomfort and need to be removed. HEMs may be presented from birth as well as developed at any point during a person's life.

B. Focal nodular hyperplasia (FNH) — is the second most common benign tumor of the liver. It is more common in females (female: male ratio 8:1) [14]. FNH is a growth of several cell types, hepatocytes, bile ductules (small ducts) and Kupffer cells. These tumors do not bleed or become cancerous. The tumors can be removed if they are very large, cause symptoms, or are located in unfavorable areas.

C. Hepatic Adenoma — Liver adenomas or hepatic adenomas are uncommon benign growth of hepatocytes [15]. Many of these tumors do not cause symptoms, but they can rupture and bleed, as well as become cancerous. Therefore, this tumor is usually removed when it is found.

D. Hepatic Cyst — is a cavity in the liver that contains fluid. On sonographic appearance, simple cyst present anechoic, well defined and posterior acoustic enhancement artifact [16]. Single or multiple liver cysts are common, especially with advancing age. While liver cysts are usually benign, a cyst may become enlarged or infected, requiring treatment.

2.1.3 Liver Cancers

Primary malignancy of the liver is a common cancer in many parts of Eastern countries. It affects more men than women and is usually found in people with chronic liver disease. There are five main types of primary liver cancer as follows.

A. Hepatocellular carcinoma (HCC) is the most common type of primary liver cancer. HCC develops from hepatocytes, the major liver cell (70-85%) [17]. It is usually found in people with chronic liver disease, such as a patient with hepatitis B virus (HBV) or C virus (HCV) infection or alcoholic cirrhosis.

B. Fibrolamellar carcinoma (FLC) or Fibrolamellar hepatocellular carcinoma is a rare sub type of HCC. It tends to develop in teens and adults under 40 years old. It is not usually linked with cirrhosis or infection with hepatitis B or C.

C. Cholangiocarcinoma (CCA) is bile duct cancer. One unique liver function is to produce bile. Bile is a yellowish-brown fluid produced continuously by the liver. The bile travels from liver cells through the gallbladder through small tubes scattered in the liver and pass-through small intestine. The bile duct always comes along with the hepatic artery and portal vein in the liver known as the portal triad.

There are 2 types of CCA: intrahepatic cholangiocarcinoma (ICC or iCCA) and extrahepatic cholangiocarcinoma (ECC or eCCA). ICC and ECC are the type of cancer that starts in the duct section inside and outside the liver, respectively. Because most CCA is ICC [18]. CCA is usually implied ICC.

D. Angiosarcoma or hemangiosarcoma is an extremely rare type of primary liver cancer and can be found in people aged over 70 years. This cancer starts in the blood vessels of the liver and causes the patient to rapidly bleed to death when ruptures.

E. Hepatoblastoma is a very rare type of primary liver cancer that usually affects young children. It is mostly diagnosed in infants and children less than 3 years of age.

Globally, CCA is much less common than HCC. However, there is a high incidence of both HCC and CCA in Thailand. In this thesis, only HCC is focused.

2.1.4 Imaging Guidelines of HCC

HCC is one of the global major health problems. Imaging guidelines for HCC surveillance are designed and revised by many groups or associations such as

- European Association for the Study of the Liver (EASL),
- Korean Liver Cancer Study Group (KLCSG),
- American Association for the Study of Liver Disease (AASLD),
- Asia-Pacific Association for the Study of the Liver (APASL).

Guidelines are not rigid protocols and depend on local clinical judgment. Figure 3 shows the 2010 HCC surveillance guideline of AASLD. Ultrasound has been used with or without serum alpha-fetoprotein (AFP) test every 6 months as a screening for HCC [19]. If the ultrasound reveals any liver nodules, the radiologist will proceed as below.

- If the nodule is smaller than 1 cm in diameter, repeat ultrasound every 3-6 months.
- If the nodule is larger than 1 cm in diameter, do a further test with either 4-phase CT or MRI with contrast. If the lesion appearance shows hypervascularity with washout in the portal/venous phase, which is typical for HCC, the lesion is HCC.

2.1.5 Diagnosis of Liver Diseases

Patients with a high risk of HCC should be enrolled in the HCC screening and surveillance program. Recently, non-invasive imaging approaches have been the first-line tools for screening. The non-invasive imaging for HCC assessment is as follows:

A. Ultrasound (US)

The US is a non-invasive, economical, and non-radiative imaging technique. It can be used for the diagnosis of soft tissue, which is not possible in normal X-rays. It is suitable for patients of any age and is used as the initial screening tool for liver lesions. The US is used for the detection of early-stage HCC nodules in HCC surveillance.

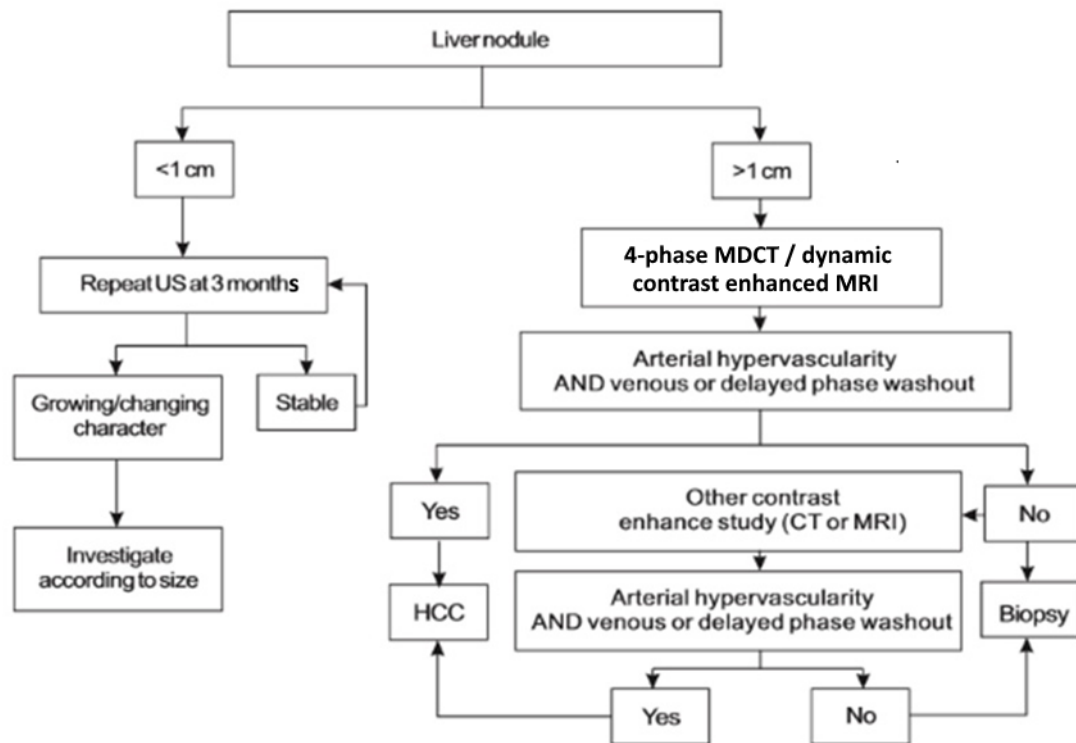


Figure 3 The HCC surveillance guideline of the American Association for the Study of Liver Diseases (AASLD) in 2010 [19].

The sound wave from the US probe has an array pattern. The sound can pass through media or organ tissue, the 4 interactions between the sound and media are reflection, refraction, scattering and attenuation [20]. Reflection is the best effect to perform an US image. Conversely, the 3-rest effect decreases sound wave and also decreases the image quality. Moreover, the sound waves pass through the low attenuating media, the rest of the sound waves are stronger than the sound waves at the same level that pass through another medium. The posterior acoustic enhancement happens when a sound wave passes through a weakly attenuating media such as water or cyst or fluid-filled structures. For example, the US wave passing through a hepatic cyst will suddenly occur the increased through transmission after the cyst [21]. The posterior acoustic enhancement or cyst artefact would help the examiner to differentiate fluid-filled structures from hypoechoic masses. Figure 4 shows the cyst artefact as the bright area below the cyst.

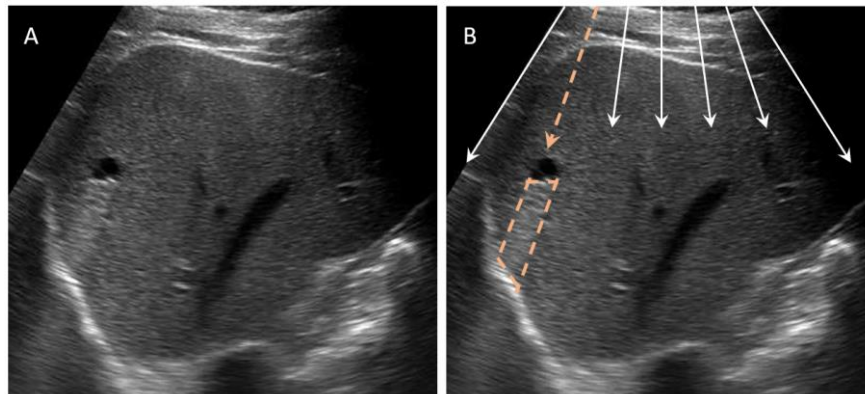


Figure 4 (A) An ultrasound image with a hepatic cyst (B) The arrows present the alignment of ultrasound beams and the dashed arrow point to the cyst (black oval) and its artefact in the dashed line area align with the ultrasound beam.

B. Computed Tomography (CT)

CT is widely used in cancer diagnosis. A patient usually lies on his/her back in the scanning device and has the X-ray beams passing through his/her body to the detectors to produce projection images at different angles. Cross-sectional images of a body are reconstructed from the projection images. CT produces a high-resolution image of soft tissue, muscles, bone, tumors, or other irregularities. It is a fast-imaging modality but has high radiation exposure which increases the cancer risk.

4-phase CT is used to identify HCC. The characteristics of HCC are as follows:

- In the non-contrast phase, the lesion appears as hypodense lesions.
- In the arterial phase (the phase of imaging 10 - 20 seconds after contrast bolus injection), the lesion becomes enhanced with contrast media.
- In the venous phase (the phase of imaging 20 - 35 seconds after the arterial phase), the lesion is rapidly washed out and returns to a hypodense state.

- In the delay or equilibrium phase (the phase of imaging 5 minutes after bolus injection), most contrast media is washed out and most tissue returns to its native structures.

C. Magnetic Resonance Imaging (MRI)

MRI is non-radiative but expensive and requires a long scanning time. It has high sensitivity and specificity for liver cancer detection; however, the liver examination requires multiple MRI sequences and may take more than an hour to finish. For example, the hepatobiliary tract is investigated in the delayed phase at 10-30 minutes with and without fat saturation suppression, and later delayed phase within 1 hour or more in some institutions. During scanning, a patient lies inside an enclosed space filled with loud noise from the magnetic coil, so he/she may become claustrophobic.

Dynamic MRI with gadolinium contrast is used to identify HCC. The characteristic of HCC is as follows.

- In the T1-weighted image, the lesion appears as hypo signal intensity lesions.
- In the T2-weighted image, the lesion has hyper signal intensity and is filled with contrast media; however, it is not too bright.
- In the dynamic phase, the lesion becomes a hyper vascular pattern as enhancing lesions filled with contrast media.
- In the hepatobiliary phase, the uptake of contrast media is no longer visible in the hepatobiliary structure.

2.1.6 Sonographic Appearance of Hepatic Tumor

There are many types of liver lesions, but a common type of liver tumor is hepatic HEM. Common type of liver cancer is HCC. The study focused on HEM and HCC as liver tumors. There are various sonographic appearances of liver tumors as shown in Figure 5.

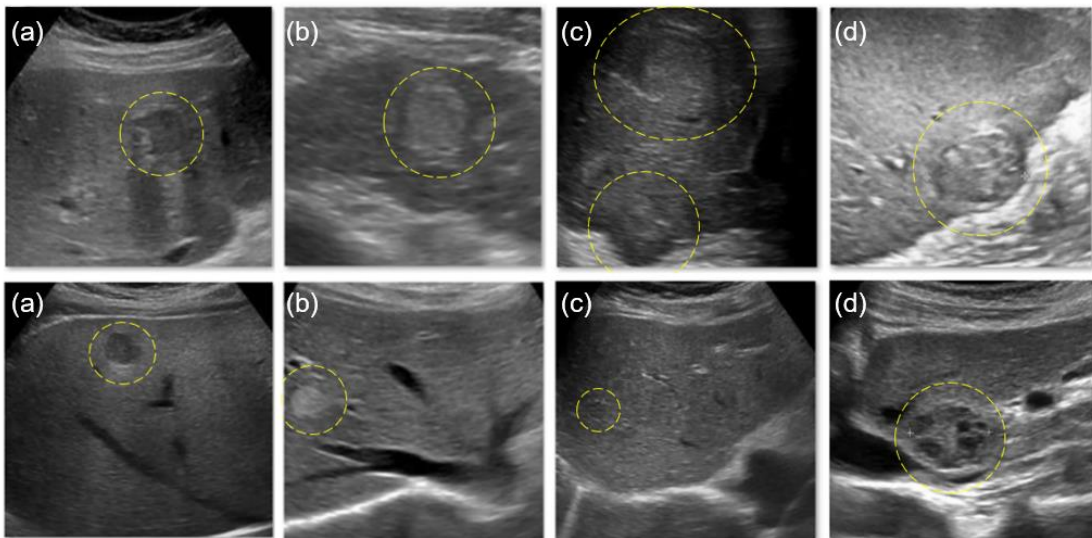


Figure 5 Echogenic patterns of hepatocellular carcinoma (HCC) and hepatic hemangioma (HEM) lesions. The top and the bottom rows show the echogenic patterns of HCC and HEM lesions, respectively. (a) hypoechoic, (b) hyperechoic, (c) isoechoic, and (d) mixed echoic.

In clinical practice, patients who have a risk of being HCC and have been found with a liver tumor larger than 1 centimeter in diameter are sent for a CT or MRI scan to determine whether the tumor is cancerous or not (Additional information can be found in topic 2.1.5). If the found tumor presents typical characteristics of a hemangioma, such as small size, uniform hyper-echogenicity, and well-defined margins, follow-up ultrasound in the next 3-6 months is recommended. On the other hand, if the found tumor presents other characteristics such as atypical hemangioma, suspicious tumor, or resembles HCC, patients are recommended for further investigation.

Ultrasound is primarily used as a tool to detect liver tumors and refer patients to other diagnostic tools accordingly.

2.1.7 Machine Intelligence

In the last ten years, visual information for the healthcare system has exponentially increased. However, the number of radiologists does not keep up with

the increase in diagnostic images. Computer-assisted diagnosis (CAD) has been introduced to help radiologists rapidly diagnose irregularity. Three common terms in machine intelligence are artificial intelligence, machine learning and deep learning.

A. Artificial Intelligence (AI)

AI is the development of computer systems to imitate human intelligence. It has been applied in many tasks in many disciplines, such as visual perception, investment, language translation and decision-making. AI is the superset of machine learning (ML) and deep learning (DL) as depicted in Figure 6.

B. Machine Learning (ML)

ML is the subset of AI. It is an approach to achieving AI. The aim is to have a computer learn to do a task without hard coding how to do it. How well the task is performed is measured by a mathematical function, called the objective function. The parameters inside the program are automatically adjusted so that the system achieves the highest performance according to the given data (training data). Examples of ML are Support Vector Machine (SVM), Artificial Neural Network (ANN), Deep Learning (DL), etc.

C. Deep Learning (DL)

DL is one of the methods for ML. It differs from the conventional ANN in that it contains a higher number of hidden layers (Figure 7). The more complex architecture allows DL to perform more complicated tasks than conventional ANN. DL has been successfully implemented for melanoma classification [22, 23].

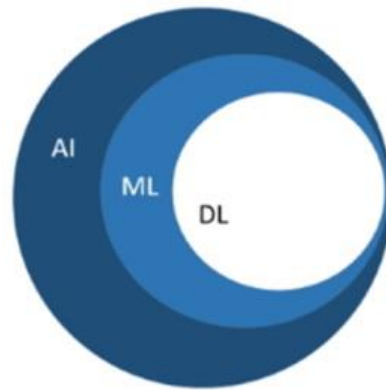


Figure 6 AI encompasses the development of intelligent machines, ML enables computers to learn and make predictions from data, and DL leverages deep neural networks for complex pattern recognition tasks.

