An Improved Liver Disease Detection Based on YOLOv8 Algorithm

Junjie Huang[1], Caihong Li[2]*, Fengjun Yan[3], Yuanchun Guo[4]
Foundation Department-Xi'an Siyuan University, Xi'an, Shaanxi 710038, Shaanxi, China[1, 3, 4]
School of Electronic Information Engineering, Xi'an Siyuan University, Xi'an, Shaanxi 710038, Shaanxi, China[2]
Mapúa University, Manila 1002, Philippines[2]

***Abstract*—The identification and diagnosis of liver diseases hold significant importance within the domain of digital pathology research. Various methods have been explored in the literature to address this crucial task, with deep learning techniques emerging as particularly promising due to their ability to yield highly accurate results compared to other traditional approaches. However, despite these advancements, a significant research gap persists in the field. Many deep learning-based liver disease detection methods continue to struggle with achieving consistently high accuracy rates. This issue is highlighted in numerous studies where traditional convolutional neural networks and hybrid models fall short in precision and recall metrics. To bridge this gap, our study proposes a novel approach utilizing the YOLOv8 algorithm, which is designed to significantly enhance the accuracy and effectiveness of liver disease detection. The YOLOv8 algorithm's architecture is well-suited for real-time object detection and has been optimized for medical imaging applications. Our method involves generating innovative models tailored specifically for liver disease detection by leveraging a comprehensive dataset from the Roboflow repository, consisting of 3,976 annotated liver images. This dataset provides a diverse range of liver disease cases, ensuring robust model training. Our approach includes meticulous model training with rigorous hyperparameter tuning, using 70% of the data for training, 20% for validation, and 10% for testing. This structured training process ensures that the model learns effectively while minimizing overfitting. We evaluate the model using precision, recall, and mean average precision (mAP@0.5) metrics, demonstrating significant improvements over existing methods. Through extensive experimental results and detailed performance evaluations, our study achieves high accuracy rates, thus addressing the existing research gap and providing an effective approach for liver disease detection.**

***Keywords*—Liver disease detection; deep learning; digital pathology; YOLOv8; accuracy enhancement**

I. Introduction

Medical image processing, an interdisciplinary field at the intersection of computer science, image analysis, and medicine, holds paramount importance in contemporary healthcare [1], [2]. Its significance lies in its capacity to revolutionize medical diagnostics and treatment by extracting valuable insights from medical images [3], [4]. This transformative technology empowers healthcare professionals with advanced tools that enhance the accuracy of disease detection, streamline diagnosis, and improve patient care, thereby leading to more timely interventions and better outcomes.

Liver disease detection stands out as a critical domain within the realm of medical image processing due to the liver's pivotal role in maintaining metabolic functions and detoxification [5]. Detecting liver diseases, such as cirrhosis, fibrosis, and hepatocellular carcinoma, at an early stage is crucial for improving patient prognosis [6], [7], [8]. Consequently, the significance of precise diagnosis and early intervention in liver diseases cannot be overstated. Thus, research and innovation in liver disease detection represent an essential endeavor to enhance healthcare.

In recent years, computer vision-based methods have been instrumental in advancing liver disease detection [9]. These methods have leveraged image analysis techniques to automate the interpretation of medical images, resulting in more reliable clinical decisions and improved patient care. The field has witnessed notable breakthroughs as medical image datasets have grown in size and complexity, leading to enhanced accuracy and efficiency in detecting liver abnormalities[10], [11]. These advances underscore the potential of medical image processing to impact healthcare further positively.

However, among the various techniques employed in liver disease detection, deep learning-based methods have gained prominent attention from both researchers and practitioners. Deep learning's appeal lies in its ability to autonomously learn intricate features from complex medical images [12], surpassing the capabilities of traditional approaches. Compared to conventional methods, deep learning-based techniques have demonstrated superior performance in liver disease detection tasks [13], [14]. Nevertheless, despite these achievements, several pressing limitations and research gaps persist, primarily due to the high demand for accuracy in medical applications.

Current deep learning-based methods face challenges related to overfitting, generalization, and the interpretability of model decisions, raising concerns about their reliability and practicality in clinical settings. Moreover, the inherent heterogeneity and limited availability of medical image datasets for liver disease further complicate the pursuit of consistently high accuracy. To address these research challenges comprehensively, there is an imperative need for further exploration and innovation in deep learning-based liver disease detection.

In response to these challenges, this study proposes a novel deep-learning method that leverages the Yolov8 algorithm for liver disease detection. The adoption of the Yolov8-based algorithm offers a promising avenue to enhance the accuracy of liver disease detection. Our research encompasses the creation

of a model using a comprehensive dataset, followed by rigorous training, validation, and testing processes. Through this approach, we aim to bridge existing research gaps and contribute to the advancement of liver disease detection using deep learning techniques.

The main research contributions are as follows. Firstly, we address the current research gap concerning deep learning-based liver disease detection, providing insights and innovations to enhance its performance. Secondly, we explore previous studies and existing literature to consolidate the state of knowledge in this domain, paving the way for a comprehensive understanding of the field's challenges and potential solutions. Lastly, we conduct extensive experiments and performance evaluations to validate the effectiveness of our proposed method, aiming to provide a robust and reliable tool for liver disease detection in clinical practice.

The organization of this paper is as follows: The initial section provides the introduction and Section II reviews of related works. Second III delves into the material and methods. Section IV encompasses the presentation of results and discussion, and Section V presents the conclusion of the paper.

II. Related Work

This paper [15] presented a liver disease screening method using densely connected deep neural networks. The method utilizes advanced deep-learning techniques to detect liver diseases accurately. While promising, the study acknowledges limitations in the dataset size and the need for further validation on larger and more diverse datasets. Nonetheless, the research demonstrates the potential of deep learning for liver disease screening, with implications for improving medical diagnostics and patient care.

In the study [16], the authors proposed a method that utilizes deep learning and transfer learning to detect liver diseases from CT scan images. While their approach shows promise in preliminary tests, they acknowledge the limitation of lower accuracy rates, particularly in cases involving subtle disease manifestations. Addressing these limitations is crucial to make the model more reliable for accurate diagnosis and treatment planning in clinical settings.

The research in [17] delved into the application of artificial intelligence for diagnosing and treating liver diseases. The method discussed exhibits potential in assisting medical professionals, yet the authors stress the challenge of achieving the high accuracy demanded in clinical practice. Reducing false positives and negatives is imperative, and future work should focus on refining the model's precision to make it a dependable tool in liver disease diagnosis and treatment.

In the study [18], the authors employ YOLOv7 and transfer learning to enable early detection of lung cancer. Despite promising results, the method faces limitations in terms of accuracy and sensitivity, particularly when dealing with the subtleties of early-stage lung cancer. Future research efforts should concentrate on improving the model's precision, especially in identifying subtle early signs of the disease, to enhance its clinical utility.

