International Journal of Computational and Experimental Science and ENgineering (IJCESEN)

Vol. 11-No.2 (2025) pp. 2053-2063
http://www.ijcesen.com

ISSN: 2149-9144

Research Article

Advanced Liver Tumour Detection Using Optimized YOLOv8 Modules

Mahendran S[1*], Venkatasekhar D[2], Shanmugasundaram G[3]

[1*] Research Scholar, Department of Information Technology, Annamalai University, Chidambaram, India
* **Corresponding Author Email:** smpdy963@gmail.com **ORCID:** 0009-0004-9733-5751

[2] Professor, Department of Information Technology, Annamalai University, Chidambaram, India
Email: ramavenkatasekar@yahoo.co.in **ORCID:** 0000-0001-8721-5028

[3] Professor, Department of Computer Science and Engineering, Chennai Institute of Technology, Chennai, India
Email: anitguy2006@gmail.com **ORCID:** 0000-0002-3231-4923

Article Info:

DOI: 10.22399/ijcesen.1613
Received : 23 February 2025
Accepted : 04 April 2025

Keywords :

liver tumour detection,
benign tumour,
malignant tumour,
YOLOv8.

Abstract:

Health fraternity is invariably challenged with early diagnosis, detection, identification, classification, treatment and convalescence of globally prevalent and life-threatening fatal diseases as liver cancer. The early detection of liver cancer through medical image processing technique is so challenging that an iota of deviation conspicuous among healthy tissues, benign tumour and malignant tumour tissues is a matter of wake up call. This work is entailed with introduction of a novel, optimized YOLOv8-based model for liver tumour detection, harnessing the strengths of transformer-based feature extraction, global attention mechanisms, and advanced feature aggregation techniques. The model was subjected to rigorous performance with relevant methods and messages as parameters time and again for repeated refinements. Eventually, it was concluded that the proposed model surpasses all the models in extant now in terms of precision, recall, and means average precision (mAP). This is ascertained by inference drawn from the model's achievement of attaining 95.34% precision, 96.49% recall, and 97.31% mAP@0.5. In regard to tumour classification, the proposed model excels in differentiating normal cases, benign tumours, and malignant tumours. These innovations represent a significant step toward improving the accuracy of automated liver tumour diagnosis systems, with the potential to revolutionize clinical workflows and enhance patient outcomes.

1. Introduction

The global medical field invariably encounters with ubiquity of potentially fatal hepatocellular carcinoma. It is one of the most critical illnesses that are causing death. World Health Organization has disclosed that hepatic cancer is next to top most cause of death of all cancerous diseases [1]. It demands a futuristic method to detect tumours in liver for effective diagnosis and remedy. Accordingly, precision in review of tumors and analysis of them before becoming in critical phase are vital. It is nothing but a challenge to the stakeholders to distinguish between healthy tissues and tumor tissues as there are a little bit noticeable deviations towards detection of liver tumor [2]. Harnessing with Artificial Intelligence and medical imaging advancements, medical images show great promise of diagnosing with hepatic tumors. A slew of methodologies has been employed to address the challenges in detecting hepatic tumors by the employment of image processing techniques, machine learning approaches, and, lately, deep learning-based frameworks [3]. In contrast to traditional methods, machine learning techniques get rid of the limitation of ability in generalization and robustness by making use of decision trees and support vector machines. Machine learning approaches have fine-tuned the abilities to diagnose [4]. At the same time, they have their own constraints due to dependence on feature engineering and precision in quality data feeding.
Intense learning has been incipient as the mainstream approach for target detection algorithms by taking into account of its excellent potentiality in comparison with traditional methods [5]. The algorithms pertaining to deep learning-based detection comprise of two-stage and one-stage approaches. Such two-stage algorithms as Faster-RCNN series, entail region generation by pre-

selecting boxes that may contain objects, followed by feature map generation and sample classification through convolutional networks, achieving high accuracy but with longer detection times. But YOLO series of one-state algorithms kind get rid of the necessity for region generation by an immediate prediction of object classification and location by means of convolutional networks, considerably lessening detection time [6].

The object detection has revolutionized by leveraging the potentials of YOLO series of models in combination by means of combined parameters of accuracy and speed in a single-stage detection framework. The updated version of YOLOv8 sets up on its earlier architectures by exploiting state-of-the-art improvements in feature extraction, detection accuracy, and computational efficiency. The architecture of YOLOv8 is depicted by a three-part design. The primary one is the Backbone meant for feature extraction; the secondary is the Neck meant for feature aggregation; and the tertiary is the Head meant for final object classification and bounding box regression. Apart from its robustness, YOLOv8's potential in detecting small and irregularly shaped objects, such as liver tumours, can be further augmented by the employment of targeted optimizations [7,8].

The emergence of deep learning has made a radical change in detecting liver tumour by the engagement of end-to-end learning and automation of feature extraction out of raw imaging data. Convolutional Neural Networks (CNNs) have proved to be effective in the tasks of segmentation, classification, and object detection related to medical image analysis by its state-of-the-art performance [9,10]. Nevertheless, the multifarious tumours in respect of its sizes, irregular shapes and heterogeneous image contrast pose significant challenges. The bottlenecks of these call for refining the optimization of existing frameworks to attain higher accuracy and robustness in diverse clinical scenarios.

Improvement of medical image analysis has acquired new avenues with recent advancements in transformer-based architectures and attention mechanisms [11]. The employment of these components into framing of detection modes has exhibited a good result in understanding unique challenges of liver tumour detection [12]. Especially, the properties of efficiency and accuracy in object detection tasks that have been inherited from its earlier version of YOLOv8 framework, serve as a strong foundation for incorporating these advanced methodologies [13]. This work concentrates on attaining maximum optimum of YOLOv8 by the amalgamation of the Swin Transformer, Global Attention Mechanism (GAM), augments to the Path Aggregation Network (PANet), and attention mechanisms post-up-sampling layers to combat the challenges inherent in liver tumour detection and ramp up diagnostic outcomes.