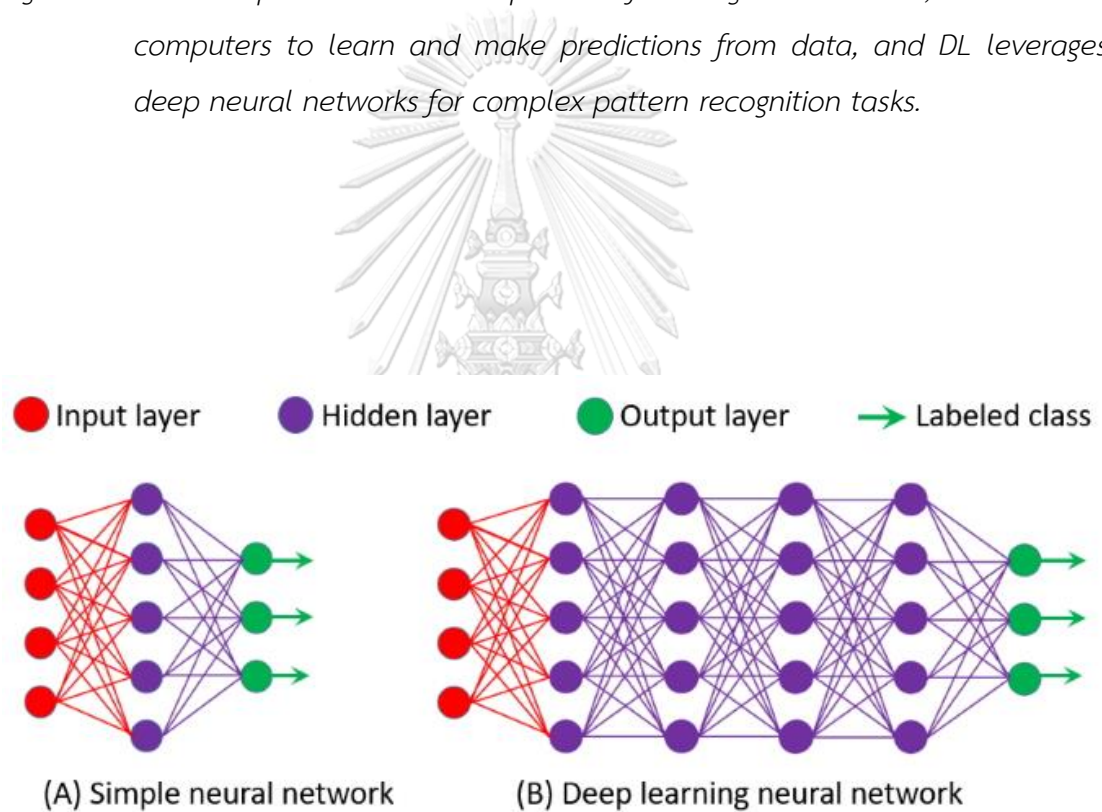


Figure 7 ML works as a simple neural network (A) only a few hidden layers (B) DL works as multiple hidden layers.

2.1.8 Convolutional Neural Network (CNN)

A convolutional neural network is an algorithm that consists of multiple layers. However, it is different from the conventional ANN in that its data are organized into 3-D tensor-like structures. It consists of four major layer types as follows.

- Input layer accepts the image as a 3-D structure where the depth (the third dimension) is the color channel. It is used to pass the information, so there is no function applied to the input.

- Convolution layer is used to extract image features. The layer of the size of $n \times n \times m$ can be considered as m different $n \times n$ filters, where the filter parameters (kernels) are learned from the training pattern instead of fixed value as in conventional filters. Each $n \times n$ filter is applied to its input from left to right and top to bottom. The filter can be applied to every pixel as well as skip with a certain stride value. m filters produce m outputs and are structured as the tensor with the depth of m .

- Pooling layer is applied to subsample its input. Its major role is to reduce the overfitting problem as well as the number of the required parameters. It is usually placed after the convolutional layer. The subsample is performed on each depth separately, so the output of the pooling layer has the same depth as the input, but smaller width and height.

- Fully connected or dense layer is usually the final layer of CNN, where all features all connected to form the representation of the required pattern. The function of this layer depends on the application. For example, 'Softmax' for the classifier (one type per input), sigmoid for multiple classifiers (multiple types per input).

CNN has been shown to be a powerful image classifier. Popular CNN includes Alexnet, GoogLeNet, VGGNet, ResNet.

2.1.9 ResNet

ResNet or residual network made its debut in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) and was able to win the ILSVRC2015 competition. After that ResNet became the neural network of the 21st century. Its top-5 error was 3.57%, which was less than the average human error of 5. ResNet has become a pretrain model that is widely accepted and used today.

ResNet, a kind of artificial neural network (ANN) based on biologically inspired computational networks, is a computational model consisting of many layers, input layer, several processing layers and deliver outputs. ResNet contains nonlinearities activation function called the rectified linear unit or ReLU.

ResNet-50 is a variant of ResNet models which has 48 convolution layers along with 1 max pool and 1 average pool layer as shown in Table 1 [24].

ResNet-50, originally trained on ImageNet with standardized 224 x 224 pixels images, requires resizing of images during training and testing for several reasons. First, resizing ensures input size consistency as ResNet-50 has a fixed input size. By resizing all images to the same resolution, the network maintains uniformity during both training and testing stages. Second, since ResNet-50 is often utilized as a pre-trained model with weights learned from ImageNet, resizing the images to the same size used during pre-training allows for effective utilization of these pre-trained weights. Lastly, resizing images to a smaller resolution, such as 224 x 224 pixels, enhances computational efficiency by reducing the computational requirements during training and testing. Smaller images necessitate fewer computations, resulting in faster processing and efficient utilization of memory. It is worth noting that while ResNet-50 typically requires a 224 x 224 input size, other network architectures or custom implementations may have different input size requirements based on their design and the training dataset used.

Table 1 Architectures for ResNet-50 (Adapted from [24]).

layer name	output size		50-layer
conv1	112×112	7×7, 64, stride 2	1
conv2_x	56×56	3×3 max pool, stride 2	9
		$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	
conv3_x	28×28	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	12
		$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	
conv4_x	14×14	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	9
		$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	
conv5_x	7×7	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	9
	1×1	average pool, 1000-d fc, softmax	1

2.1.10 Object Detection Techniques

In the learning process of the model for lesion localization in liver ultrasound images, object detection plays a significant role. Object detection is a computer technology that involves image processing, enabling computers to detect objects and classify them into different classes. This process encompasses two distinct tasks: classification and localization. Each object class possesses unique features that aid in accurate classification. Convolutional Neural Networks (CNN) are commonly employed in classification tasks due to their exceptional feature extraction capabilities [25, 26].

For localization, the selection of essential features from the "ground truth" becomes a critical step. Ground truth refers to the precise and accurate annotation of objects in the image database. It involves drawing bounding boxes around the objects of interest, assigning them class labels, and capturing relevant information such as their boundaries and pixel intensities. Objects within the same class exhibit similar boundaries and pixel intensities, which aids in their classification. The models learn to extract the ground truth from the labeled data provided by researchers. In

this particular study, the ground truth includes three labels for hepatic lesions: hepatic cyst, hepatic HEM, and HCC.

The models not only learn the features of the objects but also grasp the concept of object localization. Once the models have learned the ground truth bounding boxes, establishing an object detection network becomes crucial for successful lesion localization. This network utilizes the learned ground truth to accurately identify and classify objects of interest in liver ultrasound images. By leveraging the ground truth annotations, the model can effectively detect and localize hepatic lesions, enabling accurate diagnosis and analysis in medical applications.

A. Two-Stage Object-Detection Network.

The two-stage detector consists of the detector and the classification stage. R-CNN family models are among the widely adopted networks. It can be described as follows.

- Detector stage is used to extract the region of interests (ROI). In the original R-CNN [27], selective search is applied to find 2000 ROIs. In the fast R-CNN [28], CNN without fully connected layer is applied to extract the feature. Selective search is applied to find ROIs from the feature map. In the faster R-CNN [29], the selective search is replaced by the tiny network called region proposal network.
- Classification stage, the output of the detector stage is warped to a fixed size. The output consists of both the class and the bounding box for the object. For the R-CNN family, the CNN in the classification stage is a shallow network.

B. Single-Stage Object-Detection Network.

Two-Stage Object-Detection network has the disadvantage of time-consuming from sliding window to find an object. To solve the problem, a

single-stage object-detection network was used. A single neural network trained end-to-end learning, takes directly input images and predicts bounding boxes and labels for each bounding box straightly. The technique offers faster but lower predictive accuracy. Firstly, the models work by splitting the input image into a grid of $S \times S$ cells, such as 7×7 cells, where each cell predicts a bounding box. The width and height of the bounding box are related to the information of ground truth entered. Each grid cell may predict 2 bounding boxes. If the bounding box map to a class, it is labeled as the class. In an input image, the object detections in each grid cell are predicted simultaneously, thus saving time. The YOLO (you only look once) family used these techniques.

C. Region Proposal Network and Bounding Box Regression.

Multi-Stage Object-Detection network works by creating a small window at the top left corner of the ultrasound images and moving the window by sliding the window 1 or 2 pixels to the right. The models were trained with a binary classifier, which determines whether the presented object was "positive" or "negative". Then, they generate a region proposal as a bounding box and perform classification and regression on the bounding box. The models would produce a list of object categories present in the ultrasound image. If a bounding box aligned with the position and scale of each object category, the models presented object 'positive'. The sliding window is a method used to object detection tasks in the early stages, and it is time-consuming but more accurate. The region-based convolutional neural networks (R-CNN family) used these techniques.

D. Intersection over Union (IoU)

IoU is a metric used to define positive or negative in object detection using IoU threshold for making decisions. IoU can be computed from the area of intersection between predicted bounding box and ground truth bounding box over area of union as shown in Figure 8. In the system if model was trained with low IoU threshold, a lot of noise were allowed to be detected [30].

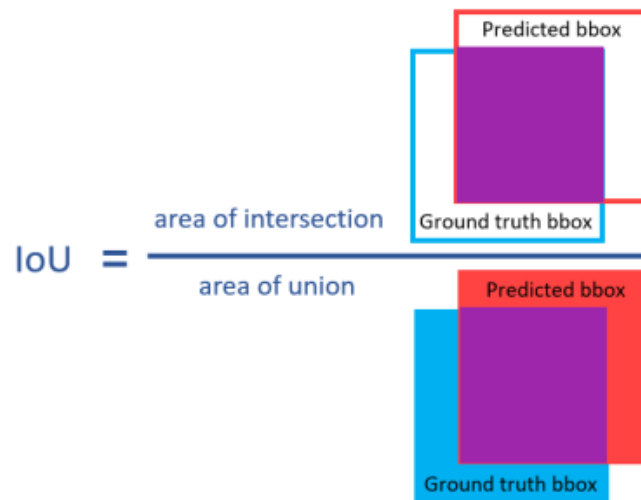


Figure 8 The image shows IoU calculation.

Setting the appropriate threshold allows the model to capture objects effectively. Generally, the default IoU threshold is greater than or equal to 0.5, so it is counted as the object detected.

E. Mean Average Precision (mAP)

Mean Average Precision (mAP) is a metric that is commonly used to evaluate object detection models such as R-CNN, YOLO, etc. The mean of average precision values is calculated over recall values from 0 to 1. It is shown as a graph as shown in Figure 9. Then, mAP is calculated area under the curve as shown as Equation (1)

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (1)$$

Where i represents the number of test data from 1 to N number

AP represents average precision that testing at number i

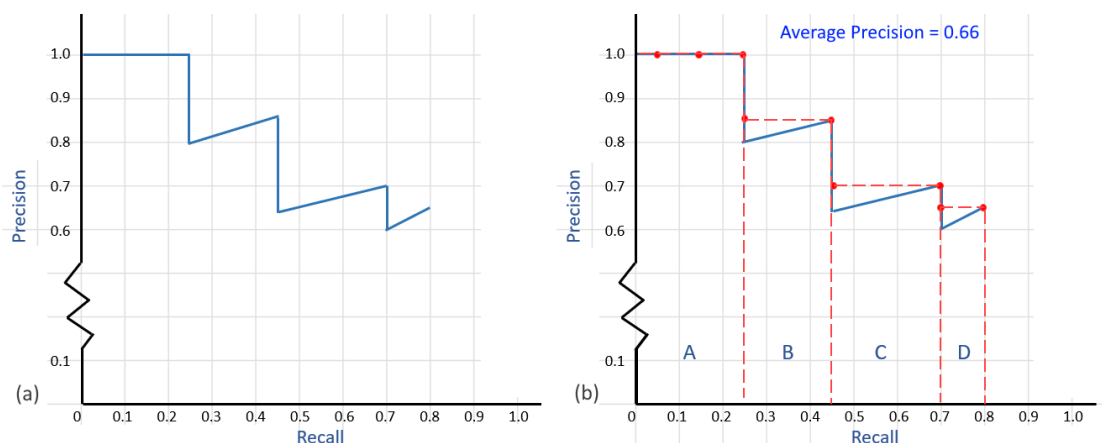


Figure 9 (a) represents a precision – recall curve (b) shows mAP could be calculated by summarization area A to area D.

$$\begin{aligned}
 \text{mAP} &= A + B + C + D \\
 &= (0.25 \times 1) + (0.20 \times 0.85) + (0.25 \times 0.70) + (0.1 \times 0.65) \\
 &= 0.25 + 0.17 + 0.175 + 0.065 \\
 &= 0.66
 \end{aligned}$$

F. The confidence score

The confidence score represents the algorithm's level of certainty or confidence that the predicted bounding box contains an object of a particular class. It reflects the model's belief that the detected object truly exists in the specified location.

Typically, the confidence score is represented as a value between 0 and 1, where 1 indicates high confidence or certainty, while 0 indicates low confidence. During the inference process, YOLOv4 predicts multiple bounding boxes with associated confidence scores for various objects present in the image.

The confidence scores can be used to filter out detections based on a threshold value. For example, if the confidence score for a detected object is below a certain threshold (e.g., 0.5), it might be considered as a false positive and disregarded in subsequent processing steps.

Overall, the confidence scores in YOLOv4 provide a measure of the model's confidence in its predictions and assist in decision-making regarding the presence and accuracy of detected objects.

2.1.11 Object Detectors

Object detection is an advanced form of image classification by pointing out in the form of bounding boxes. It refers to the detection and localization of objects in the image. The models that have the ability of object detection known as object detectors and the well-known object detectors involved in this study are R-CNN and YOLO.

A. R-CNN

The region-based convolutional neural networks or R-CNN that stand for 'regions with CNN features' developed by Ross Girshick, et al. [28]. R-CNN is a two-stage object detector that can work with deep convolutional network to localize and classify object proposals. Multi-stage object-detection has presented more accuracy in object detection but more time-consuming than one-stage [31-33].

B. YOLO

YOLO is a one stage object detector or can be called an end-to-end neural network. YOLO makes predictions of bounding boxes and compute probabilities all at once using IoU and threshold. YOLO has become a famous object detector among the computer vision community due to its speed along with good accuracy. Then, YOLO has been released several versions and YOLO 3 versions (v.4, v.5 and v.6) had released in continuing each year in 2016 - 2018.

Muhammed Enes Atik et al studied on the DOTA aerial photograph dataset consisting of 9 classes object. YOLOv2 has also been used and published the most in 2018 [34]. Moreover, YOLOv2 gave results in 5 out of 9

classes better than YOLOv3 [35]. Additionally, YOLOv2 gave highest mAP among not only Fast R-CNN, Faster R-CNN, SSD300, SSD500, and YOLO [36].

YOLOv4 had been used as object detector on Microsoft Common Objects in Context (MS COCO) dataset. Among several methods: Learning Rich Features (LRF), Receptive Field Block (RFB), Single shot multibox detector (SSD), YOLO3 even Focal Loss for Dense Object Detection (RetinaNet) and etc., YOLOv4 gave highest mAP [37]. YOLO version 4 was used in this study.

2.1.12 Statistical Analysis and Model Evaluation

Model evaluation is a critical process in the development of machine learning models. In this study, the performance of our models on two tasks: lesion detection and lesion classification, were assessed. In the following paragraphs, we will present the parameters that were evaluated in this thesis.