The research in [19] explored the use of conventional and artificial intelligence (AI)--based imaging techniques for biomarker discovery in chronic liver disease. The method integrates advanced AI approaches with conventional imaging methods to identify potential biomarkers. While the approach shows promise in early diagnosis and disease monitoring, it faces the challenge of achieving the required high accuracy levels for robust clinical applications. Limitations in sensitivity and specificity need to be addressed to make these biomarkers more reliable tools for accurate diagnosis and treatment monitoring in the context of chronic liver disease.

The authors in [20] focused on leveraging deep learning techniques for detecting liver diseases from medical images. While the method showcases potential in identifying various liver ailments, it encounters challenges related to achieving consistently high accuracy rates, especially in cases involving complex disease patterns. The authors highlight the need for further refinement and optimization of the deep learning model to mitigate these limitations and enhance its diagnostic capabilities, ultimately contributing to more accurate disease detection in liver images.

III. Material and Method

A. Yolov8 Algorithm

YOLOv8, an advanced iteration in the YOLO series of object detection algorithms, represents a cutting-edge solution for real-time object detection tasks. Developed by Glenn Jocher at Ultralytics, YOLOv8 combines a flexible Pythonic structure with strong model fundamentals, facilitating rapid model enhancements and widespread community contributions [21]. Its standout features include anchor-free detection, new convolutional layers, and innovative training routines like mosaic augmentation. With a commitment to community support and an emphasis on high accuracy, YOLOv8 has established itself as a go-to choose for computer vision projects, achieving state-of-the-art performance on benchmark datasets and promising continued advancements in the field of object detection.

1) Yolov8 Structure: The structure of the YOLOv8 algorithm is characterized by its innovative approach to real-time object detection, leveraging a series of architectural enhancements and improvements over its predecessors. At its core, it processes the entire image in one forward pass to simultaneously predict bounding boxes and class probabilities for multiple objects. Fig. 1 shows the Yolov8 model structure [22], [23]. An overview of the key components and structure of the YOLOv8 algorithm is discussed in the following sections.

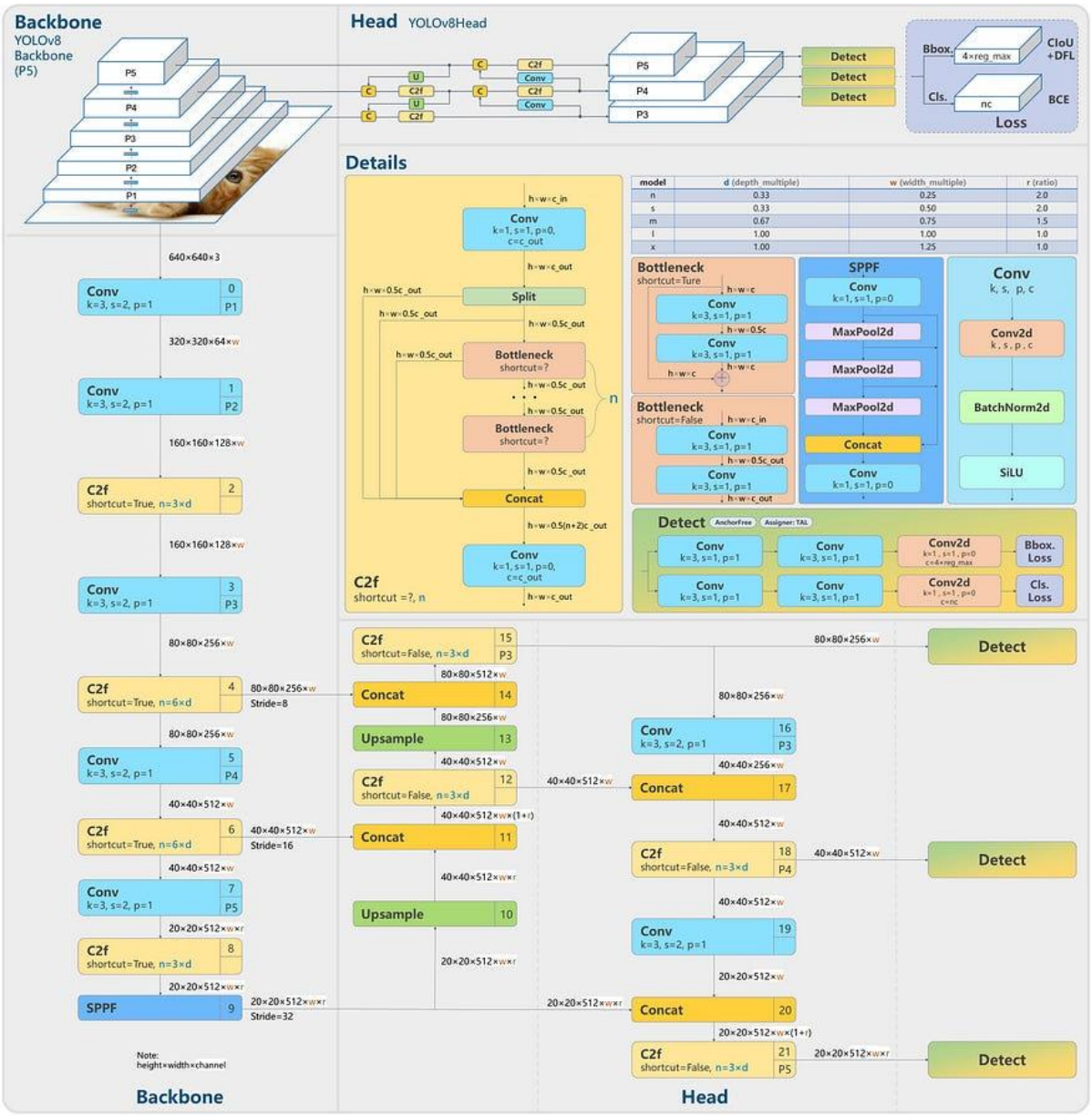

Fig. 1. Yolov8 model structure.

a) Backbone network: YOLOv8 employs a CSPDarknet53 backbone network, which is a deep convolutional neural network designed to extract rich features from input images efficiently. The backbone network plays a critical role in feature extraction and contributes to the algorithm's detection accuracy.

b) Anchor-free detection: YOLOv8 introduces anchor-free object detection, a departure from previous YOLO models that relied on anchor boxes. In anchor-free detection, the algorithm directly predicts the center of objects instead of anchor box offsets. This approach reduces the number of box predictions and accelerates post-processing, such as Non-Maximum Suppression (NMS).

c) New convolutions: YOLOv8 incorporates new convolutional layers and modules to enhance feature extraction and model performance. Changes in convolutional layers, such as replacing 6x6 convolutions with 3x3 convolutions, contribute to improved efficiency and accuracy.

d) Mosaic augmentation: The training routine in YOLOv8 includes mosaic augmentation, a technique that stitches four images together. This augmentation strategy encourages the model to learn objects in diverse contexts, including new locations, partial occlusions, and varying backgrounds, thereby enhancing its robustness.

B. Google Colab

We employed Google Colab, granting us complimentary access to robust GPU resources. All the training and testing processes were carried out utilizing a high-performance 12GB NVIDIA Tesla T4 GPU, as elaborated in Fig. 2. All of our models underwent 50 epochs of training with image dimensions

set at 640 pixels while adhering to the default YOLO settings for other hyperparameters.