In order to leverage the hierarchical feature extraction capabilities amalgamation of the Swin Transformer is set at the P3 position, thereby permitting the model to capture both local and global contextual information effectively. This is paramount to liver imaging where obscured variations in texture and structure can point out the presence of disease. Furtherance to the above, the phenomenon of Global Attention Mechanism (GAM) at position P5 ramps up the capability of the model to pay attention on relevant features across the entire image, allowing it to grab long-range dependencies and contextual information. The model is aided the GAM to enhance the ability of detecting liver tumours and heterogeneities.

Better performance of multi-scale feature fusion is smoothened by another critical optimization which is attained by augmenting the PANet structure in the Neck section. This stepping up makes sure that the model can potentially make use of both low-level and high-level features, which is necessary for accurate detection of tumours of diverse sizes and characteristics. Besides, combining attention mechanisms after up-sample layers in the Neck section permits the model to fine-tune its point of concentration on salient features, thereby enhancing detection accuracy and robustness against noise and artefacts commonly found in medical images.

This work presents a new-fangled, maximum effective, perfective and functional YOLOv8 framework for detection of hepatic liver lesion, conglomerating the potentials of transformer-supported feature extraction, all stakeholder's attention mechanisms, and state-of-the-art feature aggregation techniques. These innovations manifest a momentous step toward escalating the accuracy and integrity of automated hepatic lesion diagnosis systems, with the capability to revolutionize clinical workflows and augment patient outcomes.

The rest of this paper is systematically depicted as under. Section 2 imparts an inclusive review of pertained work, exposing salient features of developments in liver lesion detection and neo-innovations in YOLO-based object detection models. Section 3 outlines the methodology, entailing the architectural modifications and bringing into use of strategies. Section 4 talks about the outcome obtained from comparison of the performance of the proposed model with baseline YOLOv8 and state-of-the-art approaches. Eventually, Section 5 concludes the work, abstracting the implications of the findings and future research directions.

2. Literature Review

The procedures in vogue now for adoption of detecting hepatic lesions commonly comprehend an estimation of medical past events, physical scrutiny, lab test observations, and medical imaging [14]. The primary stage in diagnosing hepatic malady entails getting an inclusive medical history, comprehensive symptoms, ancestral history, and risk factors pertained to liver disease [15]. A physical scan can detect hepatic distension or tenderness, while haematic tests evaluate hepatic function and correlate damage markers.

Therapeutic imaging agents like ultrasound, MRI, and CT scans are helpful to radiographic visualization of the liver and the identification of abnormal functionalities. Ultrasound is preferred to examine the hepatic architecture as it involves no invasive procedure [16]. On the other hand, CT scans produce complex cross-sectional views for lesion identification and other liver malfunctions. The procedure involving these techniques finds widespread employment in medical diagnostics. Conventional methods have their virtues, but state of the arts in medical technology and research have made conducive for the enhancement of more advanced and less invasive techniques for liver disease detection and diagnosis.

Numerical methods, entailing artificial intelligence, machine learning, and data analytics, have considerably augmented the identification and diagnosis of hepatic disorders. These methods are able to have a thorough analysis of medical data, recognizing patterns, and coming out with accurate predictions. Amalgamation of these technologies is fructuous to patient care and consequences. Machine learning algorithms, directed on huge datasets of medical records, laboratory results, and imaging data, can identify patterns pertained to various hepatic illnesses. These algorithms can categorize new patient data and help in determining specific hepatic conditions [17]. Deep learning, a subset of machine learning, makes use of neural networks to make out complex representations from data. Convolutional neural networks (CNNs) have been employed to medical imaging to ascertain the detection of liver lesions and abnormalities. These methods can automatically locate biomarkers from intricate datasets suggestive of hepatic illness, which can be put into use of diagnostic models or well-in-advance warning indicators [18]. Compilation of data from clinical data, laboratory results, medical imaging, and genetic information can extend the frontier of the accuracy and reliability of hepatic malady detection and diagnosis [19,20].

Chen et al [21] put forward a modern method to trace the edge of stroke lesions utilizing YOLOv5. Sapitri et al [22] presented a real-time fetal cardiac substructure detection approach on the basis of the YOLO frame work, explicating the versatility of the YOLO architecture in puzzling out complex medical imaging tasks. These advancements in YOLO-based medical image analysis underline the scaling up importance and prospect of this object detection framework in driving innovative solutions for clinical decision-making and patient care.

Deep learning techniques have proved to be significant potential in detecting hepatic illnesses through medical imaging [23], accurately recognizing a slew of hepatic conditions. Nonetheless, sustaining consistently great accuracy poses a challenge, especially with complex disease patterns. Iterative refinement and optimization are prerequisite to augment model performance, manifest these limitations, and enhance diagnostic precision, eventually supporting better clinical decision-making.

Existing state of the arts in machine learning and computer vision have stimulated interests in automatic translation of medical image. Bygone research works portray the efficacy of deep learning in assorting lesion types, but works on YOLOv8—a real-time object detection model—for liver tumour screening lays restricted. Majority of prevalent research work attunes at deep learning architectures or general object detection, accentuating the need for further examination into YOLOv8's efficacy in hepatic lesion detection. AI-propelled methods, especially CNNs, have enhanced lesion detection and categorization in histopathology images. A thorough analysis of YOLOv8's speed and efficiency is necessitated to manifest specific challenges in hepatic lesion diagnosis.

This research work underscores the salient features of YOLOv8's architectural augmentations, comprising improved feature of extraction, potentially heightened detection accuracy in medical imaging. The desegregation of transformer architectures additionally ameliorates the model's capability to grab complex features in liver scans, while attention mechanisms upgrade the point of concentration on relevant tumour characteristics. Comparative analyses suggest that YOLOv8 supersedes the performance over traditional CNNs and other object detection frameworks in terms of both digital correctness and computational speed. These advancements juxtapose atop YOLOv8 as a most probable succeeding tool for real-time clinical applications in hepatic lesion diagnosis.

3. Methodology

Figure 1 demonstrates the configuration of YOLOv8 model [24] composed of three major constituents

namely, Backbone, the Neck, and the Head. The Backbone is assigned with feature extraction from input images. Centred in the middle of the Backbone and Head sections, the Neck part escalates the expressive capacity of these components by deducing and ameliorating information across many layers. Heightening of this capacitates the Head section to determine the exact locations and classifications of varied objects inside the image.

