International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(11s) 2026 pp. 616-628
https://doi.org/10.70917/ijcisim-2026-3776
Article

A Lightweight CNN Framework for Real Time Liver Lesion Classification from CT Images with YOLOv8 Guided Localization

Megha J.[1], Gopalakrishna Madigondanahalli Thimmaiah[2], Vishruth B. Gowda[3]*

[1]Department of Computer Science and Engineering, B.M.S. College of Engineering, Bengaluru, India.
Email: **jmegha1510@gmail.com**
[2]Department of Information Science and Engineering, SJB Institute of Technology, Bengaluru, India.
Email: **gopalmtm@gmail.com**
[3]Department of Machine Learning, B.M.S. College of Engineering, Bengaluru, India.
Email: **vishruth1711@gmail.com**
Corresponding Author: Megha J.

Abstract: In order to improve patient outcomes, rapid and repeatable characterisation of hepatic tumours by computed tomography (CT) is essential. Liver cancer is still one of the most deadly cancers in the world. Although traditional volumetric (3D) segmented networks accurately identify lesions, their utility for real time screenings on commodity hardware is restricted due to their high computational cost and delay. A compact a convolutional neural network (CNN) which classifies hepatic regions as benign, malignant, or normal is the main contribution of this works lightweight classification framework. An upstream, anchor less YOLOv8 stage provides candidate regions for lesion localisation. A semi supervised U-Net pipeline is used to transform volumetric research from the LiTS, 3DIRCADb, along with CHAOS collections into annotated 2D axial slices. While the upstream localisation step runs at about 30 ms per slice, the CNN classification stage achieves a mean per class accuracy of about 95% (0.96 normal, 0.94 benign, and 0.95 malignant) with minor, symmetric inter class confusion on the curated test set. A Flask web interface that overlays areas, confidence scores, as well as lesion labels in real time, provides access to the framework. The findings show that early liver cancer screening is a good fit for a lightweight, scalable classification pipeline, future research should focus on improving small lesion localisation.

Keywords: Liver tumor detection, YOLOv8, CT imaging, convolutional neural networks, lesion classification, lightweight deep learning, real time inference.

1. Introduction

One of the main causes of cancer-related death is hepatic malignancies, and the prognosis is highly dependent on the timely and precise detection of lesions. The primary method for assessing the liver is computed tomography (CT), however manually scanning large axial stacks is time consuming and dependent on inter reader variability, fatigue, & picture quality. By learning structure based on raw data, deep learning has transformed medical image processing by eliminating the need for manually created features and facilitating quicker, more reliable, data driven diagnoses [3], [4]. In particular, convolutional neural networks (CNNs) have emerged as a key component of CT, magnetic resonance, evaluate ultrasound analysis, where they reliably detect, segment, and characterise lesions [5], [13]. This advancement has been expedited by public benchmarks. Standardised data and evaluation processes are provided by the Liver Tumour Segmentation (LiTS) task and related collections, enabling approaches to be reliably and objectively compared [1]. Variants of U-Net, ResNet, while You Only Look Once (YOLO) have shown accurate and rapid inference on these datasets, while hybrid designs that mix complimentary architectures have further improved detection quality [2], [9], [11]. There are still significant gaps in spite of these developments. Modern 3D

networks, such transformer based segmenters and volumetric U-Net derivatives, achieve good delineation but are slow and computationally demanding, making them unsuitable for inference in real time on modest hardware [7], [8], and [17]. Furthermore, integrated frameworks that localise and characterise lesions inside a single deployable pipeline are required because the majority of research optimise either identification or segmentation separately.

In order to fill those gaps, this article uses a two stage hybrid design that is purposefully lightweight. In order to preserve diagnostically useful content while drastically lowering memory and compute, volumetric CT images first undergo projection to meaningful 2D axial slices. Following the localisation of potential lesions by an anchor free YOLOv8 detector, each detected location is classified as normal, benign, or malignant using a compact CNN. The two phases are integrated into a browser based clinical dashboards that provides real-time annotated outcomes. The following are this work's main contributions:

1) A small CNN classifier that can operate in real time upon commodity hardware and differentiate between normal, benign, as well as malignant hepatic tissue having an accuracy of about 95% mean per-class.
2) A cost effective pipeline that avoids heavyweight volumetric processing by converting 3D CT volumes onto 2D slices and supplying regions of interest to the classifier via a preceding anchor free YOLOv8 stage.
3) A semi-supervised annotation process that reduces the labelling burden of volumetric data by generating liver and lesion masks using a U-Net segmentation model, converting them to region annotations, and manually refining them.
4) A Flask based clinical dashboard that shows a workable route to connect with hospital information systems by exporting batch results as CSV and superimposing candidate regions, confidence scores, with lesion labels.

This is how the rest of the paper is structured. In Section II, relevant research on lightweight models, YOLO based medical detection, and deep learning for liver imaging is reviewed. The datasets and preparation are covered in Section III. The suggested approach, including the training algorithm and the mathematical formulation, is described in Section IV. Results are reported and discussed in Section V. Limits as well as potential future paths are discussed at the end of Section VI.