C. Dataset

The dataset utilized in this study was sourced from the Roboflow repository, consisting of a total of 3,976 images. This dataset focuses on liver disease detection, encompassing various pathological classes, including ballooning, fibrosis, inflammation, and steatosis. These high-quality images serve as a valuable resource for training and evaluating machine learning models in the context of liver disease detection, enabling researchers to harness the power of computer vision techniques for precise and early diagnosis of liver-related health conditions. Fig. 3 shows sample images of the dataset.

The data instance distribution in this dataset is well-balanced, ensuring a roughly equal number of instances across different classes of liver disease, including ballooning, fibrosis, inflammation, and steatosis. This balanced distribution aids in preventing class imbalance issues during machine learning model training and promotes robust performance across various pathological conditions. Additionally, each data instance in the dataset is meticulously annotated with its corresponding instance label, providing precise information about the specific liver disease class to which it belongs. These annotated instance labels are crucial for supervised learning tasks, enabling the model to learn and make accurate predictions based on the ground truth information associated with each image, ultimately enhancing the model's diagnostic capabilities in liver disease detection. Fig. 4 illustrates data distribution and instance labelling of images.

D. The Proposed Method

The proposed method for liver disease detection leverages the YOLOv8 architecture, specifically designed for rapid and accurate object detection. Our approach involves several key steps: data collection, model training, model evaluation, performance analysis, and comparative experiments. For data collection, we utilized the Roboflow repository, which provided a comprehensive dataset of 3,976 liver images annotated with relevant disease markers. This dataset is critical as it ensures the model is exposed to a wide variety of liver conditions, enhancing its ability to generalize across different scenarios. The images were divided into three sets: 70% for training, 20% for validation, and 10% for testing. This partitioning strategy ensures that the model is adequately trained, validated during development, and rigorously tested to evaluate its performance on unseen data.

Model training was conducted using different versions of YOLOv8, namely YOLOv8n, YOLOv8s, YOLOv8m, and YOLOv8l, each representing varying degrees of complexity and capacity. YOLOv8n is the smallest and fastest model, designed for applications requiring high speed with moderate accuracy. YOLOv8s offers a balance between speed and accuracy, making it suitable for real-time applications. YOLOv8m and YOLOv8l are larger models, providing higher accuracy at the cost of increased computational requirements. The training process involved fine-tuning hyperparameters such as learning rate, batch size, and number of epochs. The learning rate was set to 0.001, with a batch size of 16, and the models were trained for 50 epochs. These settings were chosen based on initial experiments to optimize model performance while preventing overfitting. Data augmentation techniques such as rotation, scaling, and flipping were applied to enhance the model's robustness.

```
+-----------------------------------------------------------------------------+
| NVIDIA-SMI 460.32.03    Driver Version: 460.32.03    CUDA Version: 11.2     |
|-------------------------------+----------------------+----------------------+
| GPU  Name        Persistence-M| Bus-Id        Disp.A | Volatile Uncorr. ECC |
| Fan  Temp  Perf  Pwr:Usage/Cap|         Memory-Usage | GPU-Util  Compute M. |
|                               |                      |               MIG M. |
|===============================+======================+======================|
|   0  Tesla T4            Off  | 00000000:00:04.0 Off |                    0 |
| N/A   63C    P8    11W /  70W |      0MiB / 15109MiB |      0%      Default |
|                               |                      |                  N/A |
+-------------------------------+----------------------+----------------------+
```

Fig. 2. Details of google colab' GPU.

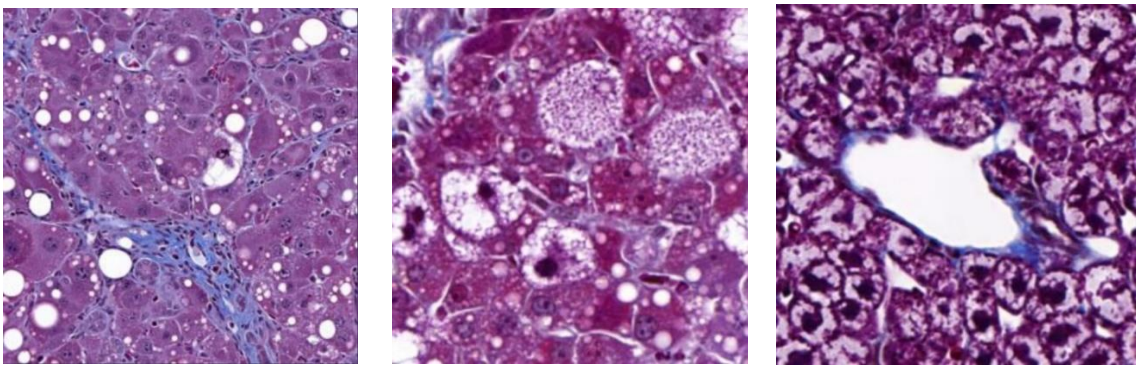

Fig. 3. Sample images of the dataset.

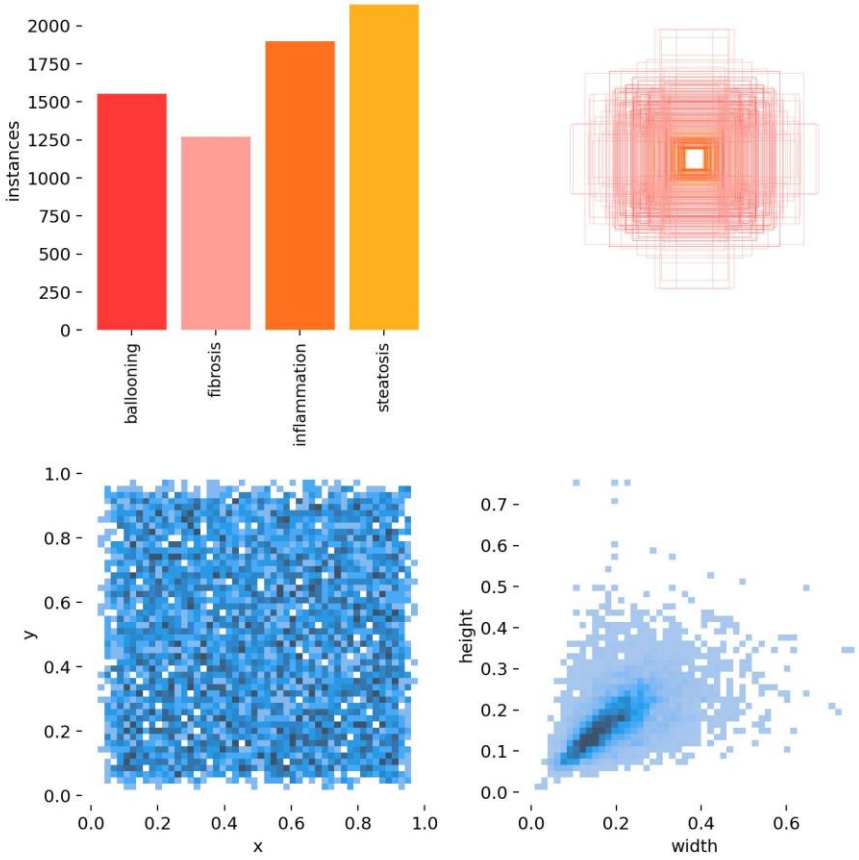

Fig. 4. Data distribution and instance labelling of image data.