A. Statistical Parameters

A confusion matrix was used to calculate sensitivity (true positive rate, TPR), specificity (true negative rate, TNR), and accuracy. Receiver operating characteristic (ROC) curves were plotted, which is a graph of sensitivity (Y-axis) versus 1-specificity (false positive rate, FPR, X-axis). The area under the ROC curve (AUC) was used to compare the performance of different systems.

The following metrics were computed for each task: %Sensitivity or Recall, %Specificity, %Accuracy, F1-score, Positive predictive value (PPV) or Precision, Negative predictive value (NPV), ROC, and AUC. These scores were computed as the mean of N-fold cross-validation on the training set, and the data was presented as mean and standard deviation.

The %sensitivity, %specificity, %accuracy, F1-score, Positive predictive value (PPV) and negative predictive value (NPV) are defined according to Equation (2) – (7).

$$\%sensitivity = \frac{TP}{TP + FN} \times 100, \quad (2)$$

$$\%specificity = \frac{TN}{TN + FP} \times 100, \quad (3)$$

$$\%accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100, \quad (4)$$

$$F1 \text{ score} = \frac{2 * precision * recall}{precision + recall} \quad (5)$$

$$positive \text{ predictive value} = \frac{TP}{TP + FP} \quad (6)$$

$$negative \text{ predictive value} = \frac{TN}{TN + FN} \quad (7)$$

Here, TP, TN, FP, and FN represent the number of true positives, true negatives, false positives, and false negatives, respectively. Sensitivity (also known as recall) refers to the model's ability to correctly identify true positive cases. Positive predictive value (also known as precision) measures the proportion of true positive cases out of all predicted positive cases.

B. N-fold Cross Validation

To ensure robust evaluation, we employed N-fold cross-validation. Specifically, the dataset was partitioned into N subsets, and the model was trained and evaluated N times, with each subset serving as the test set once. The mean values of the aforementioned metrics were computed from the results of these iterations, providing an overall performance estimation.

C. Variation Range

The evaluation of model performance goes beyond traditional metrics, such as accuracy, precision, recall, F1-score, NPV, and mAP recall rate. To gain a comprehensive understanding of the model's performance, we incorporated

the analysis of the variation range in addition to calculating mean values using a 10-fold cross-validation approach.

The mean values of the performance metrics were computed by averaging the results obtained from the 10 iterations of the cross-validation process. Furthermore, we examined the variation range in percentage, which captures the differences between the mean value and the minimum or maximum values obtained from those 10 iterations. This analysis enables us to explore more deeply the model's consistency and robustness.

Interpreting the variation range provides valuable insights into the model's performance. A large variation range from the mean percentage suggests that the model's performance varies significantly across different iterations. This lack of consistency implies that the model may not be reliable and could exhibit unpredictable behavior in real-world scenarios.

Conversely, a small variation range from the mean percentage indicates that the model consistently performs close to the mean value across different iterations. This stability in performance indicates that the model can be trusted to make accurate predictions consistently, enhancing its practical utility.

By considering both the mean values and the variation range, we can obtain a more comprehensive evaluation of the model's performance. This approach provides researchers with a deeper understanding of the model's behavior, highlighting its strengths and weaknesses and guiding future improvements.

2.2 Review of related literature

Various AI techniques have been applied for the segmentation, detection, and classification of medical images. Table 2 summarizes some of those works.

Though DL had shown great potential in image classification, AlexNet, ZFNet, GoogLeNet and ResNet, DL won the ‘Imagenet Large Scale Visual Recognition Challenge (ILSVRC) in 2012, 2013, 2014 and 2015 respectively. The adoption of DL in medical image processing faces difficulty in the lack of training data. Zhang et al. [38] showed only 10 images per class which were augmented into 500 patches by moving windows and lead to 2000 training examples for 4 classifications. The available images may also be biased toward some classes.

In recent years, there has been a significant amount of research focused on applying artificial intelligence (AI) to liver ultrasound images. In 2017, Tarek Hassan et al [39] employed a stacked sparse auto-encoder architecture to segment and classify liver images into four classes: cyst, Hem, HCC, and normal. Their study demonstrated promising results. Similarly, in 2019, Schmauch et al [4] utilized a multi-step approach involving binary class models to classify liver lesions into six classes: normal liver, cyst, HEM, HCC, liver metastasis, and FNH. Their studies also achieved excellent outcomes. However, it is worth noting that both studies utilized relatively small datasets, with fewer than 50 images in each class of liver lesion, as indicated in Table 3.

Table 2 Summary of works using AI in medical imaging analysis.

Authors	AI Technique	Application
Jamieson et. al. [11]	Adaptive Deconvolutional Networks (ADN)	Breast tumor classification in the US AUC = 0.83 (95% CI)
Wu et. al. [40]	Convolutional Neural Network (CNN)	Classification of fetal abdominal standard plane. Accuracy > 91%, Sensitivity > 94%, Specificity > 81%
Azizi et al. [41]	Deep Belief Network (DBN)	Detection and grading of prostate cancer AUC > 0.70, Accuracy = 70%, Sensitivity = 70%, Specificity = 71%
Zhang et al [38]	Convolutional Neural Network (CNN)	Fatty liver classification (normal, low-grade, moderate grade and severe fatty livers) Accuracy = 90%, Sensitivity = 81%, Specificity = 92%
Hassan et al. [39]	Stacked Sparse Auto-encoder (SSAE) and Softmax classifier	Classification of Focal Hepatic lesion (FLL) Accuracy > 93%, Sensitivity > 95%, Specificity > 92%
Schmauch et al. [4]	ResNet and DenseNet	Detection and classification of FLL to malignant and benign lesions as well as differentiation between six types of liver lesions, AUC > 0.89
Shen et al. [42]	Convolutional Neural Network (CNN)	Detection breast cancer in full-field digital mammography (FFDM) images AUC = 0.98, Sensitivity = 86%, Specificity = 96%
Monsi et al. [43]	ResNet	Classification of 14 thoracic diseases in 112,120 chest X-rays images Accuracy > 90%,

Table 3 The distribution of available lesions in Hassan et al. [39] and Schmauch et.al. [4].

Classifications	Number of US images	
	Hassan et al. [39]	Schmauch et al. [4]
Homogeneous liver	16	258
Hepatic cyst	44	30
Hepatic hemangioma	18	17
HCC	32	6
Liver metastasis	-	48
Focal nodular hyperplasia	-	8
Total	110	367

In 2021, Tiyyarattanachai et al [5] reported that they used ResNet-50 followed by RetinaNet to learn the same amount of normal liver and hepatic lesion images totally 40,397 images with 2414 HCC images. Hepatic lesions consisting of cyst, HEM, focal fatty sparing (FFS), focal fatty infiltration (FFI), and HCC were collected from 3 hospitals in Thailand. Their model could detect and diagnose the hepatic lesion with high performance. They also reported the model had sensitivity for HCC of 73.6%.

RetinaNet, an object detector, is a one-stage object detector, has focal loss layer that could solve the class imbalance problem and correct misclassified example. It has Feature Pyramid Network (FPN) structures as shown in Figure 10, which can only be gave more accuracy than two-stage detector but also spent less computation time as one detector [44, 45]. Moreover, the RetinaNet has also achieved the highest AP winner among Faster R-CNN, YOLO, SSD and DSSD as well [45].

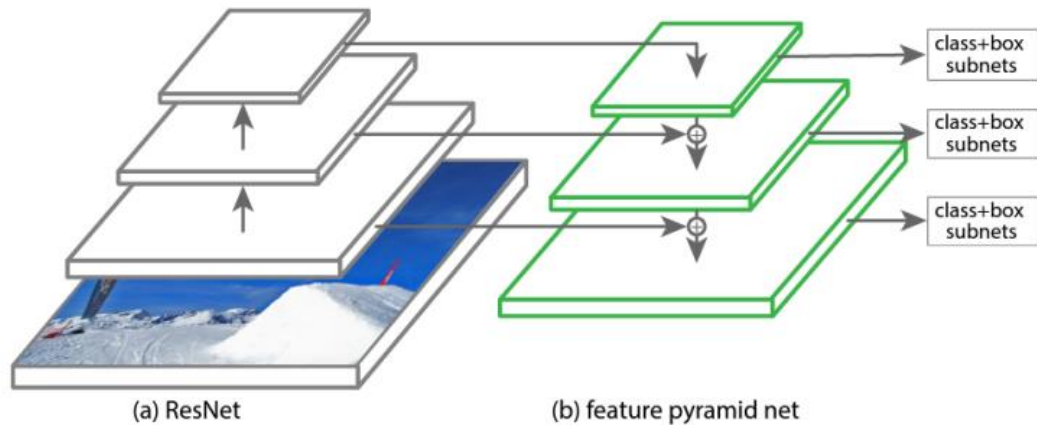


Figure 10 (a) ResNet, a convolutional neural network and (b) feature pyramid net (FPN) is a key component of RetinaNet and plays an important role in the model's ability to detect objects at different scales and resolutions. [45].

2.3 Using YOLOv4 for Classifying HCC and HEM in Ultrasound Images with Limited Dataset

This section of the thesis sheds light on the limitations encountered when utilizing AI models, specifically YOLOv4, for the classification of HCC and HEM in ultrasound images. While models like ResNet-50 demonstrated their ability to classify a vast number of objects into numerous classes on natural images, the same level of performance cannot be expected in ultrasound images due to inherent differences.

ResNet-50's remarkable performance on natural images is attributed to its exposure to a massive dataset containing 14,000,000 objects across 1000 diverse classes. Each object in this dataset exhibits various characteristics such as shape, size, color, movement, pose, and other details. The extensive training on this large dataset equips ResNet-50 with the potential to accurately classify objects. However, when transitioning to ultrasound images of HCC and HEM, the available dataset poses a significant challenge.

The ultrasound image dataset used in our study comprises 543 instances of HCC and 1665 instances of HEM, further divided into 704 instances of atypical HEM and 961 instances of typical HEM. While this dataset provides valuable samples for

analysis, it falls short in terms of quantity for training an AI model effectively. To achieve optimal results with models like YOLOv5, which is recommended by 'Ultralytics' for superior training outcomes, it is suggested to have more than 1500 images per class [46]. Unfortunately, in our study, the available dataset only contains a few hundred images per subclass.

The limited size of the dataset poses a challenge for AI models to capture and differentiate the unique characteristics of HCC and HEM beyond human capability. The scarcity of training samples restricts the ability of YOLOv4 to accurately classify and distinguish between these two types of hepatic tumors. Therefore, it is crucial to acknowledge the limitations imposed by the dataset size and consider alternative approaches or strategies to enhance the performance of AI models in this specific context.

The thesis highlights the limitations associated with utilizing YOLOv4 for classifying HCC and HEM in ultrasound images with a limited dataset. The insufficient number of training samples prevents AI models from effectively capturing the distinctive features and characteristics necessary for accurate classification beyond human capability. These limitations call for further exploration and potential improvements in dataset collection or alternative AI techniques for better performance in ultrasound image analysis.

2.4 Application of Modified Object Detectors on ResNet-50 Architectures for Liver Lesion Detection and Classification in Ultrasound Images

In this section, we described the models applied in the study for liver lesion detection and classification in ultrasound images. Two different architectures were utilized: a modified R-CNN on ResNet-50 for cyst detection and a modified YOLOv4 and ResNet-50 for HCC and HEM detection and classification.

A. R-CNN on ResNet-50

For cyst detection, the model architecture combined the R-CNN framework with the ResNet-50 backbone. This modified architecture included specific changes to certain layers to enhance the performance of cyst detection as shown in Figure 11. These changes involved modifications to the network's structure, highlighting how they contributed to improving the accuracy and effectiveness of cyst detection in ultrasound images.

R-CNN is a two-stage object detection architecture consisting of two stages: 1) region proposals with feature extraction, 2) object classification with bounding box regression as shows in Figure 12. Here's a simplified explanation of each stage:

- Region Proposals with feature extraction:

The first stage involved generating regions in an image that could potentially belong to an object (hepatic cyst for this study). This was done using a method called selective search. It worked by segmenting the image into smaller regions based on color, texture, size, and shape, and then combining similar regions to form objects. The result was a set of region proposals, usually represented by rectangular bounding boxes, for each image. After that, a feature vector of length 4096 was extracted from each region proposal using ResNet-50, a CNN. ResNet-50 was first trained on a large dataset for image classification to learn basic image features. Then it was fine-tuned using a smaller dataset specific to the detection task, where the last classification layer was replaced with a new layer to identify object classes. The region proposals were transformed into fixed-sized inputs and fed into the ResNet-50 to extract features.

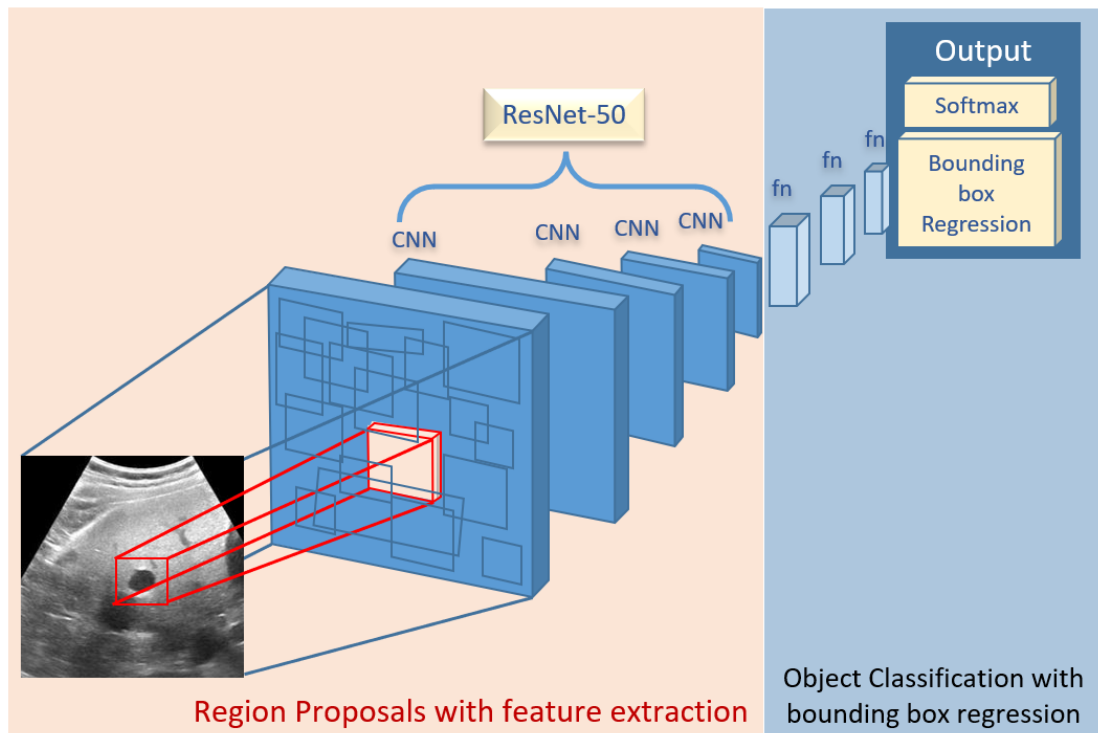


Figure 11 Architecture R-CNN on ResNet-50 for cyst detection. The architecture consists of region proposals with feature extraction and object classification with bounding box regression.

○ Object Classification using SVM with bounding box regression:

After obtaining the features from the ResNet-50, an individual linear Support Vector Machine (SVM) classifier was trained for the hepatic cyst class. The SVM classifier determined whether a cyst class was present or absent in each region proposal. The training labels were defined based on the overlap (IoU) between the region proposals and the ground-truth bounding boxes. Region proposals with low IoU were considered negatives, while those with high IoU were positives. To improve the accuracy of object localization, a bounding box regression step was performed. This step learned corrections to the predicted bounding box locations and sizes. For each positive region proposal, a more accurate bounding box was generated based on the SVM predictions.

The output of the R-CNN architecture was a set of positive object proposals for each class, along with their corresponding bounding boxes. These object proposals were obtained by combining the region proposals, extracting features, classifying objects using SVMs, and refining the bounding boxes through regression.

B. Two-stage Model

The model architecture used in this study for the detection and classification of HCC and HEM involves a two-stage model as shown in Figure 12. The first stage is a hepatic tumor detection model, and the second stage is a classification model. The goal is to differentiate between HCC and HEM, two lesions with various characteristics but shared sonographic appearances. This differentiation is important because there is a dataset available for both types of lesions.