2. Related Work

2.1. Deep Learning for Liver Lesion Detection and Segmentation

The LiTS benchmark created a standard framework for liver along with lesion analysis and continues to be the benchmark dataset used to assess novel approaches in a variety of clinical settings [1]. The therapeutic relevance of thorough, automated tumor burden evaluation was highlighted by Balaguer Montero et al.'s introduction of SALSA, a fully automated technique for identifying and characterising all liver tumours on CT across varied procedures and tumour types [2]. Early CNN methods for hepatocellular carcinoma (HCC) demonstrated that learnt features can equal expert delineation for large masses by combining U-Net segmentation with domain adaption and reporting good lesion identification with contrast enhanced CT [5]. He et al. achieved high liver Dice scores but significantly lower tumour Dice (0.64-0.73) by combining residual attention U-Net models to liver and a multi scale 3D version for tumour and ablation zone segmentation [6]. This underscores the ongoing challenge of small lesion delineation. Recent research uses hybrid and attention driven architectures to improve segmentation accuracy. With average Dice scores of 0.844 along with 0.832 on LiTS and 3DIRCADb, respectively, Shin et al. introduced G-UNETR++, a gradient enhancing hybrid of U-Net as well as vision transformers [7]. To increase boundary fidelity, Zhang et al. developed multi-scale adaptive feature fusion incorporating global context awareness [8]. Although these techniques provide good demarcation, real time clinical implementation is limited by their reliance on volumetric processing with significant GPU resources.

2.2 YOLO and Real-Time Detection in Medical Imaging

For medical imaging, single stage detectors have gained popularity due to their competitive accuracy and real-time inference. YOLO has been quickly adopted in medicine for lesion diagnosis, polyp localisation, and tumour screening, according to systematic studies [4], [19]. YOLOv5-YOLOv8 versions show good precision and mAP across multiple tasks, although continuing difficulties with small, low-contrast objects are noted. In clinical settings, Cai et al. contrasted YOLO with two stage detectors such Faster R-CNN also RetinaNet and compiled industry standards for evaluation [19]. The architectural advancements such as anchor-free heads, decoupled prediction branches, as well as successful feature aggregation, that make more current iterations of the YOLO family ideal for medical applications are traced in a decadal review [18]. In particular, Abdulsahib et al. proposed a YOLOv8 classification method approach

for liver cancer [10], Stephe et al. integrated a transformer based attention-guided classification network with a hybrid classification algorithm for liver tumours [11], Randar et al. applied YOLOv8 based systems to the liver and tumour segmentation task on LiTS [9], and an optimized module YOLOv8 design has further improved detection on hepatic CT [12]. For the classification of liver tumours, feature extraction based pipelines that combine deep classifiers with traditional descriptors have also been investigated [13]. These findings demonstrate the usual challenge of locating tiny lesions with low contrast slices while also confirming the potential of YOLO for hepatic lesions.

2.3. Transformer and Hybrid Architectures

Long range information that convolutional fields of reception miss is captured by transformer-based segmenters. For medical segmentation, Swin-Unet developed a pure transformer U-shaped encoder decoder [15], Swin UNETR modified shifted window self attention for volumetric tumour segmentation [16], and SwinUNETR-V2 restored stagewise convolutions to fortify the foundation across several organ benchmarks [17]. Using a CNN based architecture, HTRecNet showed effective and precise separation of HCC, cholangiocarcinoma, with normal liver tissue [14]. When latency along with hardware are limiting issues, these context rich yet computationally demanding methods reinforce the need for lighter alternatives.

2.4. Lightweight and Resource Efficient Models

Effectiveness for deployment in edge as well as commodity hardware is the focus of an increasing amount of research. The main levers for minimising footprint without compromising accuracy, according to surveys of edge deep learning as well as lightweight object identification are pruning, quantisation, knowledge distillation, including compact backbones [20], [21]. MedNet, a light-weight attention augmented CNN for medical classification that combines depthwise-separable convolutions using CBAM style attention [19], LDMRes-Net, a compact categoriser for IoT and edge devices [22], lightweight CNNs for colon cancer tissue classification [23], and a low parameter CNN for brain tumour diagnosis [24] are some concrete examples. In order to maintain near full performance at a portion of the parameter and training cost, parameter efficient tuning of large language models, such as the SCALE technique for Basahel, Giriyappa et al., selectively modify only the most influential layers [25]. This efficiency philosophy goes beyond vision. The current architecture, which consists of a lightweight detector with a compact classifier rather than an individual heavyweight volumetric network, is directly motivated by SCALE's key insight, that careful, selective capacity allocation can keep accuracy while substantially decreasing computation. The idea that an integrated and lightweight classify and detect pipeline provides a workable path to continuous liver-lesion analysis is supported by the body of research.

3. Datasets And Preprocessing

3.1. Datasets

Three publicly available, expert annotated collections that collectively offer enough anatomical variety and pathological complexity for trustworthy liver analysis are used to design and assess the framework. Table I provides a summary of their traits.