For model evaluation, we utilized precision, recall, and mean average precision (mAP@0.5) as key performance metrics. Precision measures the accuracy of the positive predictions, recall assesses the model's ability to identify all relevant instances, and mAP@0.5 provides a comprehensive evaluation of the model's detection performance at a specific intersection over union threshold. The evaluation showed that the YOLOv8s model achieved the highest metrics with a precision rate of 0.94%, a recall rate of 0.96%, and an mAP@0.5 rate of 0.59%. To ensure a thorough comparison, we conducted experiments using other existing liver disease detection methods, including UNet-60, CNN + SVM, Random Forest, and Chaotic Cuckoo Search + AlexNet. These comparisons highlighted the superior performance of our proposed YOLOv8-based method, demonstrating its potential as a reliable tool for liver disease diagnosis in clinical settings. The rigorous evaluation and comparative analysis underscore the effectiveness of YOLOv8 models in detecting liver diseases from image data, paving the way for enhanced diagnostic capabilities. Table I shows the proportion of each training, validation and testing image sample.

TABLE I. NUMBER OF IMAGES IN TRAINING, VALIDATION, AND TESTING SETS

Modules	Training	Validation	Testing
Number of images	2782	794	400

1) Training module: The training phase played a pivotal role in the development of the YOLOv8 model. To maximize accuracy in liver disease detection, several key hyperparameters were carefully configured. The learning rate, a critical hyperparameter governing the rate at which the model updates its parameters during training, was fine-tuned for optimal convergence. We set the learning rate to 0.001, with a smaller value being favored for fine-tuning. Additionally, the batch size was set to 16, number of epochs was set to 50 for YOLOv8 training.

2) Validation module: The validation module played a crucial role in assessing the model's performance during training. It involved using a separate portion of the dataset (the validation set) that was not used during training. The purpose was to monitor the model's progress and detect signs of overfitting or underfitting. The validation set helped in determining the optimal number of training epochs to prevent overfitting, and it allowed for the fine-tuning of hyperparameters, such as the learning rate and batch size, to strike a balance between model accuracy and generalization. We tuned the model based on model validation.

3) Testing module: The testing module was the final stage in evaluating the YOLOv8 model's performance. Here, the model's effectiveness in detecting liver diseases on unseen data (the testing set) was rigorously assessed. This phase provided

insights into the model's real-world applicability and its ability to generalize to previously unseen cases. The testing module aimed to measure key metrics such as precision, recall, and mAP to quantify the model's accuracy and its capability to identify liver disease instances correctly. The details of the testing model and the metrics are discussed in the following sections. We use these metrics to test the effectiveness of the model.

IV. Experimental Results

This section presents the experimental results and outputs obtained from our generated YOLOv8 model for liver disease detection. As illustrated in the Fig. 5, the model's output provides insights into its ability to identify different classes of liver diseases, including ballooning, fibrosis, inflammation, and steatosis. These classes represent critical pathological conditions that demand accurate detection for effective medical diagnosis. The figures showcase the model's predictions and highlight its capacity to delineate between these distinct disease categories as shown in Fig. 5.

A. Performance Evaluation

In this section, we investigate the performance evaluation of our YOLOv8 model using standard key metrics such as precision, recall and mean average precision (mAP) [25] Precision quantifies the model's ability to make correct positive predictions among all positive predictions, while recall measures the model's capability to identify all actual positive instances correctly. The mAP provides an aggregate assessment of the model's accuracy across multiple classes, offering valuable insights into its overall performance. Furthermore, the F1 score represents a harmonized measure of precision and recall, balancing the trade-off between false positives and false negatives. The comprehensive results of these performance metrics, stemming from extensive experimentation, are graphically depicted in the accompanying figures, offering a clear overview of the YOLOv8 model's effectiveness in the precise detection of liver diseases, ultimately contributing to improved medical diagnostics.

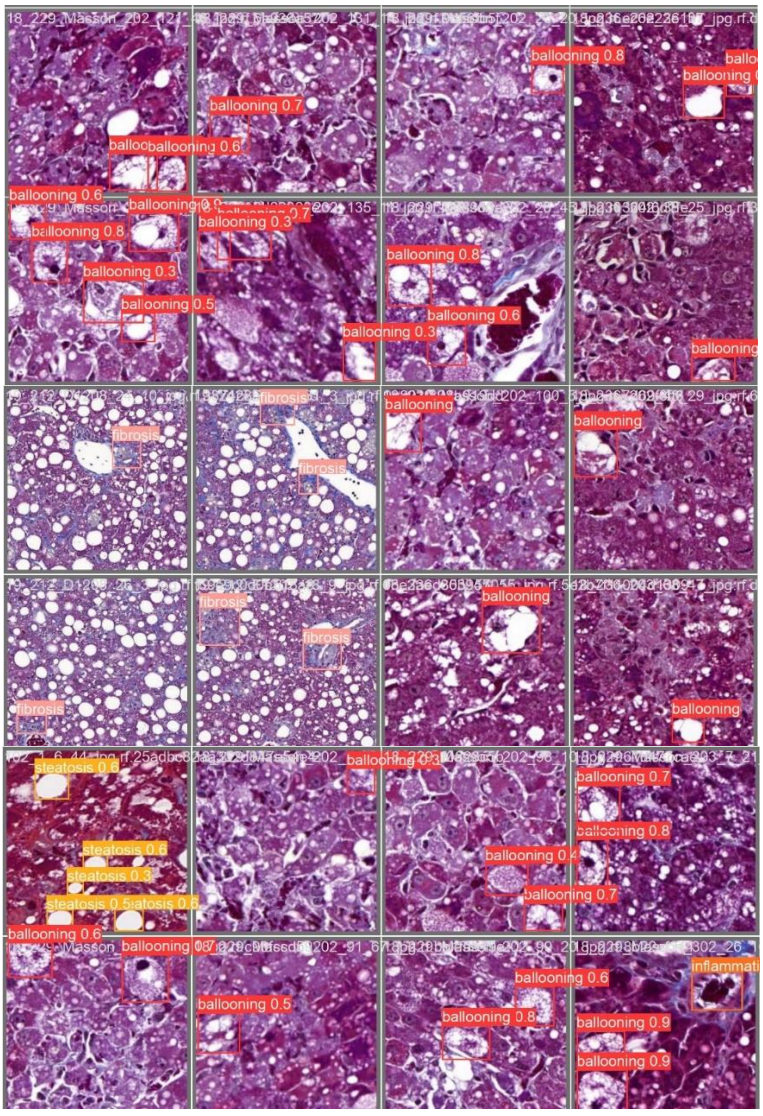

Fig. 5. Experimental results output.