In the first stage, the model combines YOLOv4 and ResNet-50. YOLOv4 is a highly efficient and real-time object detection algorithm that aims to detect all tumors (including HCC and HEM) in ultrasound images. ResNet-50 serves as the backbone network for this detection stage. To optimize the performance of the YOLOv4 and ResNet-50 architecture for the detection task, specific modifications are made to certain layers.

YOLOv4 is an object detection method that utilizes deep convolutional neural networks and includes the Darknet backbone network. It has undergone enhancements compared to YOLOv3, such as the inclusion of a focal-loss function for improved sample balancing and detection accuracy and speed.

In this study, the original Darknet backbone network of the YOLOv4 model is replaced with ResNet-50. ResNet-50 is a residual convolutional neural network architecture known for its accuracy and computational efficiency. The modified model's structure is illustrated in Figure 13, allowing for high accuracy while reducing the number of parameters and computational requirements.

In the second stage, the ResNet-50 model is used as a classifier for the classification task. The transfer learning technique is applied, where the pretrained ResNet-50 model is utilized. The tumors detected in the first stage are passed to this classifier, which then categorizes them into three classes: typical HEM, atypical HEM, and HCC.

To adapt the pretrained ResNet-50 model for the classification task, the last three layers of the original model are removed. These layers consist of a fully connected layer, a softmax activation layer, and a classification layer. In their place, three new layers are added to the model. The first layer is a fully connected layer, followed by a softmax activation layer. Finally, an output layer with a size of 3 is added to accommodate the three classes of data (typical HEM, atypical HEM, and HCC) as shown in Figure 14.

By modifying the ResNet-50 model in this way, it becomes tailored to the specific classification task at hand, enabling it to accurately classify the detected tumors into the appropriate categories.

Different models were utilized for liver lesion detection and classification in ultrasound images. Specifically, the modified R-CNN on ResNet-50 architecture was employed for cyst detection, while the modified YOLOv4-ResNet-50 architecture was used for hepatic tumor detection, and ResNet-50 was utilized for HCC and HEM classification. Detailed descriptions of the architectural modifications in each model would be provided to emphasize how they enhanced the accuracy and performance of liver lesion analysis in ultrasound images.

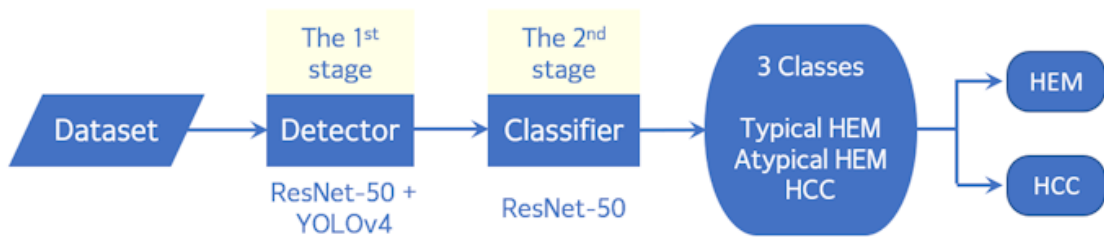


Figure 12 Two-stage model for hepatic tumor detection and classification.

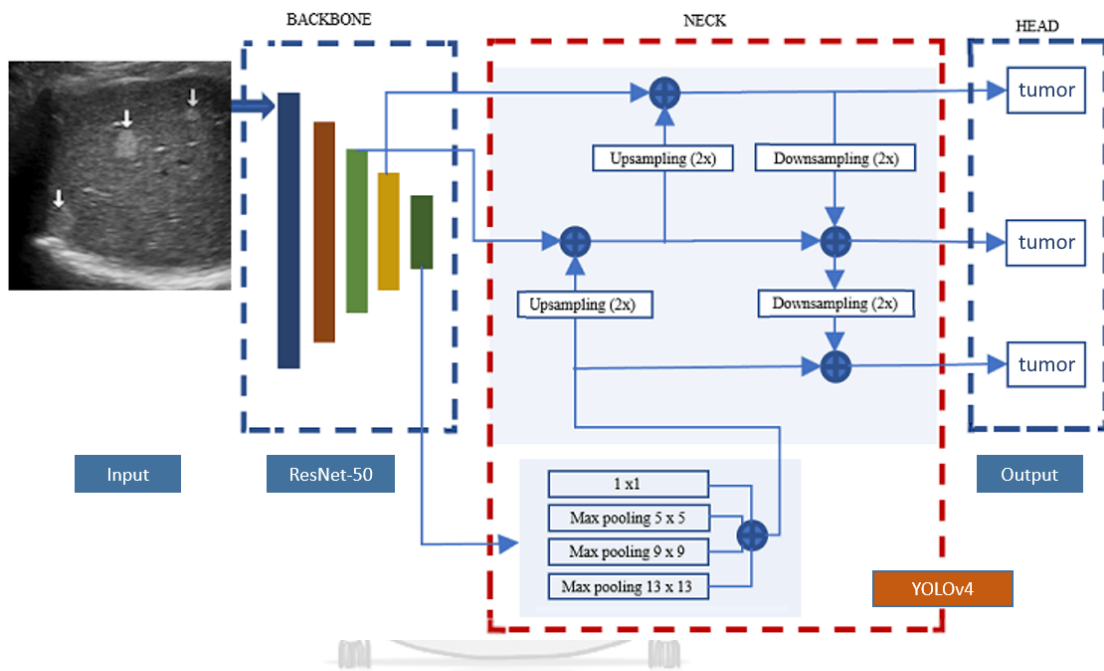


Figure 13 Architecture YOLOv4 and ResNet-50 as backbone for hepatic detection and classification in the first stage.

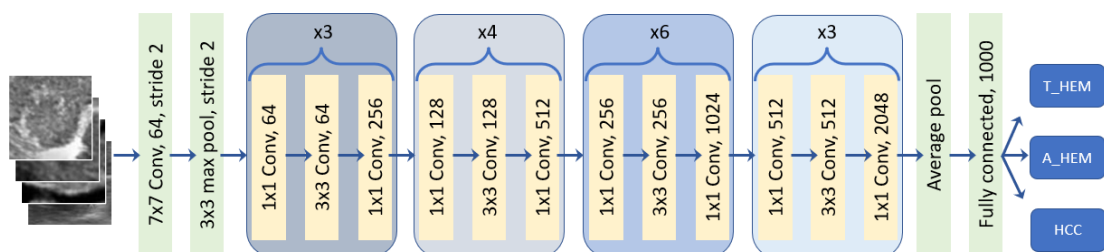


Figure 14 Architecture Resnet-50 classifier for classify hepatic tumor into typical hemangioma (T_HEM), atypical hemangioma (A_HEM), and hepatocellular carcinoma (HCC) in the second stage.

2.5 Normalization Techniques for Liver Ultrasound Images

Normalization is a crucial step in deep learning for medical imaging, specifically in the field of normalization of medical images. It involves scaling the intensity values of the image data to a common range or distribution. This process is essential because it eliminates variations in pixel intensity that can occur due to differences in acquisition protocols, imaging devices, or patient-specific factors.

Normalization has several benefits in medical imaging. Firstly, it enhances model convergence by reducing the impact of varying pixel intensities. Secondly, it helps reduce bias in the image data, allowing for more accurate analysis. Thirdly, normalization improves the generalization capability of deep learning models, enabling them to perform well on new and unseen data. Lastly, it mitigates the risk of overfitting, which can occur when a model becomes too specialized for the training data. Different approaches can be used for normalization in medical imaging, such as z-score normalization, min-max scaling, or histogram matching. The choice of method depends on the characteristics of the data and the specific requirements of the deep learning task.

For CT images, min-max scaling is an effective normalization technique. In this method, the image intensities are mapped to a predefined range, typically between 0 and 1. By doing this, the highest CT number, which represents bone, is set to 1, while other intensities are scaled accordingly. However, min-max scaling may not be suitable for normalizing liver ultrasound images because it is challenging to determine the organ that should have the highest intensity value.

In liver ultrasound images, the top part of the image always represents the abdominal wall as shown in Figure 15, including the skin and subcutaneous fat. The liver itself appears as a medium gray scale. The diaphragm is visible as a white line, but the surface of calcium or some artifacts can also create bright areas. Moreover, the level of the brightest area in each image varies due to different contrast and brightness settings used by ultrasound users. To address these challenges, the intensity normalization method was employed in this study.

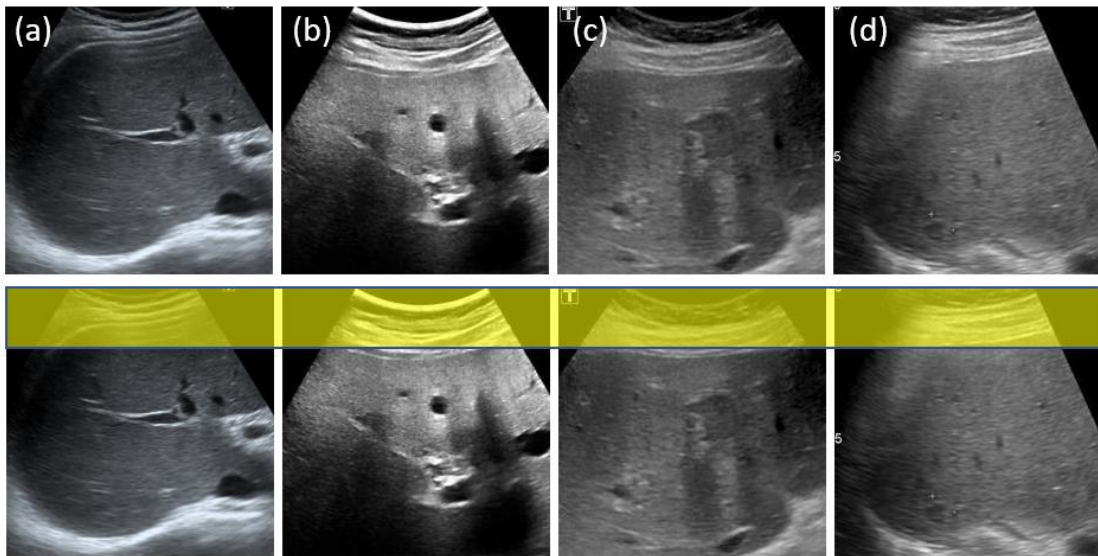


Figure 15 Top row: Liver ultrasound images - (a) Homogeneous liver, (b) Hepatic cyst, (c) Hepatocellular carcinoma (HCC), and (d) Hepatic hemangioma (HEM). Bottom row: Abdominal wall indicated by yellow box.

2.6 Intensity Normalization

During the normalization process of liver ultrasound images, it is important to exclude the non-ultrasound areas, such as the background or setting, from the analysis. Thus, the normalization is focused on the top 20% region of the image, which corresponds to the abdominal wall (Figure 15 bottom row: abdominal wall indicated by yellow box), excluding the black areas.

The adapted intensity normalization has been shown in '3.1.2 Observation Method.' Within this top 20% region, a consistent intensity pattern ranging from 30 to 110 units out of a total of 256 (8-bit depth) is observed. The median intensity value within this region, representing the characteristic intensity level of the abdominal tissues, is set to 70 [4]. By establishing this median as the reference point, the entire image is then normalized accordingly, ensuring that the intensity values of the abdominal tissues align with the desired normalization scheme.

CHAPTER 3

MATERIALS

3.1 Computer System

- (1) A personal computer with an Intel® Core™ i7-8700 CPU, 16 GB RAM, NVIDIA GeForce GTX 1060 6GB Display Card, and MS Windows 10 OS., was used for Cyst Detection.
- (2) A personal computer with CPU: Intel Xeon W-2275 @ 3.30GHz, Memory: 128 GB, Graphics Card: NVIDIA Quadro RTX 5000 16GB, and Operating System: CentOS Linux 7, was used for the hepatic tumor detection and classification.

3.2 MATLAB Software

MATLAB (MATrix LABoratory) is a programming language developed by MathWorks (The Mathworks, Inc., Natick, Massachusetts, USA). MATLAB was built based on engineering, science, and economics. Many toolboxes were created for MATLAB and are available for multi-tasking. For this study, we used the 'Image Processing Toolbox' to analyze, process, visualize and access ultrasound imaging data.

In a crucial process for ground truth labeling, the 'Image Labeler' app provides an easy way to create rectangles as a region of interest (ROI) labels in each image. These labels contain two components:

- The region of object: For this study, the object was considered a hepatic lesion.
- The label name: For this study, we labeled the name "Cyst", "Hem" for hemangioma and "HCC" for hepatocellular carcinoma.

3.3 RadiAnt DICOM Viewer Software

In this study, ultrasound images were obtained from hospitals for further analysis using MATLAB. The DICOM files were first converted to JPEG format to prepare them for analysis. To perform this conversion, we utilized a specialized medical image viewer called 'RadiAnt DICOM Viewer' developed by Medixant in Poznan, Poland. This software allowed us to export the DICOM files as images.

It is important to note that ultrasound imaging systems typically encode RAW DICOM files using 12 bits without compression. However, when exporting ultrasound images, they are usually converted to 8-bit screenshots and saved in JPEG format. These exported images include patient information displayed in the upper part of each image, as well as embedded parameters such as dynamic range, ultrasound frequency, and gray level. When importing the ultrasound images into the PACS (Picture Archiving and Communication System), the users have the option to select different levels of image quality, ranging from 60% to 100% for JPEG files. It is worth mentioning that when exporting the images, the software ensures that the image quality is maintained at 100% without any loss.

One important consideration is that the exported images may not display the complete image without patient information, as the information is typically included in the screenshots.

To summarize, this study involved collecting ultrasound images from hospitals, converting them from DICOM to JPEG format using the 'RadiAnt DICOM Viewer' software, and exporting them without compromising image quality or including any sensitive patient information.

3.4 Dataset

The dataset used for the liver ultrasonography study is presented in this section. The dataset met the inclusion criteria and consisted of liver ultrasound images categorized into four different classes: Homogeneous Liver, Hepatic cyst, HEM, and HCC. The selection of liver ultrasound images was performed based on specific criteria, taking into account the radiology reports of the patients.

To identify cases of HCC or HEM lesions, the radiology reports were reviewed, indicating that these cases had been confirmed through further imaging techniques such as CT or MRI, which followed the HCC screening program. Only ultrasound images corresponding to patients with confirmed HCC or HEM lesions were included in the dataset for analysis.

Regarding hepatic cysts, specific criteria were applied to select ultrasound images of these lesions. The criteria included the presence of an echo-free or anechoic area within the liver, a well-defined margin around the lesion, and posterior acoustic enhancement. Images meeting these criteria were included in the dataset for the hepatic cyst class.

Some of the images in the dataset were captured in a dual pattern, where the ultrasound monitor displayed two windows. To ensure consistency, each image from each side (left or right) in the dual pattern was separated and treated as a single image. Depending on whether the images met the inclusion criteria, one dual pattern could be one or two images.

After the images were collected and the selection criteria were applied, they were categorized into the following four classes based on their nature:

- a. Homogeneous Liver - 2500 images
- b. Hepatic cyst - 615 images
- c. HEM - 1665 images
- d. HCC - 543 images

CHAPTER 4

METHODS

4.1 Cyst Detection

4.1.1 Data Preparation

The data preparation played a crucial role in ensuring the accuracy and reliability of the deep learning models. The dataset used in this study consisted of ultrasound images of the liver, which were collected from various sources and annotated by an experienced sonographer. To address the low image quality and significant noise in ultrasound images, various preprocessing techniques were applied to the dataset, including image normalization, denoising, and resizing.

The imaging dataset comprised ultrasound images obtained from different sonographic equipment and facilities, resulting in variations in pixel size and image dimensions. As part of the data preprocessing, image size normalization was performed to convert the randomly sized images into a standard size. It was observed that ResNet-50, the chosen deep learning architecture, required a minimum input size of 224 x 224 pixels for effective performance. Therefore, all images were resized to a consistent dimension of 224 x 224 pixels, ensuring compatibility with the ResNet-50 architecture.

Furthermore, in order to focus solely on the liver region, which is relevant to the classification task, all images underwent a cropping process to retain only the liver region of interest. This cropping resulted in a standardized image size of 224 x 224 pixels, as depicted in Figure 16. By applying these preprocessing steps, the dataset was appropriately prepared for subsequent training and evaluation of the deep learning models.