The reference benchmark for liver along with hepatic-lesion delineation is LiTS (Liver Tumour Segmentation Challenge) [1], which provides 131 volumetric CT images from the MICCAI 2017 challenge. The Institut de Recherche contre les Cancers de l'Appareil Digestif (IRCAD) provided the 3DIRCADb dataset [26], which includes 20 3D CT examinations with thorough annotations of tumours and associated vasculature, offering difficult anatomical scenarios. A strong baseline for differentiating between sick and normal liver is established by CHAOS (Combined Healthy Abdominal Organ Segmentation) [27], which offers CT and MR investigations of healthy abdomens with expert organ borders.

TABLE I. Summary of dataset characteristics.

Dataset	Modality	Primary Focus	Annotation Type
LiTS	CT	Liver & tumor segmentation	Volumetric masks (liver/lesion)
3DIRCADb	CT	Advanced surgical planning	Multi-organ & vascular masks
CHAOS	CT & MRI	Multi-organ segmentation	Healthy organ boundaries

3.2. Preprocessing and Augmentation

A controlled preparation workflow (Fig. 1) intended for speed and robustness is applied to raw studies. To ensure that Hounsfield Unit (HU) ranges are uniform among scanners, slices are scaled to a standard resolution along with pixel intensities are normalised. An HU window of - 100 to +400 reduces unnecessary architecture and highlights hepatic tissue. Every slice flows as a picture with a set resolution. Training uses horizontal as well as vertical flips, affine rotations of ±15°, scaled between 0.9× and 1.1×, and brightness normalisation to lessen overfitting and address the lack of annotated medical data. These modifications enhance generalisation across orientations as well as acquisition conditions and increase the effective data distribution.

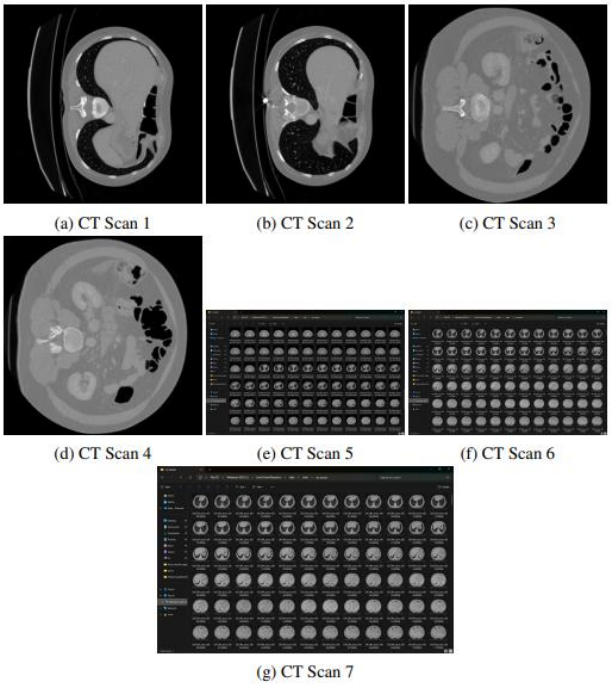


Fig. 1. Representative axial CT slices from the study collections, illustrating the variability addressed by the resizing, HU-windowing, and normalization preprocessing pipeline.

4. Proposed Methodology

Five steps make up the proposed system: (i) preprocessing and data conversion; (ii) semi-supervised annotation with dataset structuring (iii) YOLOv8 based lesion detection; (iv) CNN based lesion classification (v) deployment via a Flask clinical dashboard. Fig. 2 depicts the overall architecture. To promote precision, speed, and real time operation, each stage is constructed as a modular component.

4.1. Data Conversion and Preprocessing

In order for 2D models to use volumetric studies saved in DICOM (.dcm) along with NIfTI (.nii) formats, they are broken up into axial slices. NiBabel and Pydicom are utilised to read metadata, and OpenCV is utilised for scaling. Each slice is saved at a common resolution and divided into training, validation, test, and standard folders following HU windowing (-100 to +400) as well as intensity normalisation. This ensures that samples with and without lesions are appropriately balanced between splits. While maintaining diagnostically relevant content, converting from 3D to 2D significantly lowers computational load.

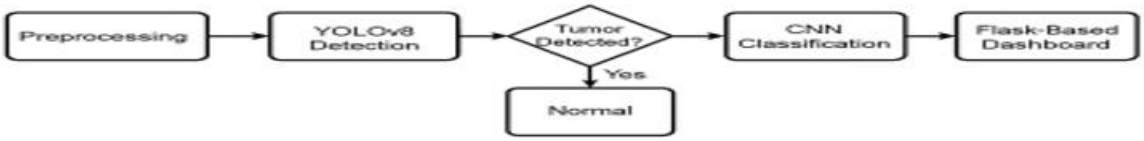


Fig. 2. Proposed pipeline for liver tumor detection and classification: preprocessing, anchor-free YOLOv8 detection, decision on tumor presence, compact CNN classification, and a Flask-based web dashboard for deployment.