The default YOLOv8 model's standard Backbone component makes use of Feature Pyramid Network (FPN) architecture to inherit features extraction from the image given as input. P1 and P2 denote lower-level features. They are in possession of complicated local details, bestowing them particularly proficient in recognizing small-sized items. On the other hand, P4 and P5 possess higher-level traits that enwrap more sophisticated global information, bestowing them for larger objects identification. Unlike conventional medical pictures, CT scans meant for hepatic lesion identification display various iota-sized tumour areas. Nevertheless, the default Neck component of YOLOv8 insufficiently exploits these lower-level traits that are ideal for locating small objects. There is also a potential threat in determining the size and appearance of hepatic lesions subject to changes of parameters in CT scanning and imaging angles. In the wake of employment of optimized YOLOv8 model as shown in the Figure 2, detection of small tumours and consistent accuracy in identifying tumours of different sizes and characteristics are ensured.

The quadric-modular optimization modules entail with different components of the YOLOv8 backbone's Feature Pyramid Network (FPN) architecture and the Path Aggregation Network (PANet) structure in the Neck portion. The YOLOv8 set up files pave way for the straightforward implementation, modification, and extension of these modules. The subsequent subsections will expound optimization techniques and the specifics for implementation of each component comprehensively.

3.1. Incorporation of GAM at P5 Position for Enhanced Feature Capture

The Global Attention Mechanism (GAM) [25] is a state-of-the-art technique formulated to augment model's capability to target on the features closely akin to an image, particularly when dealing with complex tasks like hepatic lesion detection. Ascertainment of regions within having subtle, frequently small, abnormalities like hepatic lesions, which are prone to oversight because of their size resemblance to surrounding healthy tissue, is pivotal in medical imaging. GAM commissions varied attention weights across different regions of the image, permitting the model to selectively target on key areas—such as the abnormal tissue patterns or lesions—that are vital for accurate tumour detection. By escalating the representation of significant features, GAM facilitates the model to grab more precise, informative details necessary for diagnosing liver tumours, regardless of their stage or size. As illustrated in Figure 2, enhancement presents the GAM module at P5 feature level position of the Backbone section. By integrating GAM at this crucial feature level, we ensure that the network focuses on the most relevant regions at multiple scales, effectively bettering tumour detection accuracy. The SPPF module process the enriched feature maps for further optimization, which provides even more refined features for subsequent processing in the Neck section. Hepatic lesions can differ significantly in size, shape, and texture, and often appear subtle, especially in early-stage diagnoses. GAM's propensity to concentrate on censorious region facilitates YOLOv8 to grab lesions more precisely regardless of their size which are diminutive, indistinct, or adjacent to analogous structures. This results in higher detection accuracy, facilitating more precise, prompt diagnosis, and augment patient outcomes.

3.2. Integrate Swin Transformer at P3 Position for Feature Extraction

In state-of-the-art developments in vision transformers, the powerful architecture has emerged in the form of Swin Transformer [26], especially surpassing in medical image analysis and prediction tasks. In contrast to traditional transformers that utilise global self-attention throughout every aspect of an image, the Swin Transformer introduces a hierarchical architecture with shifted windows, facilitating a compromise between computational efficiency and feature representation capability. This design allows efficient multi-scale processing, making it highly suitable for hepatic lesion detection, where both small and large lesions need to be accurately identified within high-resolution medical images. Swin Transformer fundamentally relies on the concept of shifted window-based self-attention, mathematically defined as shown in equation 1.

$$Attention(Q,K,V) = Softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \qquad (1)$$

Where Q, K, and V represent the query, key, and value matrices derived from the input feature maps within each window. $\sqrt{d_k}$is the dimensionality of the key vectors, used for scaling to stabilize gradient updates. The shifted window mechanism is devised in such a way that each local self-attention window interacts