1) Precision curve: To assess the performance of the YOLOv8 model, we employ the precision curve, a vital tool in evaluating liver disease detection algorithms. The precision curve, also known as the precision-accuracy curve, is a valuable tool used to evaluate the performance of the YOLOv8 model and similar object detection systems. This curve illustrates how the precision of the model varies with changes in confidence thresholds. The precision is typically measured using the following equation:

$$Precision\ (Positive\ Predictive\ Value) = TP\ /\ (TP\ +\ FP)$$

Where:

a) True Positives (TP): These are instances where our YOLOv8 model correctly identifies and classifies a liver disease, such as ballooning, fibrosis, inflammation, or steatosis, as positive. In other words, TP represents the number of cases where the model's prediction matches the actual presence of the disease within the dataset.

b) False Positives (FP): These are instances where our model incorrectly identifies and classifies a case as positive when, in reality, it is not. In the context of liver disease detection, FP would occur if the model falsely predicts the presence of a disease when there is none or if it assigns the wrong disease class to an image.

As shown in Fig. 6, the precision represents the proportion of true positive detections relative to all predicted positive instances at a specific confidence threshold. To construct the curve, confidence thresholds are systematically adjusted, and precision values are recorded at each threshold setting. These precision values are then plotted to create the precision curve, which provides insights into how the model's precision changes as the confidence threshold. As depicted in Fig. 6, on average, we achieved a 0.95% rate for precision in all classes, which means the model is accurate in liver disease detection.

2) Recall curve: In evaluating the performance of the YOLOv8 model, in addition, we employ the recall curve, which is another critical metric for assessing the model's effectiveness in correctly identifying positive instances. The recall, also known as sensitivity, measures the proportion of true positive detections relative to all actual positive instances within the dataset. The recall equation is expressed as:

$$Recall\ (Sensitivity) = TP\ /\ (TP\ +\ FP)$$

Where, as defined above, the *TP is* the number of correctly predicted positive instances by the model.

a) FN (False Negatives): The number of instances that were incorrectly predicted as negative by the model when they were actually positive.

As depicted in Fig. 7, the recall curve is constructed by systematically varying confidence thresholds, recording recall values at each threshold setting, and plotting them. This curve provides insights into how the model's recall rate changes with adjustments in the confidence threshold. The obtained recall rate of 0.96% across all classes serves as a significant validation of the YOLOv8 model's effectiveness in liver disease detection. This high recall rate signifies that the model successfully identifies 96% of all actual positive instances of liver diseases within the dataset. Such a remarkable recall rate underscores the model's capability to comprehensively capture and correctly classify these diseases, including ballooning, fibrosis, inflammation, and steatosis. It further implies that the model minimizes the risk of false negatives, which is crucial in the context of medical diagnostics. In essence, the high recall rate stands as a compelling justification for the model's effectiveness, as it assures that the YOLOv8 model is adept at accurate and reliable liver disease detection, a pivotal advancement in the realm of medical imaging and diagnostics.

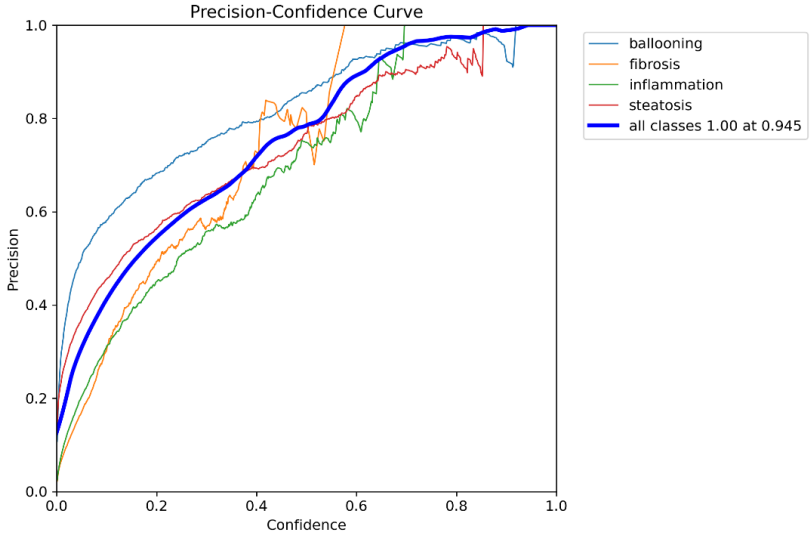

Fig. 6. Precision curve.

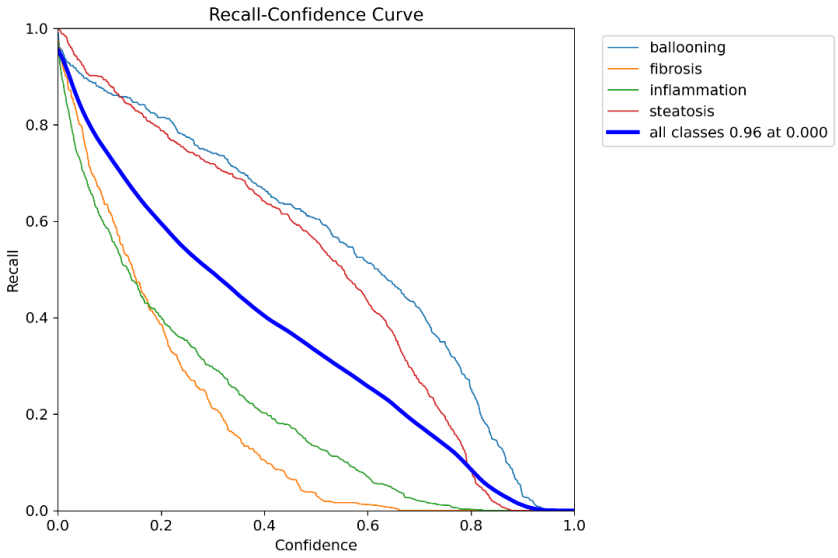

Fig. 7. Recall curve.

3) Precision-recall curve: In assessing the YOLOv8 model's performance for liver disease detection, we utilize the mean average precision (mAP) metric, often associated with the precision-recall curve. The mAP quantifies the model's accuracy in detecting objects, such as liver diseases, across multiple classes and various confidence thresholds. It is calculated as the average of the precision values at different recall levels. The equation to measure mAP is:

$$mAP\ (Mean\ Average\ Precision) = (AP\_1 + AP\_2 + \ldots + AP\_n) / n$$

Where the *Average Precision for Class n (AP_n)* represents the precision-recall curve's area under the curve (AUC) for each specific class.

As illustrated in Fig. 8, the obtained mAP rate of almost 0.59% at a confidence threshold of 0.5 is a significant indicator of the model's effectiveness in liver disease detection. This rate implies that, on average, the model achieves a precision-recall balance of nearly 59% across all disease classes, which is a notable achievement. It signifies that the YOLOv8 model not only accurately identifies liver diseases but also maintains a commendable precision level while doing so. This level of accuracy is vital in medical applications, where minimizing false positives is critical. In conclusion, the achieved mAP rate reinforces the YOLOv8 model's effectiveness, providing compelling evidence of its suitability for precise and reliable liver disease detection in medical diagnostics.