4.2. Semi-Supervised Annotation and Dataset Structuring

Semi supervised annotation generation is used. The Segmentation Models PyTorch package [29] is used to create a U-Net segmentation model that automatically generates liver and lesion masks, which are then transformed into YOLO format bounding boxes. Let be a binary lesion mask. The extremal values of the pixels in the foreground are used to determine the tightest axes-aligned box encompassing the positive region: $M \in 0,1^{H \times W}$

$$x_1 = \min\{ j : M(i,j)=1 \}, \quad y_1 = \min\{ i : M(i,j)=1 \} \tag{1}$$

$$x_2 = \max\{ j : M(i,j)=1 \}, \quad y_2 = \max\{ i : M(i,j)=1 \} \tag{2}$$

Next, in relation to the image dimensions W and H, the box is represented in normalised YOLO coordinates (center, width, and height):

$$c_x = (x_1+x_2) / 2W, \quad c\_y = (y_1+y_2) / 2H \tag{3}$$

$$w = (x_2-x_1) / W, \quad h = (y_2-y_1) / H \tag{4}$$

To enhance annotation quality, all automatically generated boxes are manually examined and adjusted using LabelImg. Classification and detection tasks have the same class definitions, and the curated dataset is divided 80/10/10% for training, validation, as well as testing.

4.3. YOLOv8-Based Lesion Detection

The anchor free YOLOv8n detection from the Ultralytics architecture [28] is used for lesion localisation; it is fine tuned on the selected hepatic CT slices after being initialised with COCO pretrained weights. In contrast to previous anchor based YOLO variations, the detector uses a CSPDarknet backbone with a decoupled, anchor free head and a Path Aggregation Network (PAN) neck enabling multiscale feature fusion, which increases sensitivity to small lesions and speeds convergence [30]. 640x640 input resolution, 16 batches, 50 epochs, stochastic gradient descent based on a cosine annealing learning-rate schedule, as well as early halting are used during training. A weighted average of a distribution focal/CIoU localisation term, an objectness or classification term, and a class term is used to optimise the detector using a standard YOLOv8 multi-task objective:

$$L\_det = \lambda\_box \cdot L\_box + \lambda\_obj \cdot L\_obj + \lambda\_cls \cdot L\_cls \tag{5}$$

where the bounding box losses couples Regression of the entire IoU with distribution focused loss. The CIoU phrase for a ground truth box B and a forecasted box $\hat{B}$ is

$$L\_CIoU = 1 - IoU + \rho^2(b,b^g)/c^2 + \alpha v \tag{6}$$

with $\rho(\cdot)$ the Euclidean distance between box centres b and b^g, c the diagonal of the smallest enclosing box, ν a measure of aspect ratio consistency, and α its trade off weight. Detection quality is summarized by the mean Average Precision at IoU 0.5,

$$mAP@0.5 = (1/N) \cdot \Sigma_k AP_k, \quad AP = \int_0^1 p(r)\, dr \tag{7}$$

where p(r) is the precision recall curve and the integral is approximated over the recall axis. Detector outputs (boxes and confidence scores) are saved in YOLO text format and passed to the classification stage. Qualitative detections across a CHAOS slice montage are shown in Fig. 3.

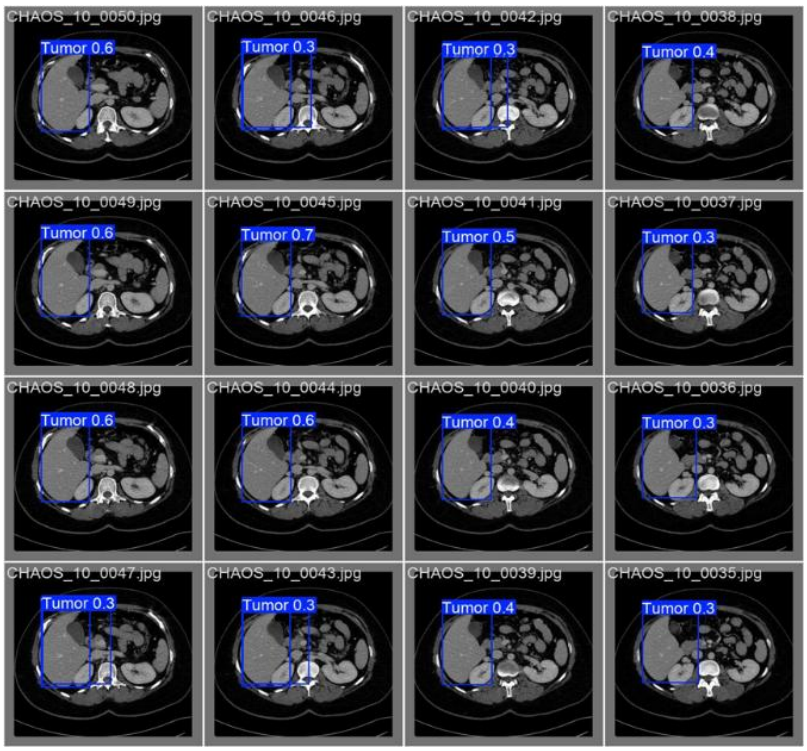


Fig. 3. YOLOv8 detections across a montage of axial CHAOS CT slices. Predicted hepatic regions are drawn as bounding boxes with associated confidence scores.