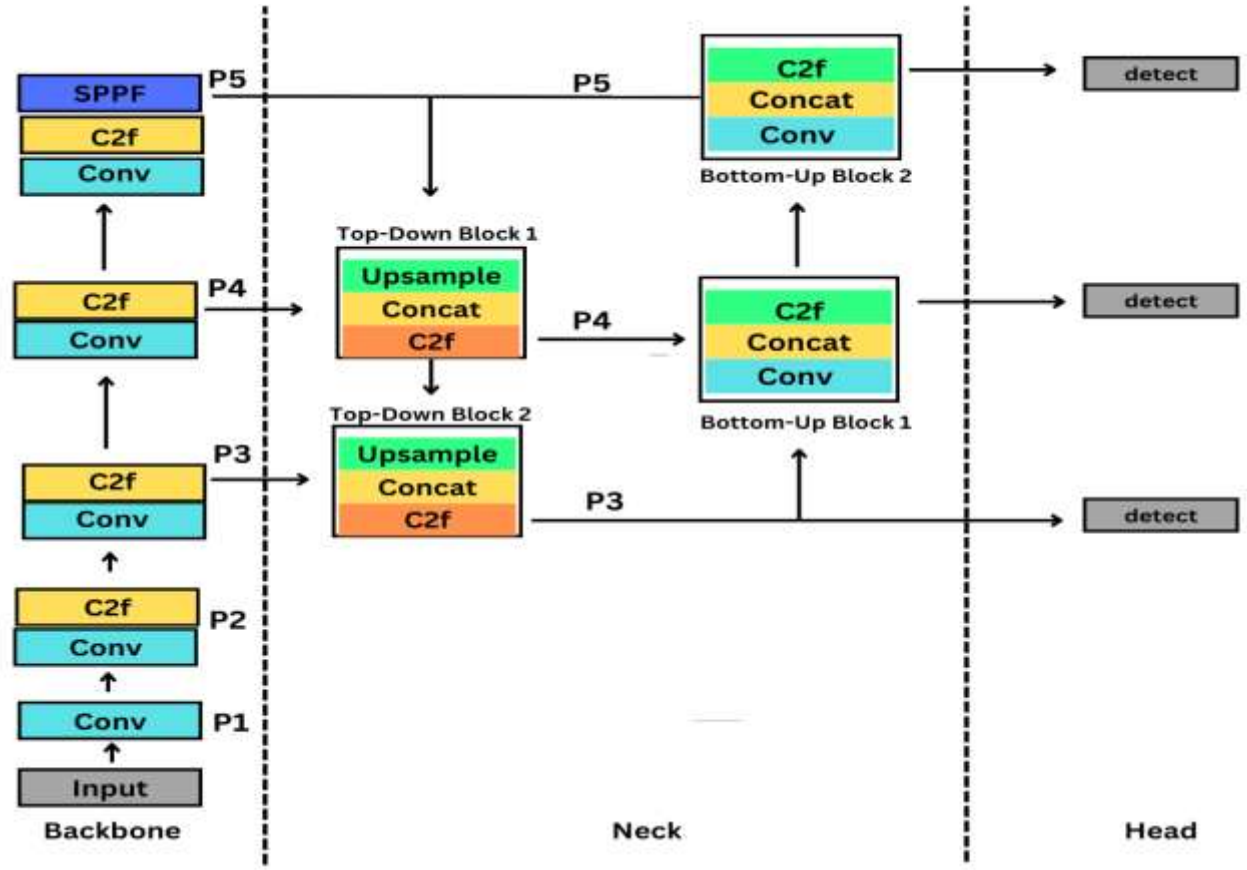

Figure 1. *Structure of YOLOv8 Architecture*

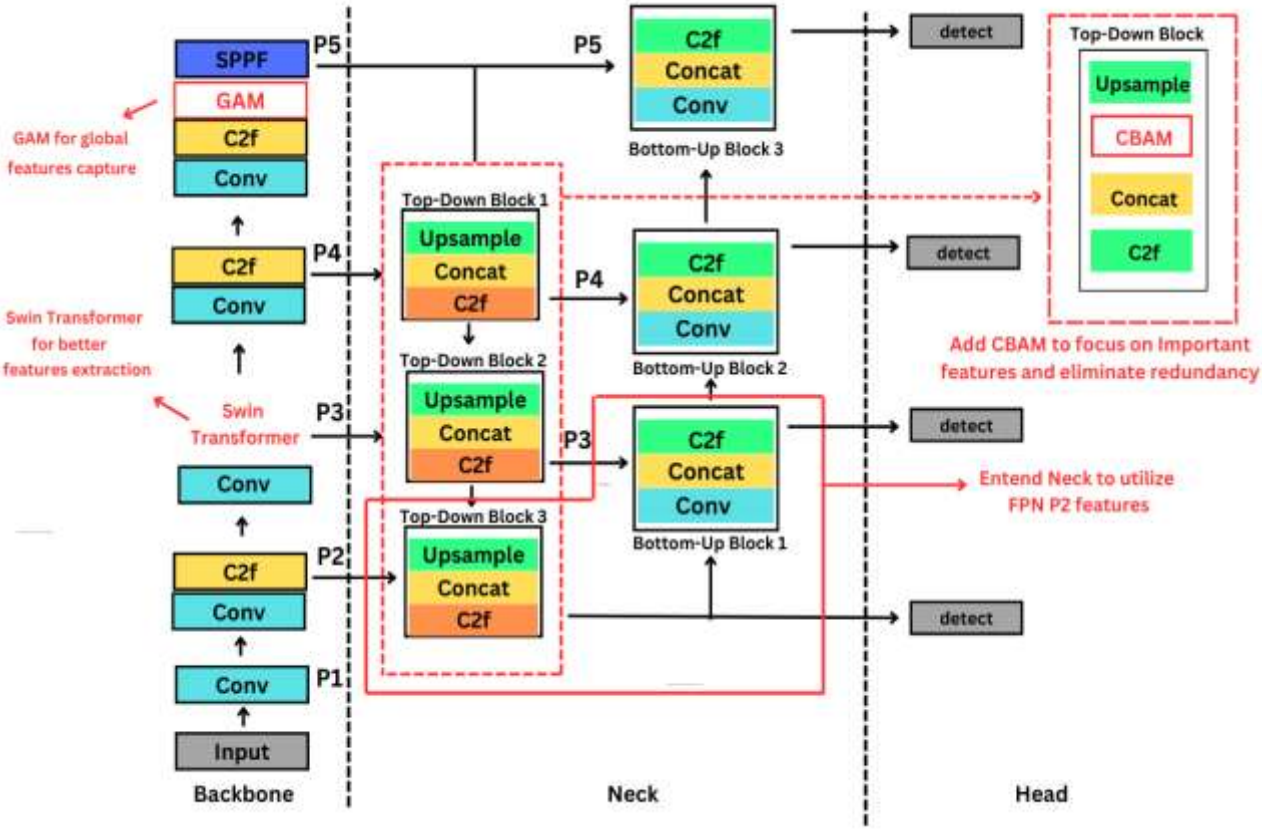

Figure 2. *Optimization Structure of YOLOv8 Architecture*

with adjacent windows, triggering the grab of long-range dependencies. This mechanism is particularly advantageous for hepatic lesion diagnosis, as it stimulates the model to store fine-grained lesion details while ascertaining that the contextual information from surrounding liver tissues is effectively leveraged. As highlighted in Figure 2, our architectural enhancement entails replacing the C2f module at P2 in the Backbone Feature Pyramid Network of YOLOv8 with the Swin Transformer module. This optimization scales up the feature extraction process, permitting the model to:

- Effectively desegregate multi-scale lesion features, bettering the detection across varying lesion sizes and locations.
- Deescalate computational complexity while saving detailed liver tissue structures, essential for early-stage liver cancer detection.
- Improve lesion boundary delineation, decreasing false positives by focusing on discriminative features in CT scans.

By making most of the hierarchical self-attention mechanism of the Swin Transformer, YOLOv8 attains advantages a more precise and adaptive feature representation, consequentially augmenting the sensitivity and specificity of hepatic lesion detection and diagnosis. This improvement is critical in clinical applications, where accurate lesion classification can directly impact patient outcomes.