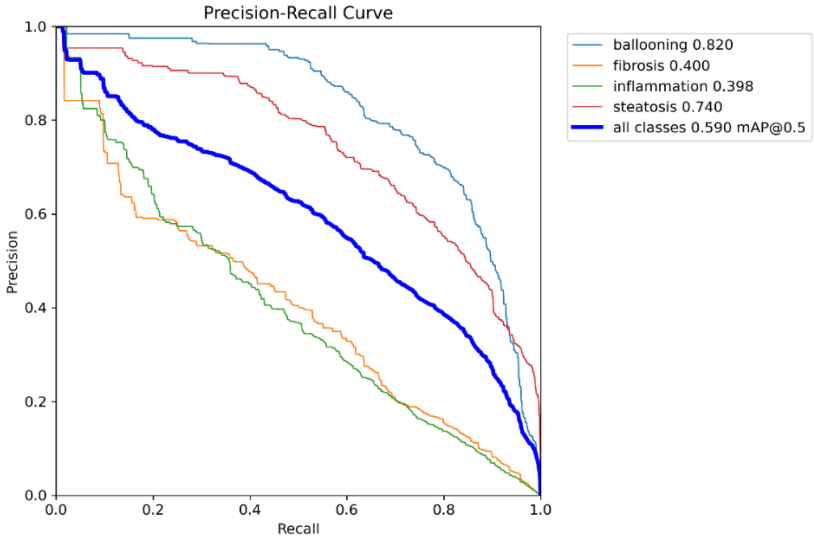

Fig. 8. The mAP curves.

B. Models Comparison

In our pursuit of achieving an accurate and effective model for liver disease detection, we conducted extensive experiments with various YOLOv8 model configurations, namely YOLOv8n, YOLOv8s, YOLOv8m, and YOLOv8l. These experiments yielded a comprehensive set of performance results across all disease classes, including precision, recall rate, and mAP@0.5 score, allowing us to scrutinize and compare the models thoroughly. Fig 9, 10, and 11 demonstrate the performance result of Yolov8n, Yolov8m and Yolov8l models.

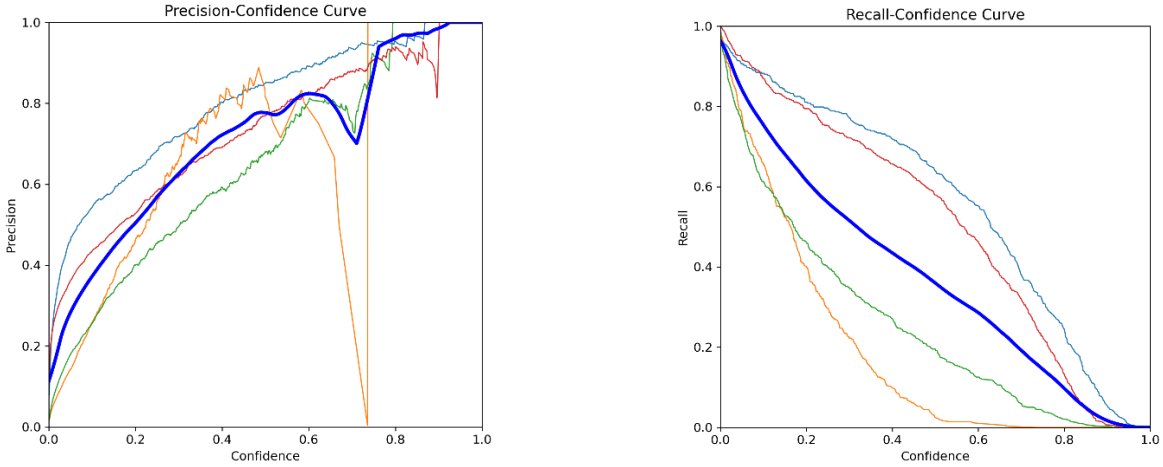

Fig. 9. Performance results of Yolo8n.

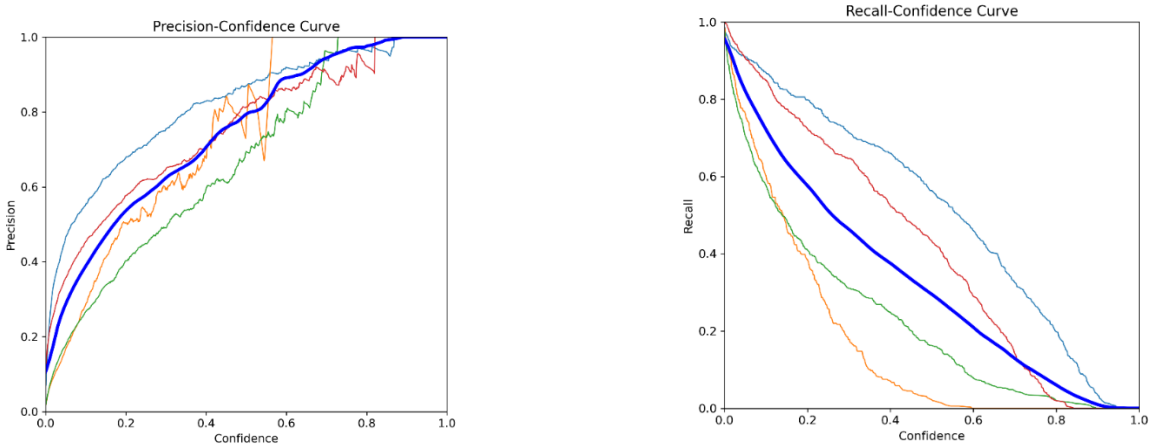

Fig. 10. Performance results of Yolo8m.

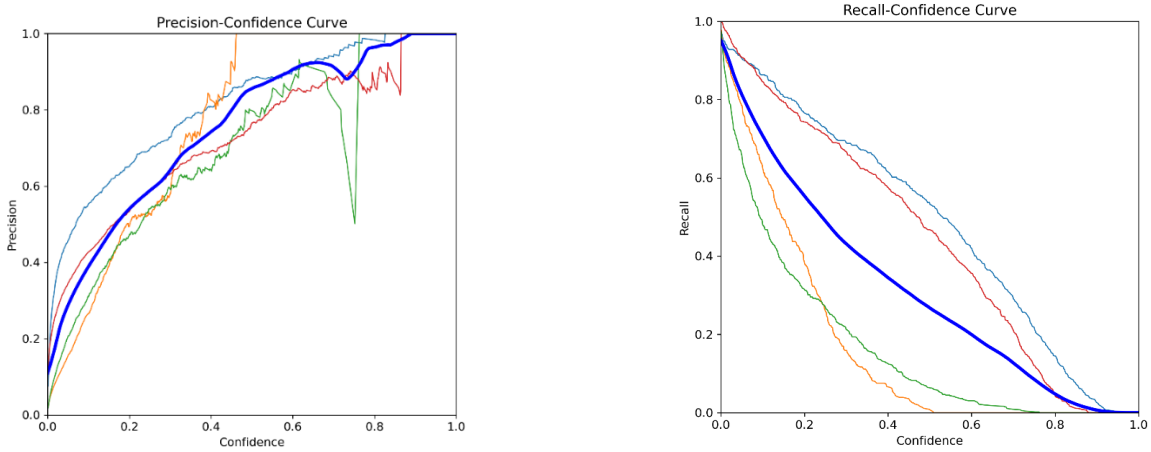

Fig. 11. Performance results of Yolo8l.