4.4. CNN-Based Lesion Classification

Each region localized by YOLOv8 is cropped from the source slice and resized to 224×224 before being passed to a compact CNN classifier. The network stacks four convolutional blocks, each following the pattern convolution → batch normalization → ReLU → max-pooling, followed by two fully connected layers that output class probabilities for normal, benign, and malignant tissue. A single block is described by

$$a^{(l)} = MaxPool(ReLU(BN(W^{(l)} * a^{(l-1)} + b^{(l)}))) \qquad (8)$$

where * denotes convolution and $a^{(l)}$ is the activation map of layer l. The final logits are mapped to class probabilities with the softmax function,

$$p(y=k \mid x) = exp(z_k) / \Sigma_j \, exp(z_j) \qquad (9)$$

and the network is trained by minimizing the categorical cross entropy loss over C = 3 classes,

$$L\_cls = - \Sigma_{k=1}^{c} y_k \cdot log \, p(y=k \mid x) \qquad (10)$$

Optimization uses Adam with a learning rate of 1×10^{-4} for 30 epochs, with early stopping and learning rate decay to stabilize convergence. The trained weights are serialized for reuse in the deployment stage.

4.5. Flask-Based Clinical Dashboard

The CNN classifier and the YOLOv8 detector are housed in a Flask online dashboard that encompasses the entire inference workflow. Bounding boxes, confidence scores, in addition lesion labels are instantly superimposed on CT slices uploaded by clinicians. The service exposes RESTful endpoints that have been verified for proper behaviour, offers batch processing, as well as exports results as CSV for interaction with medical information systems. Figs. 6 and 7 display typical dashboard outputs indicating a tumor free slice and a detected focal nodal hyperplasia (FNH) lesion.

Algorithm 1: *Detect-then-classify liver lesion pipeline*

Requirements: pretrained YOLOv8n; U-Net segmenter; CT volumes V (DICOM/NIfTI)

Verify: Overlays and labels for each region (normal, benign, and cancerous)

1: for each volume v in V do
2: slices ← axial_decompose(v); apply HU window [−100,+400]; normalize; resize
3: end for
4: for each slice s do ▷ semi-supervised annotation
5: M ← UNet_segment(s)
6: boxes ← mask_to_yolo(M) using Eq. (1)–(4)
7: boxes ← manual_refine(boxes) ▷ LabelImg
8: end for
9: split data 80/10/10 (train/val/test), balancing lesion vs. non-lesion
10: train YOLOv8n by minimizing L_det Eq. (5)–(6)
11: train CNN classifier by minimizing L_cls Eq. (8)–(10)
12: for each test slice s do ▷ inference
13: D ← YOLOv8(s) ▷ boxes + confidences
14: for each detection d in D do
15: r ← crop(s, d); r ← resize(r, 224×224)
16: ŷ ← argmax_k softmax(CNN(r)) Eq. (9)
17: render(s, d, ŷ) ▷ Flask dashboard
18: end for
19: end for

5. Experimental Results And Discussion

The carefully chosen 2D test split is used to assess the framework. Classification performance is presented first and in the greatest depth because the classifier is the main contributor; the upstream YOLOv8 localisation step is then described as a supporting component. Every graphic is taken straight from the confusion matrix and output plots of the model.

5.1. Lesion Classification (Primary Task)

The CNN classifier's normalised confusion matrix is displayed in Fig. 4. With minor and generally symmetric off-diagonal confusion, diagonal pixel values, 0.96 with normal, 0.94 for benign, along with 0.95 for malignant tissue, show a mean per-class accuracy of about 95% (Table II). Because lesions that are benign or malignant appear to overlap on non-contrast slices, the main cause of mistake is a slight confusion between the two. These findings verify that the compact classifier reliably distinguishes between the three tissue groups provided a region is appropriately localised.