Figure 16 The liver image was cropped to be a square.

4.1.2 Observation Method

A hepatic cyst is a type of lesion in the liver that appears as a black or anechoic oval area as same as cross-section of a blood vessel on an ultrasonography. During this study, it was found that the model had difficulty distinguishing cysts and hepatic vessels because both appear black on the ultrasound images [3-5]. However, radiologists and sonographers can differentiate the two structures by looking at the real time scanning of the shape - a hepatic cyst has an oval or round shape while a hepatic vessel has a tubular shape. Additionally, a cyst has posterior acoustic enhancement, which is a brighter region below the fluid-filled structure with increased echoes. In this part of the study, prior knowledge of posterior acoustic enhancement was used to train the model. The aim of the study was to determine whether including the posterior acoustic enhancement below the hepatic cyst could improve the model's detection accuracy.

Six hundred fifteen conventional B-mode liver US images were collected, all containing at least one cyst, and taken by various US machines, resulting in varying image quality. Image preparation was performed for training and detection experiments. Cysts in the training images were labelled by a sonographer with more than 10 years of experience. To improve the detection of hepatic cysts in ultrasound images, some following techniques have been implemented.

- Intensity Normalization

The 224 * 224-pixel image was normalized according to the median of abdominal tissues (shown inside the red rectangle in Figure 17). The range of the intensity median is from 30 to 110. The median is normalized to the average value of 70 [4] according to Equation (8).

$$x_{ij} \leftarrow x_{ij} \times \frac{70}{m} , \quad (8)$$

where x_{ij} and m are the intensity at (i,j) and the intensity median of abdominal tissue, respectively.

- Cyst Ground Truth Labelling

Two training sets were constructed with different regions of interest (ROI). In the first set, the ROI consisted only of the area of a cyst (the solid window in Figure 18A). In the second set, the ROI consisted of both a cyst and its distinct feature, the acoustic enhancement below the cyst (the dashed window in Figure 18B). The first and the second sets were named “Cyst Ground Truth” and “Cyst with Acoustic Enhancement Ground Truth,” respectively.

- Training Phase

ResNet-50 was adopted as the feature extractor, and the last three layers (fully connected, Softmax and the classification layers) were cut out. In R-CNN, selective search was applied to find regions in an image to be used for training. A binary support vector machine (SVM) and a regression model for constructing the bounding box of a detected object were also trained. All networks were implemented in MATLAB 2019a program on the computer with an Intel® Core™ i7-8700 CPU, 16 GB RAM, NVIDIA GeForce GTX 1060 6GB Display Card, and MS Windows 10 OS. Stochastic gradient descent method was used for the training. The mini-batch size was set to 32, and the training rate was fixed at 0.0001. All networks were trained for 10 epochs, where their training converged. Since we wanted to clearly see the effect of the inclusion of the distinct feature, the negative overlapped range was not fine-tuned. The range was fixed between 0 and 0.3.

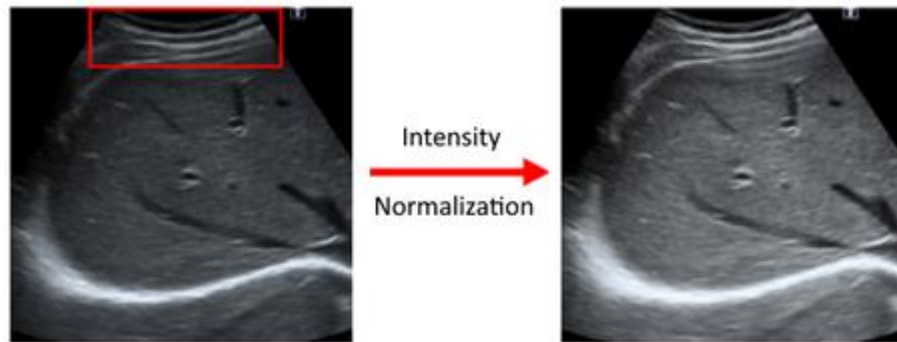


Figure 17 The normalization of the image according to the intensity median inside the abdominal tissues (shown in the red windows of the left figure). The right image is normalized left image to the intensity median of 70.

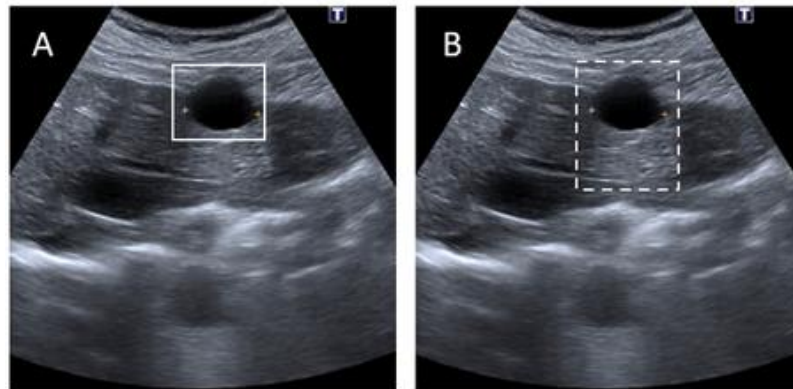


Figure 18 The samples from two training sets. (A) Displays a sample from the Cyst Ground Truth set, which contains an ROI with only one cyst. (B) A sample from the Cyst with Acoustic Enhancement Ground Truth set, which includes an ROI with both a cyst and a bright artifact located below the cyst.

- Cyst Detection Evaluation

Five-fold cross-validation was used to evaluate the classification accuracy. 615 images were randomly separated into 5 sets, each containing 123 images. Four sets were combined to form the training set, and the remaining set was used for testing. Each set was used as the testing set once, so the study was repeated 5 times. It should be noted that some images contain more than one cyst, and all cysts were used for training, so the number of training data was different in each experiment. No attempt to make the number of training data equal in all experiments because cysts from the same image have similar image characteristics. To equalize the number of training data, result in the bias from having cysts in the same image for both training and testing data.

To evaluate the accuracy, the percentage of false negatives, false positives, and true positives have been used. A positive was defined as the detection of a cyst, and the meanings of three parameters were as follows: false negative is the failure to detect a cyst in an image, false positive is the detection of other tissues as a cyst, and true positive is the correct detection of a cyst. In this study, the area of a cyst is considered for the highest confidence in an image and the confidence was larger than 0.5.

Pearson's chi squared (χ^2) statistical hypothesis was used to check whether the parameters were statistically different.

4.1.3 Selection of Detector Models

In this study, R-CNN, YOLOv2, and YOLOv4 were implemented as the detector. The training of R-CNN took much longer than YOLO family but provided higher detection accuracy than YOLO family. The best detection was achieved from R-CNN with ResNet-50 as the backbone.

4.2 Hepatic Tumor Detection and Classification

4.2.1 Data Preparation and Marker Analysis in HEM and HCC

The dataset in this study consists of 1665 HEM and 543 HCC images. The boundary region of HEM and HCC was drawn by a more than 10-year experienced sonographer. In this study, HEM is classified into typical and atypical HEMs. The sonographic appearance of a typical HEM is well-defined, homogeneous, and hyperechoic, whereas the one of atypical HEM is similar to HCC. The number of typical HEM, atypical HEM and HCC in our dataset were 961, 704 and 543 images, respectively. Compared to previous studies [4, 39], our dataset is larger, however, it is still too small to consider HEM and HCC as separate classes in the detector. For example, the recommended size per class is 1500 for YOLOv5 [46].

In the topic 4.1.1 data preparation, which involved resizing all ultrasound images to 224 x 224 pixels, was explained. In contrast to using median normalization, the 'Color Jitter' technique was applied, where the brightness, contrast, hue, and saturation of the images were adjusted by varying them between 10% to 20%. It was found that these augmentations were reasonable because liver ultrasound images can have varying image properties depending on the user, and these adjustments helped to make the dataset more robust.

The reason why the markers were not removed from the ultrasound images in our studies is because the images were obtained from patient records, and some markers were present. However, not all of the HEMs and HCCs were marked, and not all of the markers were for HEMs and HCCs, as shown in Figure 19. Although the presence of a marker could potentially aid the detector, it was believed that its impact should not be significant due to the mixed presentation of markers. Therefore, not to remove the markers from the ultrasound images in this study had been decided.

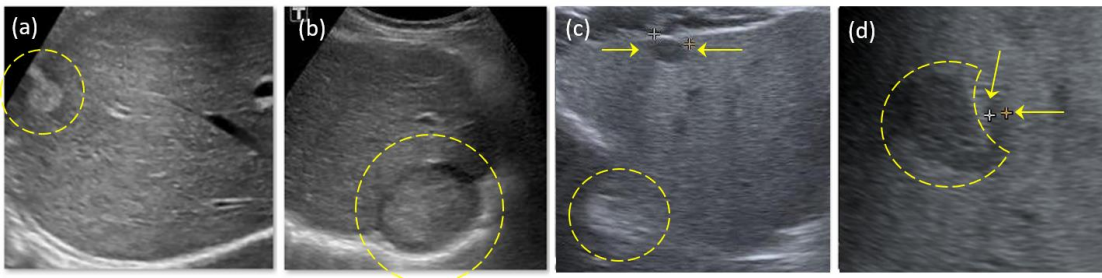


Figure 19 Examples of the lack of markers and the marker for lesions/tissues other than HEM and HCC. HEMs and HCCs were shown inside the dashed circle. (a) HEM without marker, (b) HCC without marker, (c) marker for measuring hepatic cyst in the HEM image and (d) marker for measuring the vessel near the HCC.

4.2.2 Two-Stage Method for Hepatic Tumor Detection and Classification

The proposed two-stage method is depicted in Figure 20. The detector in the first stage was trained to detect HEM and HCC like lesions. Both HEM and HCC images were grouped into one class in the training data, so the size of the training dataset was larger than previous studies where HEM and HCC were separately considered [4, 39]. In addition, the shared sonographic appearance of HEM and HCC can be exploited for better detection.

The detected lesion from the first stage was resized to 224×224 pixels and fed into the classifier in the second stage. The classifier was trained to classify the lesion into 3 classes: typical HEM, atypical HEM and HCC. Though atypical HEM and HCC were considered as two separate classes in the classifier, both would go under further investigation. Thus, in practice, it was not necessary to accurately differentiate HCC from atypical HEM. Atypical HEM mistook as HCC would be ruled out, whereas HCC mistook as atypical HEM would be treated. For screening, a high recall rate is required. Since the classification was separated from the detection, the recall rate can be achieved (at the cost of precision) by adjusting the threshold of a classifier to become biased for HCC class. However, in this study the classifier was not adjusted.

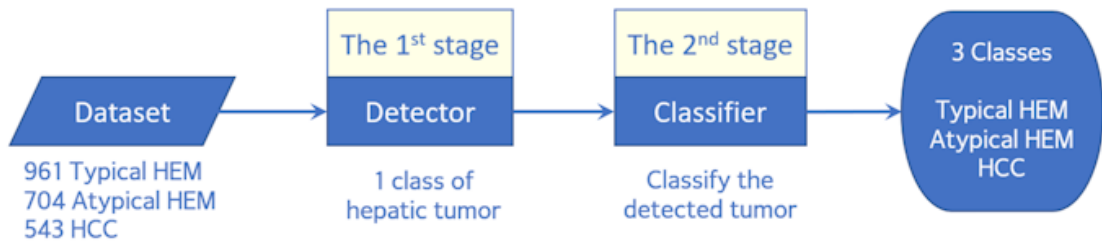


Figure 20 The proposed two-stage method and dataset.

4.2.3 Training Phase for Detector and Classifier

A. Training Phase for Detector

The original Darknet-53 backbone of YOLOv4 was replaced with ResNet-50. To develop four different models, each catering to a specific number of classes (Method 1 for 1 class, Method 2 for 2 classes, Method 3 for 2 classes, and Method 4 for 1 class), ResNet-50 was used as the feature extractor. However, to tailor the models to their respective requirements, the last three layers (fully connected, Softmax, and classification layers) of ResNet-50 were removed and replaced with customized layers suited for each method's specific class configuration. This adaptation allows each model to optimize its final layers for the intended number of classes, resulting in more accurate and specialized predictions.

All networks were implemented in MATLAB 2022 program on the computer with an Intel Xeon W-2275 @ 3.30GHz, Memory: 128 GB, Graphics Card: NVIDIA Quadro RTX 5000 16GB, and Operating System: CentOS Linux 7, was used for the hepatic tumor detection. Stochastic gradient descent method was used for the training. The mini-batch size was set to 32, and the training rate was fixed at 0.00075. All networks were trained for 30 epochs, where their training converged. To assess performance and reduce bias then 10-fold cross validation was used.

B. Training Phase for Classifier

ResNet-50 was adopted as the feature extractor, and the last three layers (fully connected, Softmax and the classification layers) were cut out. Classification layer was changed from 1000 to 3 classes: typical HEM, atypical HEM, and HCC. 10-fold cross validation was used. All networks were implemented in MATLAB 2022 program on the computer with an Intel Xeon W-2275 @ 3.30GHz, Memory: 128 GB, Graphics Card: NVIDIA Quadro RTX 5000 16GB, and Operating System: CentOS Linux 7, was used for the lesion classification. Stochastic gradient descent method was used for the training. The mini-batch size was set to 64, and the training rate was fixed at 0.00075. All networks were trained for 6 – 8 epochs, where their training converged.

4.2.4 Selection of Detector and Classifier Models

In this study, the detector and the classifier were selected from the standard models in the Deep Learning Toolbox of MATLAB R2022a. Region-based CNN (R-CNN) family, Single Shot Detector (SSD) and You-Only-Look-Once (YOLO) family were implemented as the detector. The training of R-CNN family took much longer than YOLO family but provided the same detection accuracy as YOLO family. SSD failed to accurately detect the HEM and HCC like lesions. The best detection was achieved from the YOLO family detector with ResNet-50 as the backbone. The detection accuracies of YOLOv2 and YOLOv4 were approximately the same, but YOLOv4 produced a more precise location (Figure 21). YOLOv4 was adopted as the detector for all models. For the classifier, GoogLeNet, VGG-16, ResNet-18 and ResNet-50 were implemented. The classification accuracy of GoogLeNet, VGG-16 and ResNet-18 were much lower than ResNet-50 so ResNet-50 was adopted as the classifier. All CNN networks were pre-trained using ImageNet database.

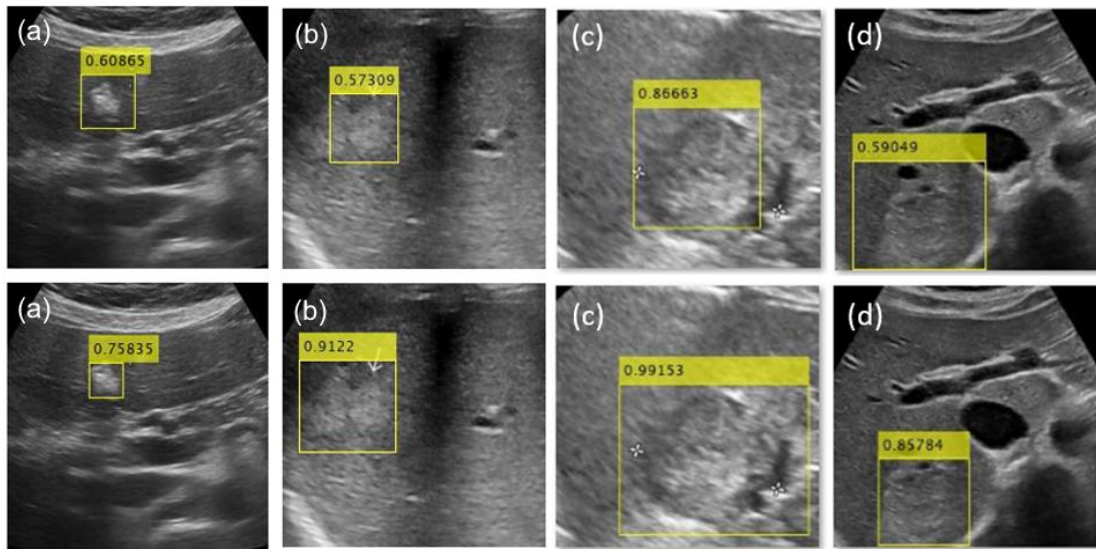


Figure 21 Comparison of detection results using YOLOv2 and YOLOv4. The top and the bottom rows display the detection by YOLOv2 and YOLOv4, respectively. YOLOv4 produced higher prediction scores (the number in the yellow box).