3.3. Desegregating CBAM in Top-down block of Neck section to eliminate redundancy and concentrate on key features

Though the Convolutional Block Attention Module (CBAM) [27,28] is lightweight, still it is an effective mechanism that escalates feature representation in deep neural networks by dynamically accentuating informative features across both channel and spatial dimensions. It consists of channel attention and spatial attention as two components working together to refine feature responses and scale up the model's discriminative power. Choosy on accentuating potential features pertaining to liver lesion while dimming non-abstract details, CBAM drives the model to concentrate effectively on lesion-specific characteristics, leading to enhanced localization and classification accuracy. This goal of refinement set for is particularly inevitable in medical imaging, where healthy tissue differentiated from tumour regions is a must for exact liver tumour detection and diagnosis.

The channel attention is employed at first to locate and amplify the top most feature channels and then spatial attention is engaged to determine the most relevant regions within the feature maps when CBAM applies a sequential attention mechanism. Computation of Channel attention takes place to feed feature map $Z \in R^{H \times W \times C}$ by the application of global average pooling (GAP) and global max pooling (GMP). These pooled features are then passed through two fully connected layers (FCL_1 and FCL_2), triggered by ReLU, and processed by function sigmoid to yield channel-wise attention weights α_{ca} as shown in equation 2.

$$\alpha_{ca} = \sigma(FCL_2(ReLU(FCL_1(GAP(Z))))) \quad (2)$$

Spatial attention M_{sa} subject to further refinement of the feature maps by precisely locating the most significant spatial regions on spurring the channel attention. This is achieved due to application of a 1×1 convolution to the concatenation of GAP and GMP outputs, making sure that both extreme and aggregate spatial cues contribute to the refinement process as shown in the equation 3.

$$M_{sa} = \sigma(Convol_{1\times1}([GMP(Z), GAP(Z)])) \quad (3)$$

The final attention-enhanced feature map is yielded via element-wise multiplication between the original input Z, the channel attention weights α_{ca} (broadcasted to match the shape of Z), and the spatial attention mask M_{sa} as shown in the equation 4.

$$Z_{CBAM} = Z \odot (\alpha \otimes Msa) \quad (4)$$

Highly refined feature representation is produced by the fusion of channel and spatial attention synergistically, paving a way for the network to concentrate on critical features eliminating inapposite information, resulting in augmenting detection accuracy. In an attempt to obtain maximum of liver tumour detection within the YOLOv8 framework, CBAM is resorted to strategic desegregation of the interval between the Upsample and Concat layers in the Neck. This placement makes sure that upsampled feature maps are subjected to choose refinement before being combined, allowing the model to hold the most pertinent spatial and channel-based information. By getting rid of noise and escalating discriminative features, CBAM ramps up feature integration, enabling subsequent processes such as concatenation and context-to-feature (C2f) transformation to better utilize the optimized feature representations. Eventually, the model yields more accurate localization and classification of liver tumours, potentially increasing its diagnostic capabilities.

3.4. Extending PANet Structure in neck part to utilize P2 features

The Neck section engages Path Aggregation Network (PANet) based structure to improve multi-scale feature fusion by refining information flow

between feature maps taken out from the Backbone's Feature Pyramid Network (FPN). The PANet structure presents additional bottom-up and top-down pathways, permitting the network to consolidate and refine features across different scales as illustrated in Figure 2. This structure is especially beneficial for liver tumour detection, where both spatial details and advanced semantic features are crucial for precise localization and classification.

In the top-down pathway, multiple modules work together to refine features:

- Upsample module ramps up resolution to retain finer details.
- Concat module proves lateral connections for improved feature integration.
- C2F (Context-to-Feature) module augments feature processing and fusion.
- Conv module parts with refined, high-quality features for better discrimination.

The connections on both end but in opposite direction make sure of perpetual advancement of refinement of features from top to bottom levels vice versa, permitting the model to make most of both detailed local structures and global contextual information, which is indispensable to lesions detection.

Nonetheless, as per Figure 2, YOLOv8 default Neck only concatenates bottom-up and top-down blocks to the P5, P4, and P3 layers of the Backbone FPN, leaving direct connections to the P2 layer, which possesses detailed spatial features. Though this structure proves effective for general object detection, the default of direct integration with the P2 layer may contain the detection of small and subtle tumour regions, detrimental to the model's sensitivity in medical imaging.

In order to surmount this shortcoming, we propose to engage bottom-up and top-down blocks in the section of Neck as shown in Figure 2, highlighted by the red solid lines. The engagement of these new blocks extending the original PANet pathway (shown in black) without any ambiguity connects and blend features from the FPN P2 layer. By incorporating low-level spatial features, this modification elevates the model's capability to detect indistinct and iota of tumours, paving a way to more and more an accurate localization and improved classification of liver tumours in medical imaging.

The enhancement of YOLOv8 for hepatic lesion diagnosis is evident from amalgamation of advanced feature extraction and attention mechanisms. GAM at P5 firms up high-level semantic features, while the Swin Transformer at P3 augments spatial feature learning. The enhanced PANet structure ensures effective multi-scale feature aggregation, and post-upsample attention mechanisms refine discriminative features for precise tumour localization. These modifications make the model more accurate and reliable for automated liver tumours diagnosis in medical imaging.

4. Result and Discussion

This section deals with presentation of observational outputs and results collected from optimized YOLOv8 model for hepatic lesion detection. From the Figure 3, the output from the model delves into deep penetrative details of its capability to locate and identify different classes, entailing normal liver, benign tumour, and malignant tumour. These classes characterize serious clinical conditions that call for the precise detection in order to medical diagnosis. The figures exhibit the model's predictions and constitute its significant ability to set out between these varied tumours.

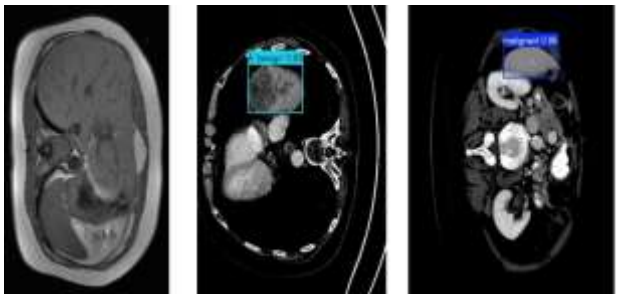

Figure 3. *(a) Normal Liver (b) Benign Tumour (c) Malignant Tumour*

A rigorous analysis on the area distribution and sample counts of varied hepatic tissue classes in the dataset was unleashed for the enhancement of experimental effectiveness as depicted in Table 1. The normal liver category tops all in terms of average area, furnishing a strong reference for classification. The benign and malignant tumours also display sufficient variability, which aids in understanding various patterns.