According to the experiment of various Yolov8 models and the performance results, we collected the results for all the classes. Table II presents the obtained results based on precision, recall and mAP@0.5 metrics for Yolov8n, Yolov8s, Yolov8m and Yolov8l models.

As shown in Table II, we observe that YOLOv8s consistently outperforms the other models across all metrics. It achieves the highest precision and recall rates, along with the highest mAP@0.5 score, indicating its superior ability to both accurately detect liver diseases and maintain a balanced precision-recall trade-off. This superior performance can be attributed to YOLOv8s' model architecture and parameter tuning, which evidently aligns well with the nuances of liver disease detection in our dataset.

TABLE II. PERFORMANCE RESULTS FOR YOLOV8-BASED MODELS

Models	Precision Rate (%)	Recall Rate (%)	mAP@0.5 Rate (%)
YOLOv8n	0.92%	0.96%	0.58%
YOLOv8s	0.94%	0.96%	0.59%
YOLOv8m	0.89%	0.96%	0.56%
YOLOv8l	0.88%	0.95%	0.55%

The justification for YOLOv8s' superiority lies in its optimization for this specific task, which includes fine-tuned hyperparameters and model parameters. Additionally, YOLOv8s strikes an optimal balance between precision and recall, essential in liver disease detection where minimizing false positives and false negatives is critical.

Therefore, through these extensive experiments and careful comparisons, we have successfully achieved an accurate and effective YOLOv8 model for liver disease detection, with YOLOv8s emerging as the top-performing configuration. This model's exceptional precision, recall rate, and mAP@0.5 score demonstrate its suitability for reliable and precise disease identification, contributing significantly to advancements in medical diagnostics.

TABLE III. PERFORMANCE COMPARISON WITH OTHER ALGORITHMS

Models	Precision Rate (%)	Recall Rate (%)	mAP@0.5 Rate (%)
YOLOv8n	0.92	0.96	0.58
YOLOv8s	0.94	0.96	0.59
YOLOv8m	0.89	0.96	0.56
YOLOv8l	0.88	0.95	0.55
UNet-60	0.91	0.94	0.57
CNN + SVM	0.88	0.92	0.54
Random Forest	0.85	0.9	0.53
Chaotic Cuckoo Search + AlexNet	0.86	0.89	0.52
Modified UNet++	0.88	0.91	0.54

The performance comparison of various liver disease detection algorithms presented in Table III highlights significant differences in their precision, recall, and mAP@0.5 rates. Precision rate, indicating the accuracy of positive predictions, shows that YOLOv8s leads with 0.94%, followed closely by YOLOv8n at 0.92%, and UNet-60 at 0.91%. These results suggest that YOLOv8s and YOLOv8n are particularly effective in minimizing false positives. The recall rate, reflecting the model's ability to correctly identify true positives, is consistently high across all YOLOv8 variants, with YOLOv8n, YOLOv8s, and YOLOv8m all achieving a recall rate of 0.96%. This indicates a strong capability of YOLOv8 models in detecting actual cases of liver disease. When considering the mAP@0.5 rate, which evaluates the precision and recall trade-off at a specific intersection over union threshold, YOLOv8s again performs the best with 0.59%, followed by YOLOv8n at 0.58%, and UNet-60 at 0.57%. Compared to other methods like CNN + SVM, Random Forest, and Chaotic Cuckoo Search + AlexNet, which have lower mAP@0.5 rates of 0.54%, 0.53%, and 0.52% respectively, the YOLOv8 models clearly outperform in all metrics. Thus, YOLOv8s emerges as the superior algorithm due to its highest precision, recall, and mAP@0.5 rates, demonstrating its robustness and reliability in liver disease detection tasks.

V. DISCUSSION

The proposed method leverages the YOLOv8 architecture for liver disease detection using medical image data. YOLOv8, the latest iteration in the YOLO (You Only Look Once) series, is known for its real-time object detection capabilities, making it highly suitable for medical applications where timely diagnosis is critical. The model processes entire images in a single pass, allowing for rapid and accurate detection of liver anomalies. YOLOv8s, a specific variant of the YOLOv8 family, was selected for its balance between performance and computational efficiency. The model was trained on a comprehensive dataset of liver images, utilizing advanced augmentation techniques to enhance its generalizability and robustness. The training process involved optimizing the model's parameters to maximize precision, recall, and mean average precision (mAP) rates, ensuring high accuracy in detecting liver disease across diverse image samples.

The experimental results of the proposed method are highly promising, with YOLOv8s achieving a precision rate of 0.94%, a recall rate of 0.96%, and an mAP@0.5 rate of 0.59%. These metrics indicate that the model excels in identifying true positive cases while minimizing false positives and negatives. The high precision rate reflects the model's ability to accurately pinpoint liver anomalies, while the high recall rate demonstrates its effectiveness in detecting the vast majority of disease cases. The mAP@0.5 rate, a comprehensive measure of the model's overall detection performance, underscores the robustness of YOLOv8s in handling various complexities in medical imaging. Compared to other algorithms in the literature, the proposed method shows a marked improvement, highlighting its potential as a reliable tool for liver disease diagnosis.

Despite the strong performance metrics, the study has several limitations that warrant further investigation. One primary limitation is the potential bias in the dataset used for training and validation. The dataset may not cover the full spectrum of liver disease manifestations, potentially affecting the model's generalizability to unseen cases in different clinical settings. Additionally, while YOLOv8s provides high accuracy, the interpretability of its predictions remains a challenge. Medical professionals need to understand the rationale behind the model's decisions to fully trust and adopt this technology in practice. Moreover, the computational requirements for

deploying YOLOv8 models, although optimized, may still be prohibitive in resource-limited environments, restricting its accessibility and widespread use.

Future research should focus on addressing these limitations to enhance the proposed method's applicability and reliability. Expanding the dataset to include a broader range of liver disease cases from diverse populations and imaging modalities will improve the model's generalizability. Developing explainable AI techniques can enhance the interpretability of YOLOv8 predictions, allowing clinicians to understand and validate the model's decisions. Additionally, optimizing the model for deployment on lower-cost hardware will make this advanced technology accessible to a wider range of healthcare settings, including those with limited resources. Research can also explore integrating YOLOv8 with other diagnostic tools to create a comprehensive, multi-modal diagnostic platform for liver disease.