4.2.5 Observation Methods

Four YOLOv4 models were trained. The first one was trained to detect HCC lesion. The second one was the conventional two-class detector and trained to detect HEM and HCC lesions. The third one was trained to detect typical HEM and suspicious lesions. It was the detector-only method that provided the same output as the proposed two-stage method. The fourth model was the first stage of our method and trained to detect HEM and HCC like lesions.

Transfer learning was used to train ResNet-50 classifier. Image augmentation was performed by ± 5 -degree rotation, vertical and horizontal flipping. Only the images in the training set were used to train the ResNet-50 classifier. Additionally, all lesions were resized to 224 x 224 pixels. And only the correct detections from tumor detector were used to test the classifier.

Ten-fold cross-validation was used in all experiments. 90% of the data was allocated for training and validation, and the remaining data was reserved for testing. The experiments were repeated ten times with different testing data. The models' performance was evaluated based on their ability to detect suspicious hepatic tumors.

The performance of the proposed two-stage method depended on the accuracy of both the detector and the classifier, so the study was divided into three parts. In the first part, the detector's performance was investigated, with a focus on the detector's error. In the second part, the classification accuracy of the ResNet-50 model was investigated. In the last part, the overall accuracy of the proposed two-stage method was evaluated against the detector-only method.



CHAPTER 5

RESULTS

5.1 Cyst Detection

The results of the classifier using only Cyst area and Cyst with Acoustic Enhancement Ground Truth sets are shown in Table 4. True positive can be considered as the classification accuracy. The addition of the artifact features helps increase the accuracy from 84.39% to 90.73%, and reduce the errors from both the false negative, 11.20% to 7.64%, and the false positive, 4.39% to 1.63%. The study indicated that the Cyst with Acoustic Enhancement was statistically better than the one without enhancement (p-value = 0.001).

Table 4 Results of the different ground truth labelling for detections.

Ground truth set	Number of images (percentage)		
	False Negative	False Positive	True Positive*
Cyst	69 (11.20%)	27 (4.39%)	519 (84.39%)
Cyst with Acoustic Enhancement	47 (7.64%)	10 (1.63%)	558 (90.73%)

*True positive was considered as the classification accuracy

5.2 Hepatic Tumor Detection and Classification

5.2.1 Performance Analysis of Detectors

This section presents the performance analysis of detectors using YOLOv4 with a ResNet-50 backbone trained under four different settings: a single-task model for detecting HCC, a two-class model for detecting HCC and HEM, a two-class model for detecting typical HEM and suspicious tumors, and a single-task model for detecting HEM and HCC-like lesions. The final model, Model 4, serves as the first stage in the proposed method. While it was not mandatory to use only clinically confirmed HCC and HEM cases for this model, both were classified into the same group. However, to ensure a fair comparison, only the confirmed cases were included. The detection results, including accuracy, precision, recall, F1-score, NPV, and mAP, are presented in Table 5, where the target lesion was considered as positive. The values in the parenthesis were the range of the matrices. Bold indicated the best result for the given evaluation matrix.

Among the performance metrics, particular attention is given to the mean recall values, which are highlighted. The recall range, obtained from 10-fold cross-validation, is indicated in parentheses. The variation range of the recall rate from the four models can be observed in Table 6, enabling evaluation as a percentage by calculating the difference between the Mean and the Minimum percent, as well as the difference between the Maximum and the Mean percent. These performance matrices contribute to the comprehensive evaluation of the detectors and their potential clinical applications in HCC screening.

Table 5 Detection results of four YOLOv4 models.

	Model 1 1 class	Model 2 2 classes	Model 3 2 classes	Model 4 1 class
Dataset	543 HCC	1) 1665 HEM 2) 543 HCC	1) 961 typical HEM 2) 1247 suspicion	2208 HEM and HCC like lesions
Accuracy	0.52 (0.39-0.61)	0.85 (0.73-0.89)	0.72 (0.70-0.77)	0.86 (0.82-0.88)
Precision	0.54 (0.43-0.61)	0.74 (0.50-0.88)	0.88 (0.84-0.93)	0.88 (0.82-0.91)
Recall	HCC 0.67 (0.54-0.86)	HCC 0.64 (0.53-0.97)	Suspicion 0.70 (0.63-0.75)	HEM and HCC 0.84 (0.79-0.89)
F1-score	0.71 (0.64-0.87)	0.68 (0.51-0.93)	0.78 (0.75-0.82)	0.86 (0.83-0.89)
NPV*	0.60 (0.48-0.70)	0.87 (0.82-0.92)	0.72 (0.65-0.76)	0.86 (0.81-0.88)
mAP**	0.50 (0.32-0.68)	0.49 (0.37-0.61)	0.60 (0.51-0.65)	0.76 (0.73-0.84)

* NPV = negative predictive value

**mAP = mean average precision

Table 6 Recall Rate and Variation Range of Four YOLOv4 Models.

	Model 1 1 class	Model 2 2 classes	Model 3 2 classes	Model 4 1 class
Recall	HCC 0.67 (19 – 28%)	HCC 0.64 (17 – 52%)	Suspicion 0.70 (7 – 10%)	HEM and HCC 0.84 (6%)

The first two models were conventional methods in which YOLOv4 was applied to detect HCC and HEM as separate classes. The values in parentheses indicate the variation range compared to the mean value. The results revealed that Model 1, trained specifically to detect HCC, achieved a higher recall rate compared to Model 2, which utilized a two-class model. However, Model 2 exhibited a large variation range, indicating inconsistent performance. The variation range of Model 1 was 19-28%, while the variation range of Model 2 was 17-52%. The implementation of Model 2 to separate HCC and HEM failed to effectively differentiate the two classes. On the other hand, the third model followed the actual clinical protocol, resulting in better detection of suspicious lesions (atypical HEM and HCC) compared to HCC-only detection. When considering the recall rate of the 4 models, Model 4 achieved the highest recall rate with lowest variation range. This indicates that the two-stage approach not only outperformed the detector-only method but also exhibited more repeatability results.

Ultrasound lesions do not present with the same characteristic throughout their area. Parts of typical or atypical HEM may have the characteristics of the other type. To address this issue, all HEM and HCC were grouped into a single class for detection in the fourth model. By doing so, the detector was able to identify the distinct characteristics of both HEM and HCC. Consequently, all evaluation matrices besides negative predictive value improved. Notably, the detector was capable of detecting HCC lesions that the first three models failed to detect, which is demonstrated in Figure 22.

It was possible that the better detection of Model 4 was due to the better HEM detection, not HCC detection, so the detection results were categorized into three classes: HCC, HEM and others. Table 7 shows the distribution of each category. Note that some images contain more than one HEM / HCC and YOLOv4 failed to detect all lesions in those images. Some lesions were detected more than once as shown in the last row of Table 7 and Figure 23. Overall, the HEM and HCC like lesion detector performed better than the other three models. The recall rate for HCC and HEM were 0.78 and 0.86, respectively.

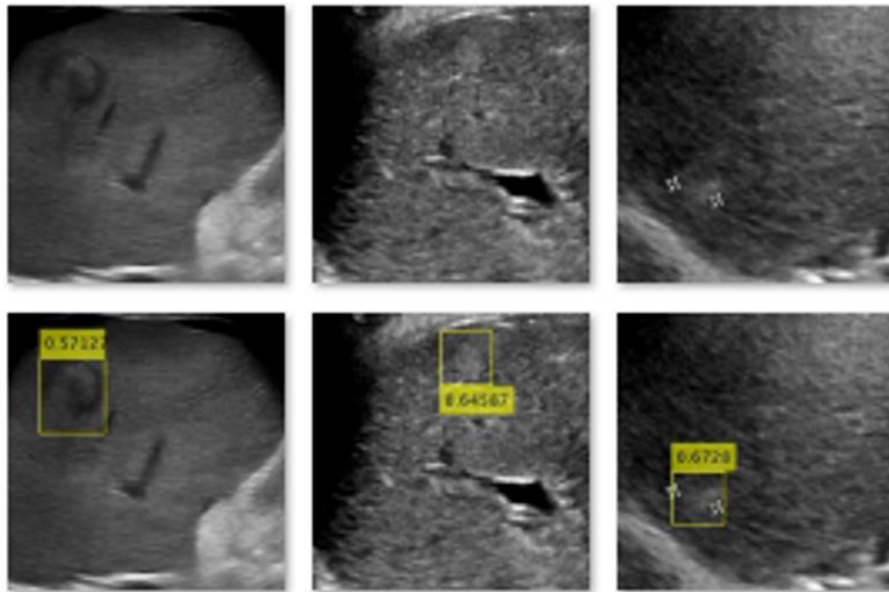


Figure 22 Training YOLOv4 to detect HEM and HCC like lesions as one class resulted in an improved HCC detection, as depicted in the bottom row of the figures. The top row displays the results of the same region when the models were trained to detect HCC or HEM as separate classes, which failed to detect the lesions.

Table 7 The detection results of Model 4 as grouped by lesion type.

Detector model 4	The number of detected lesions (actual value)			
	HCC	HEM	Others	Total*
YOLOv4				
Images	472 (543) 86.92%	1455 (1665) 87.39%	68 (0)	1927 (2208) 87.73%
Lesions	480 (615) 78.05%	1479 (1721) 85.94%	68 (0)	1954 (2336) 83.64%
Lesion + Redundancy	489 (624) 78.37%	1494 (1734) 86.16%	68 (0)	1983 (2358) 84.10%

* Total is the sum of the HCC and HEM only. Other lesions were not considered.

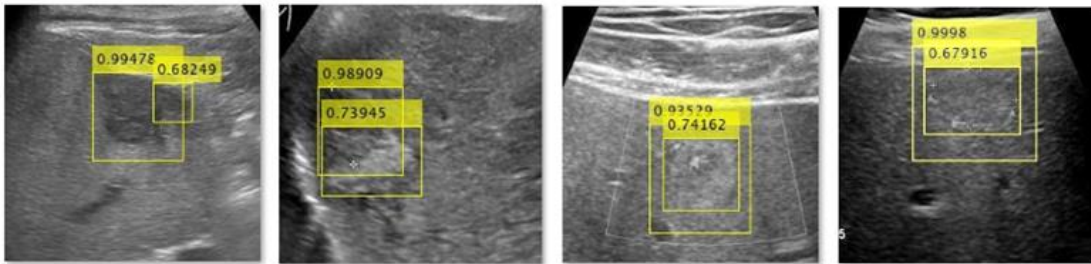


Figure 23 Examples of multiple detection of the same lesion by YOLOv4.

5.2.2 Performance Analysis of Classifier

ResNet-50 was applied to classify HEM and HCC like lesions into three classes: typical HEM, atypical HEM and HCC. The model's performance was evaluated using 10-fold cross-validation, revealing area under the Receiver Operating Characteristic (ROC) curve values ranging from 89% to 94%. In Figure 24, the average area under the ROC curve of 92% is presented. Table 7 provides detailed information regarding the HCC and HEM images. From the 472 HCC images detected by the first-stage detector, a total of 480 lesions were identified. In the case of the 1455 HEM images, there were a total of 1479 lesions, with 570 classified as atypical HEMs (if we consider redundancy, the count increases from 570 to 574) and 909 classified as typical HEMs (if we consider redundancy, the count increases from 909 to 920). Table 8 showcases the confusion matrix of the classification, including redundancy in the count of accurate detections. Note that some HCC were detected multiple times, thus, the number of HCC lesions was 489.

Atypical HEM and HCC share many sonographic appearances; thus, CT or MR examination is required for correct diagnosis. Recently, contrast-enhance ultrasound (CEUS) can also be used and be applied immediately after normal B-mode ultrasound scanning. However, CEUS is not widely adopted in Thailand. Having only ultrasound images, it is impossible to differentiate between atypical HEM and HCC. The further treatment plans for atypical HEM and HCC are the same, i.e., scheduled for CT or MRI examination. Therefore, the misdetection of HCC as atypical HEM did not pose health risk. Table 8 was modified to Table 9 where the HCC incorrectly detected as atypical HEM is accepted as the correct classification.

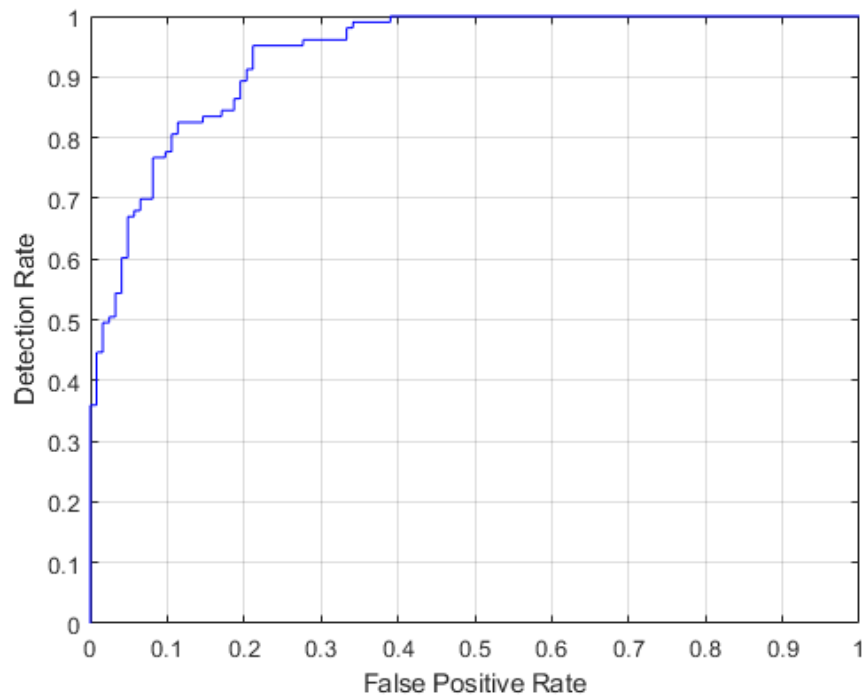


Figure 24 Average area under the ROC Curve as 92% for HEM and HCC-like Lesions.

Table 8 Results of YOLOv4 Detector and ResNet-50 Classifier on 3 x 3 confusion matrix.

Predicted Class	Actual Class				
	Class	HCC	Atypical HEM	Typical HEM	Total
HCC		337	59	98	494
Atypical HEM		111	486	135	732
Typical HEM		41	29	687	757
Total		489	574	920	1983

Table 9 Results of modified 3 x 3 confusion matrix as 2 x 2 confusion matrix.

Predicted Class	Actual Class			
	Class	HCC	HEM	Total
HCC		448	157	605
HEM		41	1337	1378
Total		489	1494	1983

According to Table 9, the 2 x 2 confusion matrix provides valuable insights into the classification performance when considering HCC as positive and HEM as negative. Among the 489 total HCC lesions, 448 (92%) were correctly classified as HCC, representing the true positive rate. The overall accuracy was 0.90, and the negative predictive value (HEM classified as negative) was 0.97. However, when considering HCC as positive, the precision dropped to 0.74 due to the challenges in differentiating HCC from atypical HEM using ultrasound imaging alone. Similarly, out of the 1496 total HEM lesions, 1337 (89%) were accurately classified as HEM, indicating the true negative rate.

However, there were 157 cases where HCC lesions were misclassified as HEM, resulting in a false positive rate of 11%. If these atypical HEM lesions were included as part of HCC instead of HEM, the precision would increase to 0.82, demonstrating the impact of incorporating atypical HEM into the HCC category.

Specifically, 41 HCC lesions were identified as HEM, with 22 of them displaying small hyper-echoic features that resembled typical Hemangioma. This resemblance may be due to capturing ultrasound images when HCC lesions appeared as regenerative or dysplastic nodules, similar in appearance to typical HEM. These misclassifications highlight the challenges in differentiating between atypical HEM and HCC based solely on ultrasound images.

To achieve accurate diagnoses, CT or MR examinations are typically required since atypical HEM and HCC share similar sonographic appearances. Although contrast-enhanced ultrasound (CEUS) is an alternative, it is not widely adopted in Thailand. Therefore, relying solely on ultrasound images makes it impossible to differentiate between atypical HEM and HCC. Consequently, patients with suspected HCC or atypical HEM would undergo further examinations, such as CT or MRI scans, for accurate diagnosis and treatment planning. Misclassifying HCC as atypical HEM does not pose a significant health risk.