Table 1*. Statistical overview of class annotations*

Statistical Data	Normal Liver	Benign Tumour	Malignant Tumour
Min. area (in pixels)	26298.31	12251.14	103.14
Max. area (in pixels)	121780.34	186034.92	76641.50
Avg. area (in pixels)	65081.92	58019.53	7270.02
Train count (images)	721	4739	6515
Val count (images)	75	458	623
Train Val ratio (%)	9.61	10.35	10.46

Further, the train-to-validation ratios maintain consistency across all categories, solidifying a well-structured dataset for model training. These characteristics are paramount importance to a comprehensive and balanced hepatic lesion identification framework.

4.1. Model-wise Performance Analysis

The evaluation of hepatic lesion detection across normal liver, benign tumour, and malignant tumour categories as illustrated in the Table 2 on performance of the proposed model and different YOLOv8 variants was conducted. For category-wise classification, the proposed model surpassed in distinguishing Normal (0.9261), Benign (0.9946), and Malignant (0.9905) cases, superseding all other models.

Table 2. *Comparative Analysis of YOLOv8 Variants and the Proposed Model*

Model	Normal Liver	Benign Tumou	Malignant Tumour
YOLOv8n	0.9000	0.9800	0.9700
YOLOv8s	0.9243	0.9938	0.9897
YOLOv8m	0.9180	0.9921	0.9882
YOLOv8l	0.9102	0.9903	0.9856
Proposed Model	0.9261	0.9946	0.9905

Analyses on the basis of crucial evaluation metrics such as recall, precision, and mean Average Precision (mAP) as illustrated in the Table 3 resulted with effectiveness of each model variants and the proposed model. The proposed model's performance was phenomenal and achieved the highest precision (95.34%) and recall (96.49%), surmounting all YOLOv8 variants. It also recorded the highest mAP@0.5 (0.9731), superseding YOLOv8s (0.9691), YOLOv8m (0.9600), and YOLOv8l (0.9500), implying superior detection capabilities. The mAP@0.5-0.95 score of 0.8604 is competitive with YOLOv8s (0.8617) and YOLOv8m (0.8600), demonstrating consistent performance across different confidence thresholds. YOLOv8s and YOLOv8m demonstrated close performance, but the proposed model proved a slight edge in precision and recall, ensuring more reliable tumour detection.

Table 3. *Model-wise Performance Analysis*

Model	Precision(%)	Recall (%)	mAP @0.5	mAP@ 0.5:0.95
YOLOv8n	92.00	96.00	0.9200	0.5800
YOLOv8s	94.14	95.42	0.9691	0.8617
YOLOv8	94.00	96.00	0.9600	0.8600
YOLOv8l	93.00	95.50	0.9500	0.8550
Proposed Model	95.34	96.49	0.9731	0.8604

The results emphasize the effectiveness of the proposed model in liver tumour detection, benefiting from optimized feature extraction and classification strategies. Its superior accuracy and consistency make it a propitious approach for automated liver tumour diagnosis.

The assimilation between different YOLOv8 models (YOLOv8n, YOLOv8s, YOLOv8m, and YOLOv8l) and the model proposed in terms of precision, recall, mAP@0.5 and mAP@0.5:0.95 scores is depicted in Figure 4. The mAP@0.5 indicates a steady increase from YOLOv8n to YOLOv8s, reaching peak performance with YOLOv8m and YOLOv8l, while the proposed model maintains a competitive score, signalling high detection accuracy. The mAP@0.5:0.95 reveals a substantial enhancement from YOLOv8n to YOLOv8s and remains stable across YOLOv8m, YOLOv8l, and the proposed model, proving strong generalization across different IoU thresholds. Overall, the proposed model performs on par with or slightly better than YOLOv8l, highlighting its effectiveness in object detection while maintaining robustness and accuracy.

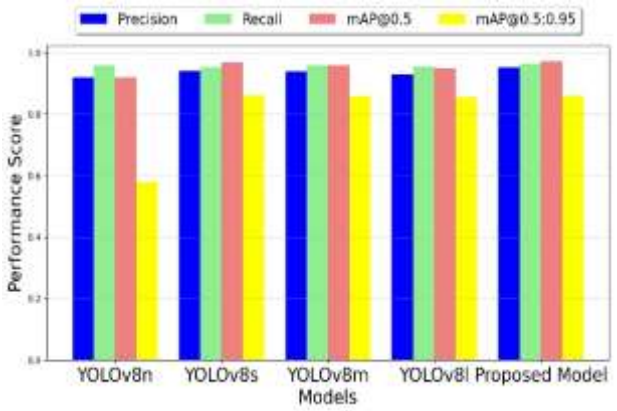

Figure 4. *Performance metrics of YOLOv8 variants and the proposed model*

4.2. Performance of Class-wise Analysis

Table 4 depicts the performance of class-wise proposed model in liver tumour detection. It was evaluated by putting into use of key metrics like recall (R), precision (P), mean Average Precision (mAP) at different thresholds, the number of images and instances per class.

The overall performance across all classes is overarching, with a precision of 95.34%, recall of 96.49%, and a high mAP@0.5 of 0.9737, insinuating the model's ability to accurately detect and classify liver conditions. The mAP@0.5:0.95 score of 0.8593 further illustrated the model's consistency across varying Intersection over Union (IoU) thresholds.

- **Normal Liver**: The proposed model was subject to rigorous performance with 70

images and instances. It yields a precision of 94.20% and recall of 92.86%. The mAP@0.5 is 0.933, while the mAP@0.5:0.95 is lower at 0.793, insinuating slightly reduced performance in capturing fine-grained variations in normal liver images.