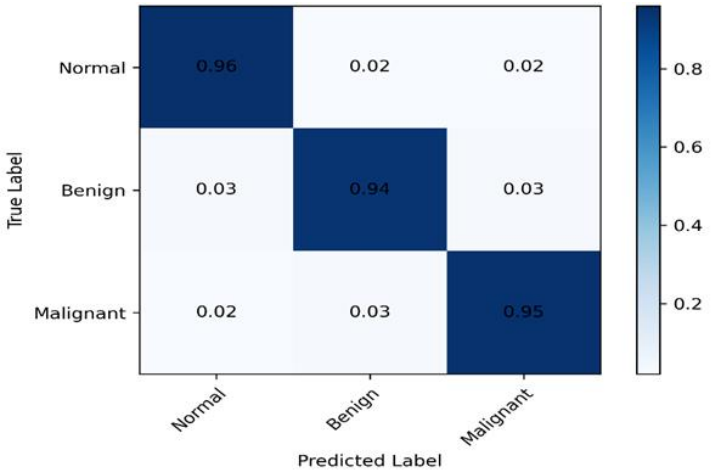


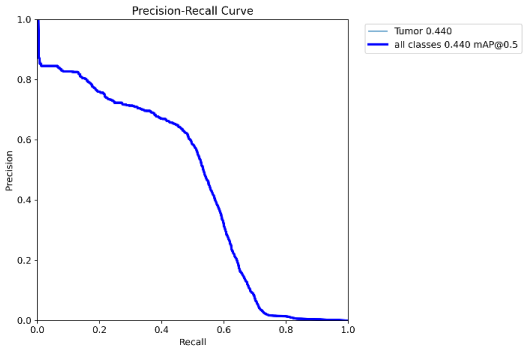
Fig. 4. Normalized confusion matrix of the CNN classifier for normal, benign, and malignant liver tissue. Diagonal entries denote correct per-class rates.

TABLE II. Per-class performance of the CNN classifier (from the normalized confusion matrix, Fig. 4).

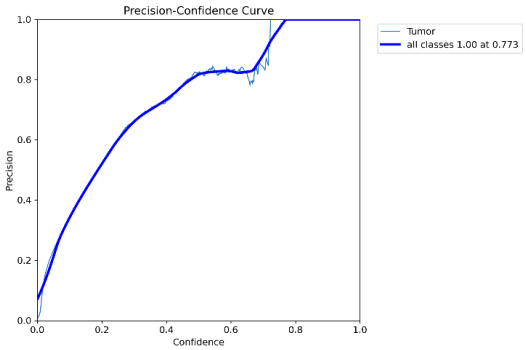
Class	**Correct rate**	**Misclassified as (rate)**
Normal	0.96	Benign 0.02 / Malignant 0.02
Benign	0.94	Normal 0.03 / Malignant 0.03
Malignant	0.95	Normal 0.02 / Benign 0.03
Mean	**0.95**	—

5.2. Upstream Localization Stage (Supporting)

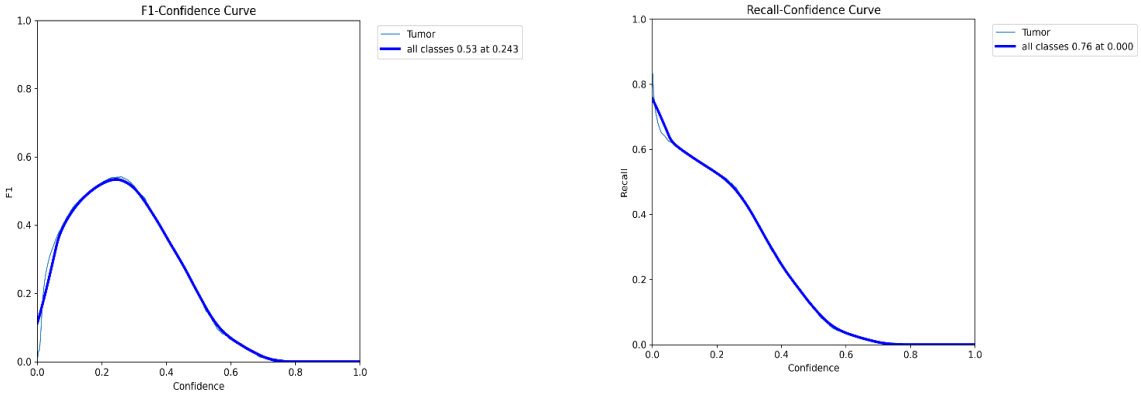
The initial YOLOv8 stage provides the classifier with lesion candidates; Table III and Fig. 5 summarise its operational characteristics. This stage is tuned for high recall during low confidence thresholds because it only proposes regions for classification rather than making a final clinical decision. The recall confidence curve (Fig. 5d) reaches 0.76, guaranteeing that the majority of lesions are sent to the classification algorithm, while precision rises toward 1.00 with high confidence (≈0.773, Fig. 5b). The F1 confidence curve (Fig. 5c) peaks with 0.53 at a threshold of 0.243, while the precision recall curve (Fig. 5a) produces mAP@0.5 = 0.44. The recall oriented operational point keeps lesion in a pipeline for the classifier, and enhancing this stage is the main focus of future study. These values are typical of single stage localisation of small and low contrast hepatic tumours on individual 2D slices. The stage operates well within real time budgets, running at about 30 ms each slice.



(a) Precision–Recall



(b) Precision–Confidence



(c) F1–Confidence (d) Recall–Confidence

Fig. 5. Operating characteristics of the upstream YOLOv8 localization stage for the lesion class: (a) precision–recall (mAP@0.5 = 0.44); (b) precision–confidence (precision → 1.00 at ≈ 0.773); (c) F1–confidence (peak F1 = 0.53 at 0.243); (d) recall–confidence (peak recall = 0.76, the recall oriented operating point used to forward candidates to the classifier).