5.2.3 Performance Analysis of Two-Stage Model

In this study, the proposed two-stage method was compared with the detector only method (YOLOv4 Model 3). Both atypical HEM and HCC have the same appearance and require further CT or MR examination, so it is unnecessary to differentiate between atypical HEM and HCC. The result of Model 3 in topic 5.2.1 was compared with the proposed two-stage method. HCC and HEM were considered as positive and negative samples, respectively. The incorrect detection of Model 4 was not classified but would be considered as getting negative (HEM) classification. The accuracy and the recall rate were calculated based on the number of actual HCC (not the number of detected area). If an HCC lesion was detected more than once, only one instance classified as HCC was enough for further examination and would be considered as correct. The results are presented in Table 10.

From Table 10, the best results, highlighted in bold, were chosen based on evaluation matrices and came from the final results of Model 4, surpassing the results of Model 3. Overall, the two-stage method (Model 4) performed better than Model 3, except for precision where both methods were equal. However, when focusing solely on detecting HCC, Model 3 had a recall rate of 0.70 for identifying suspicious tumors (atypical HEM + HCC), which dropped to 0.68 for HCC recall rate. On the other hand, Model 4 had a recall rate of 0.84 for identifying hepatic tumors (HCC + HEM) in the first stage, which then dropped to 0.78 for HCC recall rate, and achieved a correct classification rate of 0.92 for HCC in the second stage. Then the two-stage model had an HCC recall rate of 0.72.

Table 10 The results of HCC detection using both single-stage detector and two-stage methods.

	Model 3	Model 4 Two-stage method		
		Detector	Classifier	Overall
Accuracy	0.72	0.86	0.90	0.77
Precision	0.88	0.88		
Recall	Suspicion	HCC+HEM		HCC+HEM
	0.70	0.84	0.90	0.76
F1-score	HCC	HCC	HCC	HCC
	0.68	0.78	0.92	0.72
Negative predictive value	0.72	0.86		
Mean average precision	0.60	0.76		

Furthermore, Nishida et al [47] conducted a comprehensive study on liver classification using multiple AI models. They analyzed a dataset of 94,427 B-mode ultrasound images of liver tumors and focused on classifying four types: cyst, hemangioma, metastasis, and HCC.

Their findings showed a notable HCC sensitivity ranging from 0.65 to 0.68. It is important to note that their study utilized a larger dataset, while our two-stage model with 2208 liver tumor images achieved a slightly higher HCC sensitivity of 0.72. These results highlight the potential for further advancements in liver tumor classification through AI models.

CHAPTER 6

DISCUSSION

6.1 Cyst Detection

During our study, the data was not augmented for training the deep learning model to detect cysts. Nevertheless, our dataset was comparatively large and diverse, with 615 images and a wide range of equipment settings. In fact, when compared to other published articles [4, 39], our dataset stood out in terms of its size and variability, although demographic factors were not considered in this comparison. Through our analysis, four primary causes of errors in the cyst detection process were identified as follows.

- (1) The cysts and vessels were similar. In the ultrasound image, main vessels, hepatic veins and portal veins appear as black area, which was the same as a cyst. Though both classifiers mistakenly detected some vessels as cysts (Figures 25A and 26A), the classifier having the information of acoustic enhancement, incorrectly detected only the vessels that closely resembled the cyst (round with acoustic enhancement). Note that this error was also reported by Schmauch [4].
- (2) Acoustic enhancement is not the unique property of a cyst. Organs such as hepatic capsule, diaphragm and esophagus, etc. were also presented with enhanced edge (bright line). In some cases, the classifier mistakenly detected the bright line as the border of a cyst as shown in Figures 25B and 26B.
- (3) Cysts were not detected in an image with low contrast. As shown in Figure 25C and 26C, a cyst and its posterior acoustic enhancement were not distinct in the low contrast images. Furthermore, the cyst and vessels are more similar in appearance, so the classifier may also incorrectly detect vessels as a cyst. This error can be reduced if better preprocessing has been used, instead of globally normalizing image intensity based on the non-black pixels at the top 20% area. The top 20% area was fixed and may not reflect the real intensity distribution in the image.

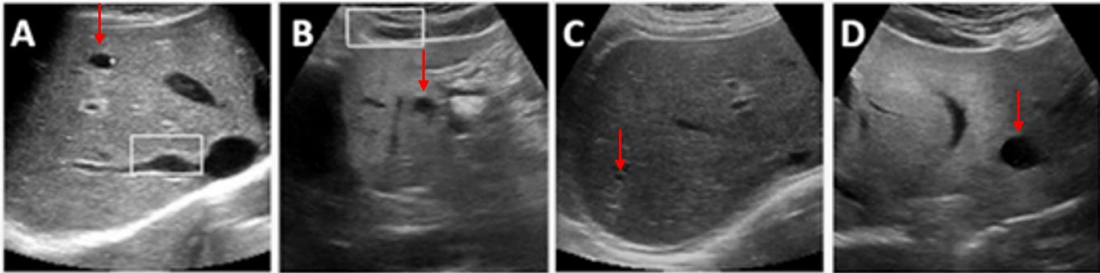


Figure 25 Examples of errors from the classifier trained with Cyst Ground Truth set. Actual cysts were indicated by arrows. The solid box was the cyst detected by the classifier. The classifier detected a branch of right hepatic vein as a cyst in (A) and a part of subcapsular as a cyst in (B). The classifier failed to detect a cyst in an image with low contrast in (C) and a cyst with incomplete border in (D).

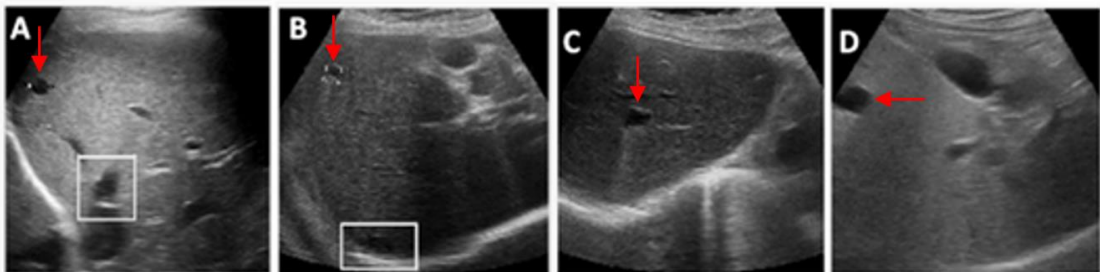


Figure 26 Examples of errors from the classifier trained with Cyst with Acoustic Enhancement Ground Truth set. Actual cysts are indicated by arrows. The windows are the cysts detected by the classifier. The classifier detected a branch of hepatic vein as a cyst in (A) and a part of diaphragm as a cyst in (B). The classifier failed to detect a cyst in an image with low contrast in (C) and a cyst with incomplete border in (D).

- (4) The border of the cyst was missing or blended with the background (hepatic parenchyma) as shown in Figures 25D and 26D.

6.2 Hepatic Tumor Detection and Classification

6.2.1 Analysis and Interpretation of Results in The First Stage Model or The Tumor Detector

In this study, there were four models to detect and classify hepatic tumors in ultrasound images (Table 5). The first model was trained on only HCC images, while the second model was trained on both HCC and HEM images. These models achieved a sensitivity of 0.67 and 0.64, respectively. The third model was trained on typical HEM and suspicious tumor images (which consisted of both HCC and atypical HEM images) and achieved a sensitivity of 0.70 for detecting suspicious tumors. The fourth model was trained on hepatic tumor images, which included both HCC and HEM images, and achieved a sensitivity of 0.84 for detecting tumors.

Our findings suggest that Models 3 and 4 would be more useful for accurately detecting HCCs, as they achieved higher sensitivities than Models 1 and 2. However, if we focus only on Models 1 and 2, our results are consistent with other studies on the detection of malignant tumors in breast ultrasound images, such as Cao et al. [48] and Tanaka et al [49]. These studies also reported relatively low detection rates for cancerous tumors in ultrasound images. It is worth noting that a recent study by Tiyarattanachai et al [5] reported a high recall rate of 0.74 for HCC detection using Retina Net. However, this was achieved by lowering the IoU threshold to 0.2, which may result in more false positives.

If focusing on Model 4 that gives highest sensitivity, there were two kinds of error in the detection. The first kind of error is the failure to detect HEM and HCC like lesions. This kind of error affects the recall of the proposed two-stage method. The second kind of error was the misdetection of other areas / lesions as HEM and HCC like lesion. This error affects the precision. Note that the second kind of error was less than 5% (the data in Table 7, the number 68 from 2208 images) of the total test data and could be easily dismissed by radiologists during the follow-up.

According to Table 7, among 135 undetected HCC, 87% did not have the sonographic appearance of HCC. Most of them were faintly visible, isoechoic tumor or had incomplete border or obscured by low image quality. This false negative problem was also reported by [5]. Nevertheless, the missing lesions were detected in further CT or MR scanning.

In the clinical protocol, if a new lesion appears where nothing was shown in the previous scan, regardless of its appearance, a CT or MR scan is requested because the patient is considered to be at high risk for HCC. Without access to previous records, it can be impossible to detect these HCC. Additionally, once one HCC is detected, the entire liver is scanned using a CT or MR system. Therefore, detecting every HCC in an image has the same outcome as detecting at least one. Based on this, it is possible to conclude that YOLOv4 could detect 86.92% of HCC patients. However, it cannot be guaranteed that at least one HCC will be detected, so all undetected HCCs were considered false negatives in this study. It can be concluded that the HEM and HCC-like lesion detector had a recall rate of 0.78 for HCC detection.

6.2.2 Analysis and Interpretation of Results in The Second Stage Model or The Tumor Classifier

According to Table 8, out of the 41 HCC incorrectly detected as typical HEM, 22 lesions closely resembled typical HEM. These 22 lesions were well defined and hyperechoic (Figure 27). Some of these lesions were detected in CT or MR examinations, because they were either presented (1) in a liver with multiple HCCs or (2) new lesions that appeared in the area without any lesions during the previous ultrasound screenings. Furthermore, some misclassification occurred, because the detector did not extract enough area of the HCC lesion as shown in the leftmost image of Figure 23.

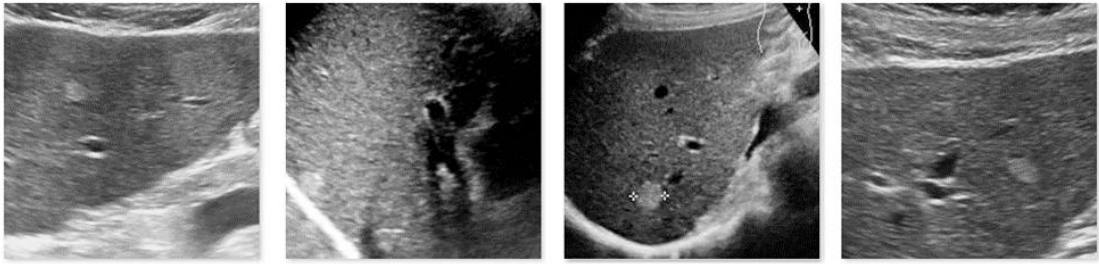


Figure 27 Small oval-shaped hyperechoic lesions resembling HCC were misclassified as typical HEM by the Classifier.

The classifier was trained by the ground truth lesions. The classification result would be better if classifier was also trained with the detection result. However, we would like to evaluate the performance independent of the detector, so the ground truth was used.

6.2.3 Analysis and Interpretation of Results of the Two-Stage Model

The experiment demonstrated the better performance of the two-stage method compared to the detector-only method. The implementation of the two-stage method led to the improvement of HCC recall from 0.68 to 0.72. Both the detector and the classifier were implemented by the out-of-the-box models which were not designed for medical imaging and were not among the latest models. However, the achieved recall rate of 0.72 was only slightly lower than the 0.74 reported by Tiyyarattanachai et al [5]. The dataset in our work is smaller so the number of the training image was much lower (615 HCC vs 2414 HCC). Furthermore, Tiyyarattanachai et al [5] reported the high recall rate by lowering the accepted IoU to 0.2, which was considered incorrect detection in our work. There were other works that demonstrated high accuracy [4, 39]. However, the database was too small to make a solid conclusion.

In the two-stage method, the detector and the classifier are trained separately. YOLO detector has been continuously improved. YOLOv4 is no longer a state-of-the-art detector. YOLOv8 is recently available at <https://ultralytics.com/yolov8>. The quick improvement of the proposed method is to replace YOLOv4 with YOLOv8.

ResNet-50 provided good classification, but it was not optimal. The best result would be achieved if the network was pretrained with medical images. In our study, the network was pretrained by ImageNet database, which contains natural images. It is impossible to train the entire ResNet-50 by a small database, so we currently develop the shallow network for lesion classification in liver ultrasound images.

6.3 Overall Discussion

The goal of this study was to develop an efficient deep learning model for the detection and classification of hepatic lesions on ultrasound images. Several deep learning architectures and transfer learning algorithms were used to train models for detecting cyst, HEM, and HCC lesions. The results showed that training separate models for each lesion can lead to higher efficiency than training a single model for all lesions.

In particular, the first model was trained using R-CNN with ResNet-50 architecture for cyst detection, achieving higher efficiency than using YOLOv2 or YOLOv4. The second model used a two-step approach for tumor detection and classification, with ResNet-50 and YOLOv4 architectures for tumor detection and a ResNet-50 architecture for classification. The 3 x 3 confusion matrix was modified into a 2 x 2 confusion matrix to optimize the model's performance, resulting in high HCC recall rates.

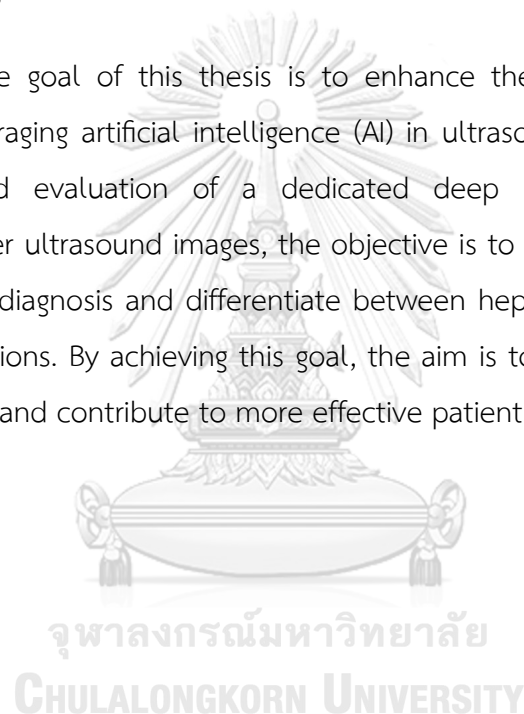
The findings of this study highlight the importance of training separate models for each lesion when detecting and classifying hepatic lesions on ultrasound images. This approach allows for a more specific and targeted approach to each lesion type, resulting in higher efficiency and accuracy. Additionally, the use of deep learning architectures and transfer learning algorithms can improve the performance of these models, leading to more effective diagnosis and treatment of hepatic lesions.

The study also has potential clinical implications. The developed models can be used in clinical practice to assist radiologists and clinicians in the detection and classification of hepatic lesions, improving diagnostic accuracy and patient outcomes.

Moreover, the study demonstrates the potential of deep learning in medical imaging, opening avenues for further research and development in this field.

The study presents a two-step approach using deep learning models for the detection and classification of hepatic lesions on ultrasound images. The findings suggest that training separate models for each lesion can lead to higher efficiency and accuracy in detecting and classifying hepatic lesions. The developed models have potential clinical implications and demonstrate the potential of deep learning in medical imaging.

The ultimate goal of this thesis is to enhance the current practice of HCC screening by leveraging artificial intelligence (AI) in ultrasound imaging. Through the development and evaluation of a dedicated deep learning model designed specifically for liver ultrasound images, the objective is to enhance the accuracy and efficiency of HCC diagnosis and differentiate between hepatic cysts, hepatic vessels, HEM, and HCC lesions. By achieving this goal, the aim is to improve the overall HCC screening process and contribute to more effective patient health care.



CHAPTER 7

CONCLUSIONS

7.1 Implications of Cyst Detection Study

In this study, the effect of the inclusion of the distinct feature around the object of interest to the training data had been investigated. The studies indicated that the inclusion of cyst artifact led to a classifier with higher accuracy. The distinct feature may not be unique to one particular tissue / organ. The classifier may still incorrectly classify other tissues/organs as the object of interest. Therefore, in addition to the area around the object, the image properties of the object should also be investigated and included in the training data.