- **Benign Tumour**: The proposed model was subject to rigorous performance with 452 images and instances. It brings about high accuracy, attaining 99.56% precision and 99.12% recall. The mAP@0.5 is 0.994, and the mAP@0.5:0.95 is 0.875, reflecting robust detection and classification of benign liver conditions.
- **Malignant Tumour**: This class tops of all instances (1100) across 625 images. The model yields a precision of 98.99% and recall of 98.27%, with mAP@0.5 of 0.994 and mAP@0.5:0.95 of 0.910. The high scores are suggestive of strong sensitivity and precision in identifying malignant liver conditions.

Table 4. *Class-wise Performance*

Class	Images	Instances	Precision (%)	Recall (%)	mAP @0.5	mAP @0.5 :0.95
All	1077	1622	95.34	96.49	0.974	0.859
Normal	70	70	94.20	92.86	0.933	0.793
Benign	452	452	99.56	99.12	0.994	0.875
Malignant	625	1100	98.99	98.27	0.994	0.910

Figure 5 depicts the performance metrics across classes such as normal liver, benign tumour, malignant tumour and all. On the whole, the results of the proposed model's efficiency in liver tumour classification, with particularly high performance in detecting benign and malignant cases are phenomenal. While the normal liver class sports slightly lower mAP@0.5:0.95, the overall accuracy remains competitive, making the model well-suited for practical medical applications. The superior performance of the proposed model can be attributed to its optimized YOLOv8 architecture, which incorporates advanced feature extraction and classification techniques. The improvements in precision (95.34%) and recall (96.49%) further highlight its reliability in detecting liver tumours with high accuracy. These results indicate that the proposed model is a highly effective solution for automated liver tumour detection, offering better generalization and robustness than existing approaches.

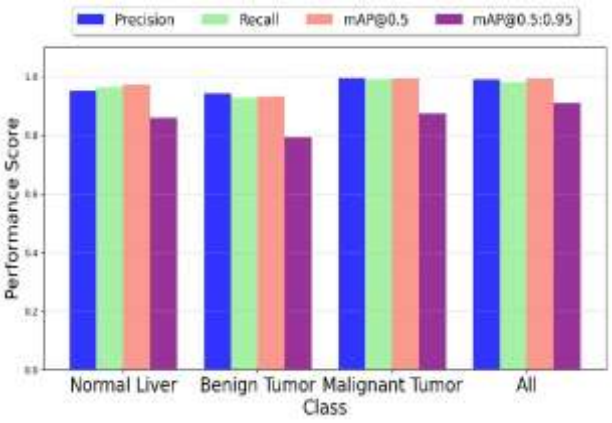

Figure 5. *Performance metrics comparison across classes*

5. Conclusion

The entirety of this work contributes to an optimized YOLOv8-based model for hepatic lesion detection, surpassing models prevalent in terms of recall, precision, and means average precision (mAP). The model proposed brought forth 95.34% precision, 96.49% recall, 97.31% mAP@0.5, and 86.04% mAP@0.5:0.95, proving superior performance compared to YOLOv8 variants and previous deep learning models.

The proposed model is up to the mark of radiologists to identify liver conditions through effective disintegration. It paves way for tumour classification by achieving 0.9261 mAP, 0.9946 mAP and 0.9905 mAP respectively for healthy cases, benign tumour and malignant tumours. The results obtained from this model proclaim its potentiality for a precise detection of liver abnormalities, which exhibits consistency in producing higher accuracy, upholding enhanced diagnostic reliability.

Significant enhancement of detection accuracy and sustained computational efficiency are the fortes of optimized YOLOv8 and hence it is preferable for real-time liver tumour screening applications. The improved model architecture features enhanced speed and precision in the detection of hepatic abnormalities, minimizing eventualities of misdiagnosis and augmenting early detection abilities.

On the other hand, there is a desperate need of iterative refinements to ensure its enhanced clinical applicability and real-world reliability. For winning clinicians' confidence in the interpretation of predictions, the model transparency has to be enhanced by harnessing explainable AI techniques and expanding the dataset to include diverse populations and imaging modalities of generalizing. The employment of minimum cost hardware by optimizing the model will pave way for flexible

accessibility in resource-limited settings, enabling broader adoption in healthcare. Furthermore, conglomerating YOLOv8 with patient clinical data, biochemical markers, and complementary diagnostic tools can deduce a multi-modal framework for liver disease assessment, improving diagnostic precision. Future studies should also explore longitudinal disease tracking to enable early detection and real-time monitoring of disease progression. Close collaboration with healthcare professionals is essential to refine the model for seamless integration into clinical workflows, ensuring its practicality in real-world settings. Addressing these future directions will further strengthen the model's potential as a comprehensive, efficient, and scalable solution for liver tumour diagnosis, ultimately improving patient outcomes.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** The authors declare that they have equal right on this paper.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

[1] Devarbhavi, H., Asrani, S. K., Arab, J. P., Nartey, Y. A., Pose, E., & Kamath, P. S. (2023). Global burden of liver disease: 2023 update. *Journal of Hepatology*, *79*(2), 516–537. https://doi.org/10.1016/j.jhep.2023.03.017

[2] Zhou, J., Sun, H., Wang, Z., Cong, W., Zeng, M., Zhou, W., et al. (2023). Guidelines for the Diagnosis and Treatment of Primary Liver Cancer (2022 Edition). *Liver Cancer*, *12*(5), 405–444. https://doi.org/10.1159/000530495

[3] Park, I., Kim, N., Lee, S., Park, K., Son, M., Cho, H., et al. (2022). Characterization of signature trends across the spectrum of non-alcoholic fatty liver disease using deep learning method. *Life Sciences*, *314*, 121195. https://doi.org/10.1016/j.lfs.2022.121195

[4] Asif, S., Wenhui, Y., Ur-Rehman, S., Ul-Ain, Q., Amjad, K., Yueyang, Y., et al. (2024). Advancements and Prospects of Machine Learning in Medical Diagnostics: Unveiling the future of Diagnostic Precision. *Archives of Computational Methods in Engineering*. https://doi.org/10.1007/s11831-024-10148-w

[5] Chai, J., Zeng, H., Li, A., & Ngai, E. W. (2021). Deep learning in computer vision: A critical review of emerging techniques and application scenarios. *Machine Learning with Applications*, *6*, 100134. https://doi.org/10.1016/j.mlwa.2021.100134