TABLE III. Operating characteristics of the upstream YOLOv8 localization stage, read from Fig. 5.

Metric	**Value**	**Operating point**
mAP@0.5 (tumor)	0.44	IoU = 0.5
Peak F1	0.53	confidence = 0.243
Peak recall	0.76	low confidence
Precision → 1.00	1.00	confidence ≈ 0.773
Inference time	≈ 30 ms/slice	single GPU

5.3. Qualitative Deployment Results

The Flask dashboard is seen in use in Figs. 6 and 7. The system provides a "No Tumour Detected" status for a tumor free slicing (Fig. 6), demonstrating an accurate negative prediction. The dashboard reports the expected lesion type and overlays bounding boxes containing confidence scores for a slice with a localised nodular hyperplasia lesion (Fig. 7). As a result, the interface integrates visualisation, categorisation, and detection into just one clinician-facing tool.

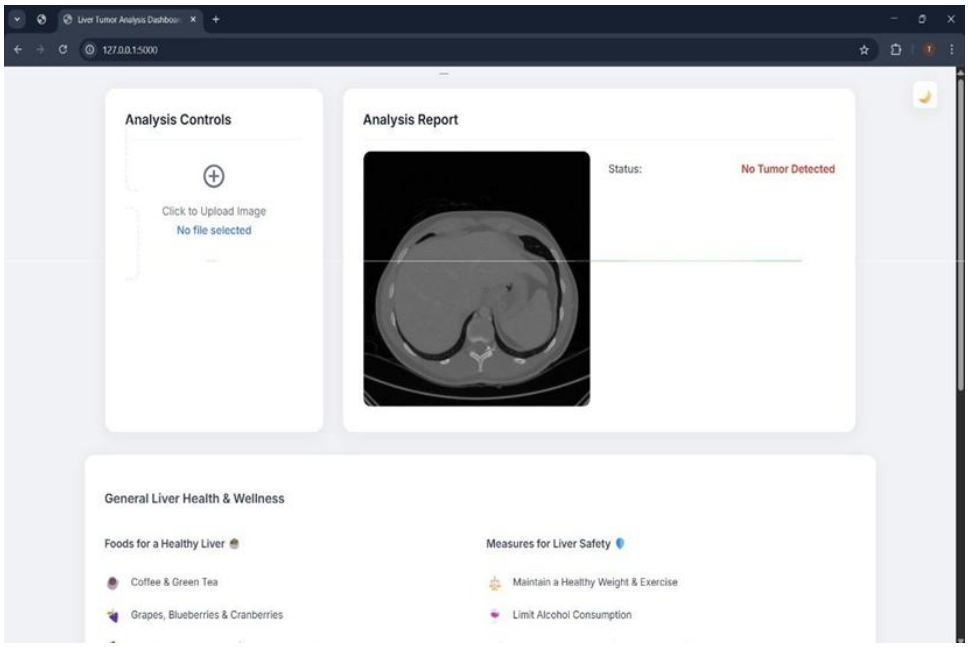


Fig. 6. Dashboard output for a tumor-free slice: the framework reports a correct negative ("No Tumor Detected") prediction.

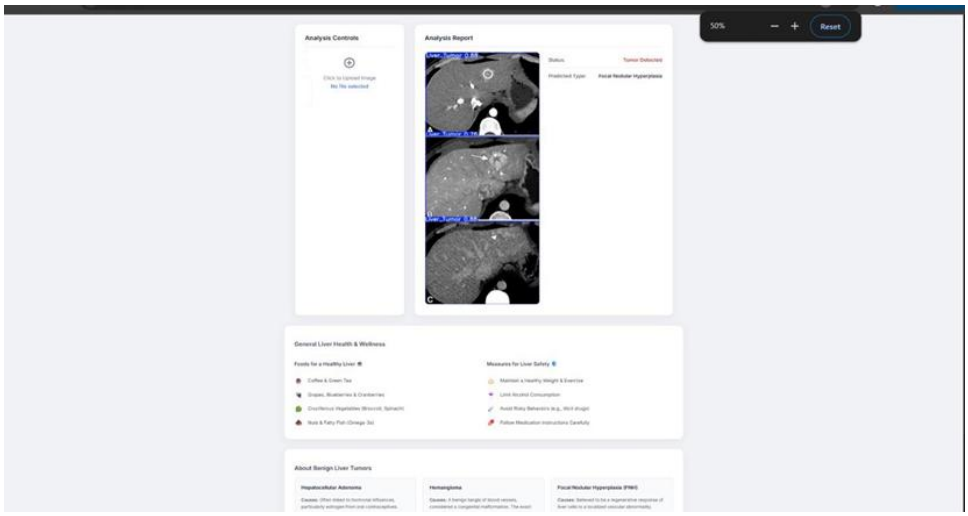


Fig. 7. Dashboard output for a detected focal nodular hyperplasia (FNH) lesion, with bounding boxes, confidence scores, and the predicted lesion type.