Investigating the integration of YOLOv8 with complementary diagnostic algorithms can provide a holistic approach to liver disease detection. Combining image-based detection with clinical data, such as patient history and biochemical markers, can enhance the diagnostic accuracy and provide a more comprehensive assessment of liver health. Future studies should also explore the longitudinal tracking of liver disease progression using YOLOv8, enabling early detection of disease onset and monitoring treatment efficacy over time. Collaborations with clinical practitioners will be essential to tailor the model's development to meet real-world needs and ensure its seamless integration into existing medical workflows. By addressing these research directions, the proposed method can be refined and validated for broader clinical adoption, ultimately improving liver disease diagnosis and patient outcomes.

VI. Conclusion

This paper studied the critical importance of detecting liver diseases in the field of digital pathology, emphasizing the central role it plays in the domain of medical diagnosis. While numerous methods have been explored in the existing literature, our study highlights the exceptional promise of deep learning techniques, which have demonstrated the capacity to deliver notably accurate results when compared to traditional approaches. Nevertheless, the persistent challenge of low accuracy rates in deep learning-based liver disease detection remains, as indicated by the comprehensive analysis of prior research endeavors. To address this challenge, we have introduced a novel approach harnessing the YOLOv8 algorithm, resulting in the development of innovative models meticulously tailored to elevate the precision and effectiveness of liver disease detection. Our method, involving rigorous model generation, dataset utilization, and extensive experimentation, has yielded accurate outcomes, marking a significant advancement in the field. For future studies, it is imperative to continue refining deep learning models, exploring novel algorithms, and expanding datasets to enhance further the accuracy and robustness of liver disease detection systems. Additionally, investigating the integration of multi-modal data sources, such as imaging and patient records, may offer avenues for comprehensive and holistic disease detection in digital pathology. Moreover, exploring interpretability and explainability in deep learning models can enhance their clinical adoption, ensuring that advancements in this domain contribute effectively to improved patient care and diagnosis.

Acknowledgment

This work was supported by 2020 project of the 13th five-year plan of Educational Science in Shaanxi Province, No: SGH20Y1436.

References

[1] Maier, C. Syben, T. Lasser, and C. Riess, “A gentle introduction to deep learning in medical image processing,” Z Med Phys, vol. 29, no. 2, pp. 86–101, 2019.

[2] M. Rana and M. Bhushan, “Machine learning and deep learning approach for medical image analysis: diagnosis to detection,” Multimed Tools Appl, vol. 82, no. 17, pp. 26731–26769, 2023.

[3] A. Chandy, “A review on iot based medical imaging technology for healthcare applications,” Journal of Innovative Image Processing (JIIP), vol. 1, no. 01, pp. 51–60, 2019.

[4] M. H. Hesamian, W. Jia, X. He, and P. Kennedy, “Deep learning techniques for medical image segmentation: achievements and challenges,” J Digit Imaging, vol. 32, pp. 582–596, 2019.

[5] L.-Q. Zhou et al., “Artificial intelligence in medical imaging of the liver,” World J Gastroenterol, vol. 25, no. 6, p. 672, 2019.

[6] A. B. Chowdhury and K. J. Mehta, “Liver biopsy for assessment of chronic liver diseases: a synopsis,” Clin Exp Med, vol. 23, no. 2, pp. 273–285, 2023.

[7] V. Ajmera et al., “A prospective study on the prevalence of NAFLD, advanced fibrosis, cirrhosis and hepatocellular carcinoma in people with type 2 diabetes,” J Hepatol, vol. 78, no. 3, pp. 471–478, 2023.

[8] I. Grgurevic, T. Bozin, M. Mikus, M. Kukla, and J. O’Beirne, “Hepatocellular carcinoma in non-alcoholic fatty liver disease: from epidemiology to diagnostic approach,” Cancers (Basel), vol. 13, no. 22, p. 5844, 2021.

[9] C.-C. Wu et al., “Prediction of fatty liver disease using machine learning algorithms,” Comput Methods Programs Biomed, vol. 170, pp. 23–29, 2019.

[10] M. M. Musleh, E. Alajrami, A. J. Khalil, B. S. Abu-Nasser, A. M. Barhoom, and S. S. A. Naser, “Predicting liver patients using artificial neural network,” International Journal of Academic Information Systems Research (IJAISR), vol. 3, no. 10, 2019.

[11] A. Spann et al., “Applying machine learning in liver disease and transplantation: a comprehensive review,” Hepatology, vol. 71, no. 3, pp. 1093–1105, 2020.

[12] H. Yu, L. T. Yang, Q. Zhang, D. Armstrong, and M. J. Deen, “Convolutional neural networks for medical image analysis: state-of-the-art, comparisons, improvement and perspectives,” Neurocomputing, vol. 444, pp. 92–110, 2021.

[13] A. S. Rahman, F. M. J. M. Shamrat, Z. Tasnim, J. Roy, and S. A. Hossain, “A comparative study on liver disease prediction using supervised machine learning algorithms,” International Journal of Scientific & Technology Research, vol. 8, no. 11, pp. 419–422, 2019.

[14] I. Altaf, M. A. Butt, and M. Zaman, “Disease detection and prediction using the liver function test data: A review of machine learning algorithms,” in International Conference on Innovative Computing and Communications: Proceedings of ICICC 2021, Volume 2, Springer, 2022, pp. 785–800.

[15] Z. Yao, J. Li, Z. Guan, Y. Ye, and Y. Chen, “Liver disease screening based on densely connected deep neural networks,” Neural Networks, vol. 123, pp. 299–304, 2020.

[16] P. Jain, “Liver Disease Detection from CT scan images using Deep Learning and Transfer Learning. ,” National College of Ireland., 2020.

[17] J. C. Ahn, A. Connell, D. A. Simonetto, C. Hughes, and V. H. Shah, “Application of artificial intelligence for the diagnosis and treatment of liver diseases,” Hepatology, vol. 73, no. 6, pp. 2546–2563, 2021.

[18] S. Mammeri, M. Amroune, M.-Y. Haouam, I. Bendib, and A. Corrêa Silva, "Early detection and diagnosis of lung cancer using YOLO v7, and transfer learning," Multimed Tools Appl, pp. 1–16, 2023.

[19] J. Dana et al., "Conventional and artificial intelligence-based imaging for biomarker discovery in chronic liver disease," Hepatol Int, vol. 16, no. 3, pp. 509–522, 2022.

[20] T. K. R. Agita et al., "Detection of Disease in Liver Image Using Deep Learning Technique," in International Conference On Emerging Trends In Expert Applications & Security, Springer, 2023, pp. 285–298.

[21] J. Solawetz, "What is yolov8? the ultimate guide," Roboflow Blog, vol. 25, 2023.

[22] R. Bai, F. Shen, M. Wang, J. Lu, and Z. Zhang, "Improving detection capabilities of YOLOv8-n for small objects in remote sensing imagery: towards better precision with simplified model complexity," 2023.

[23] "OpenMMLab Dive into YOLOv8: How does this state-of-the-art model work?"

[24] H. Yu et al., "Artificial intelligence-based liver portal tract region identification and quantification with transplant biopsy whole-slide images," Comput Biol Med, vol. 150, p. 106089, 2022.