[6] Sun, Y., Sun, Z., & Chen, W. (2024). The evolution of object detection methods. *Engineering Applications of Artificial Intelligence*, *133*, 108458. https://doi.org/10.1016/j.engappai.2024.108458

[7] Yaseen, M. (2024). What is YOLOv8: An In-Depth Exploration of the Internal Features of the Next-Generation Object Detector. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2408.15857

[8] Safaldin, M., Zaghden, N., & Mejdoub, M. (2024). An improved YOLOV8 to detect moving objects. *IEEE Access*, *12*, 59782–59806. https://doi.org/10.1109/access.2024.3393835

[9] Yu, H., Yang, L. T., Zhang, Q., Armstrong, D., & Deen, M. J. (2021). Convolutional neural networks for medical image analysis: State-of-the-art, comparisons, improvement and perspectives. *Neurocomputing*, *444*, 92–110. https://doi.org/10.1016/j.neucom.2020.04.157

[10] Sarvamangala, D. R., & Kulkarni, R. V. (2021). Convolutional neural networks in medical image understanding: a survey. *Evolutionary Intelligence*, *15*(1), 1–22. https://doi.org/10.1007/s12065-020-00540-3

[11] Li, X., Li, M., Yan, P., Li, G., Jiang, Y., Luo, H., et al.. (2023). Deep Learning Attention Mechanism in Medical Image Analysis: Basics and Beyonds. *International Journal of Network Dynamics and Intelligence*, 93–116. https://doi.org/10.53941/ijndi0201006

[12] Stephe, S., Kumar, S., Thirumalraj, A., & Dzhyvak, V. (2024). *Transformer based attention guided network for segmentation and hybrid network for classification of liver tumor from CT scan images*. https://essuir.sumdu.edu.ua/handle/123456789/97135

[13] Abdulsahib, F. I., Al-Khateeb, B., Kóczy, L. T., & Nagy, S. (2025). Liver Cancer classification approach using Yolov8. In *Lecture notes in networks and systems* (pp. 14–21). https://doi.org/10.1007/978-3-031-73997-2_2

[14] Ma, J., Xia, S., Zhang, B., Luo, F., Guo, L., Yang, Y., et al. (2023b). The pharmacology and mechanisms of traditional Chinese medicine in promoting liver regeneration: A new therapeutic option. *Phytomedicine*, *116*, 154893. https://doi.org/10.1016/j.phymed.2023.154893

[15] Nowak, S., Mesropyan, N., Faron, A., Block, W., Reuter, M., Attenberger, U. I., et al. (2021). Detection of liver cirrhosis in standard T2-weighted

MRI using deep transfer learning. *European Radiology*, 31(11), 8807–8815. https://doi.org/10.1007/s00330-021-07858-1

[16] Xu, J. (2020). RETRACTED: Medical wireless sensor network coverage and clinical application of MRI liver disease diagnosis. *Microprocessors and Microsystems*, *81*, 103688. https://doi.org/10.1016/j.micpro.2020.103688

[17] Haas, M. E., Pirruccello, J. P., Friedman, S. N., Wang, M., Emdin, C. A., Ajmera, V. H., et al. (2021). Machine learning enables new insights into genetic contributions to liver fat accumulation. *Cell Genomics*, *1*(3), 100066. https://doi.org/10.1016/j.xgen.2021.100066

[18] Ahad, M. T., Li, Y., Song, B., & Bhuiyan, T. (2023). Comparison of CNN-based deep learning architectures for rice diseases classification. *Artificial Intelligence in Agriculture*, *9*, 22–35. https://doi.org/10.1016/j.aiia.2023.07.001

[19] Cheng, N., Chen, D., Lou, B., Fu, J., & Wang, H. (2021). A biosensing method for the direct serological detection of liver diseases by integrating a SERS-based sensor and a CNN classifier. *Biosensors and Bioelectronics*, *186*, 113246. https://doi.org/10.1016/j.bios.2021.113246

[20] Yue, J., Yang, H., Feng, H., Han, S., Zhou, C., Fu, Y., et al. (2023). Hyperspectral-to-image transform and CNN transfer learning enhancing soybean LCC estimation. *Computers and Electronics in Agriculture*, *211*, 108011. https://doi.org/10.1016/j.compag.2023.108011

[21] Chen, S., Duan, J., Wang, H., Wang, R., Li, J., Qi, M., et al. (2022). Automatic detection of stroke lesion from diffusion-weighted imaging via the improved YOLOv5. Computers in Biology and Medicine, 150, 106120. https://doi.org/10.1016/j.compbiomed.2022.106120

[22] Sapitri, A. I., Nurmaini, S., Rachmatullah, M. N., Tutuko, B., Darmawahyuni, A., Firdaus, F., et al. (2022). Deep learning-based real time detection for cardiac objects with fetal ultrasound video. *Informatics in Medicine Unlocked*, *36*, 101150. https://doi.org/10.1016/j.imu.2022.101150

[23] Rana, M., & Bhushan, M. (2022). Machine learning and deep learning approach for medical image analysis: diagnosis to detection. *Multimedia Tools and Applications*, *82*(17), 26731–26769. https://doi.org/10.1007/s11042-022-14305-w

[24] Yaseen, M. (2024b). What is YOLOv8: An In-Depth Exploration of the Internal Features of the Next-Generation Object Detector. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2408.15857

[25] Liu, Y., Shao, Z., & Hoffmann, N. (2021). Global Attention Mechanism: Retain information to enhance Channel-Spatial interactions. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2112.05561

[26] Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., et al. (2021). Swin Transformer: Hierarchical Vision Transformer using Shifted Windows. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2103.14030

[27] Woo, S., Park, J., Lee, J., & Kweon, I. S. (2018). CBAM: Convolutional Block Attention Module. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1807.06521

[28]N. Sriram, Jayalakshmi V., P. Preethi, B. Shoba, & K. Shenbagavalli. (2024). Navigating the Future with YOLOv9 for Advanced Traffic Sign Recognition in Autonomous Vehicles. *International Journal of Computational and Experimental Science and Engineering*, 10(4). https://doi.org/10.22399/ijcesen.658