In addition to the distinct features, the accuracy also depends on the classifier. Currently, the original R-CNN is used. The feature was extracted by ResNet-50 which was pre-trained by non-medical images and fine-tuned for cyst detection. However, medical and non-medical images are greatly different. In this sense, the R-CNN in this study was not the best classifier. The accuracy can also be improved if networks dedicatedly trained or designed for medical images such as Nifty-Net were used.

7.2 Implications in Detecting and Classifying Hepatic Tumor

The proposed method for HCC detection from ultrasound images is a two-stage approach. In the first stage, a detector was trained to capture both HEM and HCC like lesions. In the second stage, the classifier was trained to distinguish among HCC, atypical HEM and typical HEM. Since the future plans for atypical HEM and HCC are the same, it is not necessary to strictly limit the classification of HCC to only HCC. The experiment showed that the two-stage method outperformed the detector-only method in HCC detection.

7.3 Overall Conclusions

The primary objective of this study was to design a feature extraction algorithm for lesion detection and classification in liver ultrasound images, in combination with AI. Through the development and testing of various algorithms, it has shown promising results for lesion detection and classification in liver ultrasound images. Our proposed two-stage model was able to effectively detect and classify HEM and HCC like lesions with high sensitivity and specificity. The use of training cyst with its artifact proved to be a better approach for feature extraction compared to the use of area only cyst in the previous study. The results of this study demonstrate the potential of AI and feature extraction algorithms for improving the accuracy and efficiency of lesion detection and classification in liver ultrasound images, which could have significant implications for clinical practice in the future.

The secondary objective was to investigate the most accurate deep learning structure for hepatic lesion classification when used with proposed the two-stage model. Based on the experimental results, it was found that the combination between detector and classifier using convolutional neural network (CNN) architecture, provided the most accurate and robust classification of hepatic lesions in ultrasound images.

Overall, this study has demonstrated the effectiveness of combining feature extraction algorithms with deep learning techniques for accurate hepatic lesion detection and classification in ultrasound images. These findings may have significant implications for the development of more advanced diagnostic tools for liver disease and could ultimately help to improve patient outcomes and treatment options.

7.4 Suggestions for Future Research

From the ultimate goal of improving HCC screening and the results obtained in the second study, the two-stage model demonstrated the potential to significantly reduce the number of patients who require further investigation for hepatic tumors. In this study, the total number of tumors was effectively reduced from 1983 lesions

to 605 lesions (as shown in Table 8), focusing on those with a higher risk of malignancy.

Building upon these findings, there are several avenues for future research that could further enhance the field of deep learning for hepatic lesion detection and classification. Firstly, exploring more advanced deep learning architectures, such as CNNs or recurrent neural networks (RNNs), could improve the accuracy and robustness of the models. These architectures have shown great potential in various computer vision tasks and may offer improved performance in detecting and diagnosing hepatic lesions.

Furthermore, the development of more comprehensive datasets encompassing a wider range of hepatic lesions and imaging modalities would be valuable. Such datasets would provide a rich source of information to refine and optimize deep learning algorithms, allowing for more accurate and reliable detection and diagnosis of liver disease.

Additionally, integrating additional types of data, such as patient history, laboratory test results, and clinical symptoms, into the deep learning model could further improve diagnostic accuracy and enable more personalized treatment recommendations. By incorporating a holistic approach to data integration, the model could provide more comprehensive and tailored insights for individual patients.

Finally, exploring the use of explainable AI (XAI) techniques is another important area for future investigation. As AI-assisted detection and classification of liver lesions in ultrasound, it could support radiologist's efficiency in clinical situations. XAI methods aim to provide transparency and interpretability to deep learning models, allowing clinicians and radiologists to better understand how predictions are made. This could help build trust in the model's outputs and facilitate decision-making processes, further supporting and enhancing the efficiency of radiologists.

Overall, future research should focus on refining deep learning models, expanding datasets, integrating diverse data sources, and leveraging XAI techniques. These advancements would contribute to the continuous improvement of HCC screening programs, ultimately leading to enhanced patient care and outcomes.



REFERENCES

1. Fitzmaurice C, Allen C, Barber RM, Barregard L, Bhutta ZA, Brenner H, et al. Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015: a systematic analysis for the global burden of disease study. *JAMA oncology*. 2017;3(4):524-48.
2. Kumar V, Abbas AK, Fausto N, Aster JC. Robbins and Cotran pathologic basis of disease, professional edition e-book: elsevier health sciences; 2014.
3. Dietrich CF, Serra C, Jedrzejczyk M. Ultrasound of the liver. EFSUMB—European Course Book. 2010:56-66.
4. Schmauch B, Herent P, Jehanno P, Dehaene O, Saillard C, Aube C, et al. Diagnosis of focal liver lesions from ultrasound using deep learning. *Diagn Interv Imag*. 2019;100(4):227-33.
5. Tiyarattanachai T, Apiparakoon T, Marukatat S, Sukcharoen S, Geratikornsupuk N, Anukulkarikusol N, et al. Development and validation of artificial intelligence to detect and diagnose liver lesions from ultrasound images. *PloS one*. 2021;16(6):e0252882.
6. Caturelli E, Pompili M, Bartolucci F, Siena DA, Sperandeo M, Andriulli A, et al. Hemangioma-like lesions in chronic liver disease: diagnostic evaluation in patients. *Radiology*. 2001;220(2):337-42.
7. Chou R, Cuevas C, Fu R, Devine B, Wasson N, Ginsburg A, et al. Imaging techniques for the diagnosis of hepatocellular carcinoma: a systematic review and meta-analysis. *Annals of internal medicine*. 2015;162(10):697-711.
8. Deutsch GS, Yeh KA, Bates III WB, Tannehill WB. Embolization for management of hepatic hemangiomas. *The American Surgeon*. 2001;67(2):159-64.
9. Conter RL, Longmire Jr WP. Recurrent hepatic hemangiomas. Possible association with estrogen therapy. *Annals of surgery*. 1988;207(2):115.
10. Nichols JA, Chan HWH, Baker MA. Machine learning: applications of artificial intelligence to imaging and diagnosis. *Biophysical reviews*. 2019;11(1):111-8.
11. Andrew R. Jamieson* KD, Maryellen L. Giger. Breast Image Feature Learning with

Adaptive Deconvolutional Networks. Proceedings of SPIE - The International Society for Optical Engineering. 2012.

12. Cornell B. Hepatic lobules: BioNinja; 2016 [cited 2019 July 2]. Available from: <https://ib.bioninja.com.au/options/option-d-human-physiology/d3-functions-of-the-liver/liver-structure.html#previous-photo>.
13. Marieb EN, Hoehn K. Human anatomy & physiology: Pearson education; 2007.
14. Nguyen BN, Fléjou J-F, Terris B, Belghiti J, Degott C. Focal nodular hyperplasia of the liver: a comprehensive pathologic study of 305 lesions and recognition of new histologic forms. The American journal of surgical pathology. 1999;23(12):1441.
15. Dhingra S, Fiel MI. Update on the new classification of hepatic adenomas: clinical, molecular, and pathologic characteristics. Archives of Pathology and Laboratory Medicine. 2014;138(8):1090-7.
16. Hapani H, Kalola J, Trivedi A, Chawla A. Ultrasound evaluation of focal hepatic lesions. IOSR J Dent Med Sci. 2014;13(12):40-5.
17. Nguyen BN FJ, Terris B, Belghiti J, Degott C. Focal nodular hyperplasia of the liver: a comprehensive pathologic study of 305 lesions and recognition of new histologic forms. American Journal of Surgical Pathology. 23;12:1441-54.
18. Bray F FJ, Soerjomataram I, Siegel RL, Torre LA, Jemal A. . Global cancer statistics for the most common cancers: WCRF International; 2019 [cited 2019 July 6]. Available from: <https://www.wcrf.org/dietandcancer/cancer-trends/worldwide-cancer-data>.
19. Song DS, Bae SH. Changes of guidelines diagnosing hepatocellular carcinoma during the last ten-year period. Clin Mol Hepatol. 2012;18(3):258-67.
20. Laugier P, Haiat G. Introduction to the physics of ultrasound. Bone quantitative ultrasound: Springer; 2011. p. 29-45.
21. Ortiz SHC, Chiu T, Fox MD. Ultrasound image enhancement: A review. Biomedical Signal Processing and Control. 2012;7(5):419-28.
22. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. nature. 2017;542(7639):115-8.
23. Hekler A, Utikal JS, Enk AH, Hauschild A, Weichenthal M, Maron RC, et al.

Superior skin cancer classification by the combination of human and artificial intelligence. *European Journal of Cancer*. 2019;120:114-21.

24. He K, Zhang X, Ren S, Sun J, editors. Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*; 2016.

25. Yang B, Yan J, Lei Z, Li SZ, editors. Convolutional channel features. *Proceedings of the IEEE international conference on computer vision*; 2015.

26. Jing L, Zhao M, Li P, Xu X. A convolutional neural network based feature learning and fault diagnosis method for the condition monitoring of gearbox. *Measurement*. 2017;111:1-10.

27. Girshick R, Donahue J, Darrell T, Malik J, editors. Rich feature hierarchies for accurate object detection and semantic segmentation. *Proceedings of the IEEE conference on computer vision and pattern recognition*; 2014.

28. Girshick R, editor Fast r-cnn. *Proceedings of the IEEE international conference on computer vision*; 2015.

29. Ren S, He K, Girshick R, Sun J. Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*. 2015;28.

30. Cai Z, Vasconcelos N, editors. Cascade r-cnn: Delving into high quality object detection. *Proceedings of the IEEE conference on computer vision and pattern recognition*; 2018.

31. Lu X, Li Q, Li B, Yan J, editors. Mimicdet: Bridging the gap between one-stage and two-stage object detection. *European Conference on Computer Vision*; 2020: Springer.

32. Carranza-García M, Torres-Mateo J, Lara-Benítez P, García-Gutiérrez J. On the performance of one-stage and two-stage object detectors in autonomous vehicles using camera data. *Remote Sensing*. 2020;13(1):89.

33. Yang R, Yu Y. Artificial convolutional neural network in object detection and semantic segmentation for medical imaging analysis. *Frontiers in oncology*. 2021;11:638182.

34. Jiang P, Ergu D, Liu F, Cai Y, Ma B. A Review of Yolo algorithm developments.

Procedia Computer Science. 2022;199:1066-73.

35. Muhammed Enes A, DURAN Z, OZGUNLUK R. Comparison of YOLO Versions for Object Detection from Aerial Images. International Journal. 2022(2):87-93.

36. Redmon J, Farhadi A, editors. YOLO9000: better, faster, stronger. Proceedings of the IEEE conference on computer vision and pattern recognition; 2017.

37. Bochkovskiy A, Wang C-Y, Liao H-YM. Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:200410934. 2020.

38. Lei Zhang¹ HZ, Tengfei Yang¹. Deep Neural Networks for fatty liver ultrasound images classification. IEEE. 2019:4641-6.

39. Hassan TM, Elmogy M, Sallam E-S. Diagnosis of focal liver diseases based on deep learning technique for ultrasound images. Arabian Journal for Science Engineering. 2017;42(8):3127-40.

40. Lingyun Wu J-ZC, Shengli Li, Baiying Lei, Tianfu Wang, and Dong Ni. FUIQA: Fetal Ultrasound Image Quality Assessment With Deep Convolutional Networks. IEEE TRANSACTIONS ON CYBERNETICS. 2017;47(No.5):1336-49.

41. Azizi S, Bayat S, Yan P, Tahmasebi A, Nir G, Kwak JT, et al. Detection and grading of prostate cancer using temporal enhanced ultrasound: combining deep neural networks and tissue mimicking simulations. International journal of computer assisted radiology surgery. 2017;12(8):1293-305.

42. Shen L, Margolies LR, Rothstein JH, Fluder E, McBride R, Sieh W. Deep learning to improve breast cancer detection on screening mammography. Scientific reports. 2019;9(1):1-12.

43. Monsi J, Saji J, Vinod K, Joy L, Mathew JJ. XRAY AI: Lung Disease Prediction Using Machine Learning. International Journal of Information. 2019;8(2).

44. Wang X, Cheng P, Liu X, Uzochukwu B, editors. Focal loss dense detector for vehicle surveillance. 2018 international conference on intelligent systems and computer vision (ISCV); 2018: IEEE.

45. Lin T-Y, Goyal P, Girshick R, He K, Dollár P, editors. Focal loss for dense object detection. Proceedings of the IEEE international conference on computer vision; 2017.

46. Jocher G. ultralytics [Internet]2023. [cited 2023]. Available from: https://docs.ultralytics.com/yolov5/tutorials/tips_for_best_training_results/.

47. Nishida N, Yamakawa M, Shiina T, Mekada Y, Nishida M, Sakamoto N, et al. Artificial intelligence (AI) models for the ultrasonographic diagnosis of liver tumors and comparison of diagnostic accuracies between AI and human experts. *Journal of Gastroenterology*. 2022;57(4):309-21.
48. Cao Z, Duan L, Yang G, Yue T, Chen Q. An experimental study on breast lesion detection and classification from ultrasound images using deep learning architectures. *BMC medical imaging*. 2019;19(1):1-9.
49. Tanaka H, Chiu S-W, Watanabe T, Kaoku S, Yamaguchi T. Computer-aided diagnosis system for breast ultrasound images using deep learning. *Physics in Medicine & Biology*. 2019;64(23):235013.





จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

VITA

NAME	Mrs. Sutthirak Tangruangkiat
DATE OF BIRTH	11 January 1974
PLACE OF BIRTH	Ang Thong
INSTITUTIONS ATTENDED	Master degree in Applied Radiation and Isotopes, Kasetsart University Bachelor degree in Radiological technology, Mahidol University
HOME ADDRESS	30 Soi Lat Prao Wang Hin 40, Lat Prao, Lat Prao, Bangkok Thailand. 10230
PUBLICATION	<p>1) Chaiwongkot N, Jomsak T, Thabsangthong T, Tangruangkiat S*, Siripongsakun S and Hiranrat P. Differentiation of vascular patterns in hepatic hemangioma using superb microvascular imaging (SMI) technology of ultrasound at Chulabhorn Hospital. The Thai Journal of Radiological Technology 46 (2021): 43-51.</p> <p>2) Tangruangkiat S, Krisanachinda A and Auetawekiat S. The effective methods for liver cyst labelling in ultrasound images by using R-CNN. Electronics-Proceedings of 20th AOCMP, 18th SEACOMP, 120th JSMP, 12th TMPS; 2020 Dec 3-5; Phuket, Thailand.</p> <p>3) Ritlumlert N, Tangruangkiat S, Kawvised S, Pairodsantikul P, Vidhyarkorn S. Assessment of Average Glandular Dose Received in Full-field Digital Mammography and Digital Breast Tomosynthesis. Journal of Health Science and Medical Research. 2020 Mar 30; 38(2): 115-23.</p>

4) Promduang A, Pongnapang N, Ritlumlert N, Tangruangkiat S, Phonlakrai M. A Study of Entrance Surface Air Kerma for Patients Undergoing Chest and Abdomen from Digital Radiography at Chulabhorn Hospital. Journal of Health Science and Medical Research. 2019 Feb 12;37(1): 51-60.

5) Siripongsakun S, Vidhyarkorn S, Charuswattanakul S, Mekraksakit P, Sungkasubun P, Yodkhunnathum N, Tangruangkiat S, Ritlumlert N, Sricharunrat T, Jaroenpatarapesaj S, Soonklang K. Ultrasound surveillance for cholangiocarcinoma in an endemic area: A prove of survival benefits. Journal of gastroenterology and hepatology. 2018 Jul;33(7): 1383-8.

6) Tangruangkiat S., Siripongsakul S., Vidhayakorn S., Thitisithikorn W., Ritlumlert N., Polkrai M. Ultrasonography Detection of Liver Lesion: Comparison between Radiologists vs Sonographer., 24th Annual Meeting, Thai society of Radiological Technology, 27-29 April 2016, Pattaya, Thailand.

7) Tangruangkiat S, Roongtanakiat N, Meesat R. Potential use of vetiver in industrial wastewater treatment. Warasan Din Lae Pui. 2006.

AWARD RECEIVED

-