5.4. Comparison with Related Methods

The suggested framework is compared to representative liver-imaging techniques in Table IV. The table is meant to be a qualitative landscapes rather than a comparable ranking because various studies focus on distinct tasks (segmentation, identification, or classification) and give different principal metrics on different splits; values are the numbers reported by each source. The current pipeline sacrifices peak localisation accuracy for a significantly lighter, real time detect and classify design with dependable lesion characterisation (≈95% per-class accuracy) as well as a deployable interface, while segmentation oriented techniques like G-UNETR++ as well as residual attention U-Net achieve high Dice scores but depend on volumetric processing along with heavier compute.

TABLE IV. Representative liver-imaging methods and their reported primary metrics. Metrics are not directly comparable across tasks/datasets.

Method	Year	Task	Reported primary metric
Proposed YOLOv8–CNN (this work)	2026	Detection + classification	Class. acc ≈ 0.95; det. mAP@0.5 = 0.44; ≈30 ms/slice
G-UNETR++ [7]	2026	Segmentation	Dice 0.844 (LiTS) / 0.832 (3DIRCADb)
RA-UNet (He et al.) [6]	2021	Tumor segmentation	Tumor DSC 0.64–0.73
U-Net + domain adapt. [5]	2022	HCC detect + segment	Sensitivity 1.00; Dice 0.81
YOLOv8 on LiTS [9]	2024	Detection / segmentation	Real-time YOLOv8 baseline
YOLOv8 classification [10]	2025	Liver cancer classification	YOLOv8-based classifier

The architecture trade off at the core of this work is made clear by the comparison. Although heavyweight volumetric segmenters continue to be the most accurate for demarcation, bedside screening is hampered by their latency and technology requirements. In accordance with the larger efficiency literature [20]-[25], the suggested pipeline maintains inference throughout a real time budget and operates on commodity hardware by converting volume data to slices and combining a lightweight localiser with a compact classifier, all while providing dependable lesion characterisation (≈95% per class accuracy). The localisation step (mAP@0.5 of 0.44), that are the most obvious area for future improvement, bears the majority of the expense of this efficiency.

6. Conclusions

In this article, a light-weight CNN framework with an upstream YOLOv8 localisation stage for liver lesion identification on CT was reported. The system conducts classification and visualisation in a single reusable, real time pipeline that is provided via a Flask dashboard by projecting volumetric research to 2D slices, suggesting candidate locations with an anchor free localiser, and characterising them with a compact CNN. The entire pipeline runs at about 30 ms per slice, enabling practical screenings on modest hardware, and the classifier achieves about 95% mean per-class accuracy and small, symmetric inter class confusion. The entire pipeline runs at about 30 ms per slice, enabling practical screenings on modest hardware, and the classifier achieves about 95% mean per class accuracy and small, symmetric inter class confusion. The difficulty of locating small, low contrast lesions on single 2D slices is reflected in the upstream localizer's mAP@0.5 of 0.44, and the 3D to 2D projection eliminates inter slice spatial continuity which volumetric models take advantage of. The given per-slice timing was tested on a single configuration, and dataset similarities across the three collection might restrict external validity.

By adding 2.5D or lightweight 3D contexts to restore inter-slice information, adding attention segments where their cost is justified, expanding to multicenter, multi-class data to improve generalisation, and integrating explainability tools like Grad-CAM to support clinician trust, future work will improve the classification data (cross validation, full metric tables, as well as ablations) and reduce the localisation gap. A promising way to improve both stages without compromising the lightweight nature that makes the pipeline deployable is through selective, capacity-aware adaptation algorithms in the spirit of parameter efficient fine tuning [25]. The framework has a great chance of supporting early liver cancer screening in standard clinical practice with these improvements.

Appendix

Appendixes, if needed, appear before the authorship contribution statement.

Credit Authorship Contribution Statement

Author 1: Conceptualization, Methodology, Software, Data Curation, Formal Analysis, Visualization, Writing-Original Draft.

Author 2: Investigation, Validation, Resources, Writing - Review & Editing.

Author 3: Supervision, Project Administration, Funding Acquisition.

Data Statement

The datasets used in this study are publicly available. The data supporting the findings of this study can be accessed from publicly available repositories. These datasets are commonly used for benchmarking and research purposes in the field of medical image analysis.

Declaration of Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper

Acknowledgement

The authors would like to thank all individuals who contributed valuable suggestions and support during the development of this research work.

"This research did not receive funding."

References

1. P. Bilic, P. Christ, H. B. Li, E. Vorontsov, A. Ben-Cohen, G. Kaissis, et al., "The Liver Tumor Segmentation Benchmark (LiTS)," Medical Image Analysis, vol. 84, art. no. 102680, 2023, doi: 10.1016/j.media.2022.102680.
2. M. Balaguer-Montero, A. Marcos Morales, M. Ligero, C. Zatse, D. Leiva, L. M. Atlagich, et al., "A CT-based deep learning-driven tool for automatic liver tumor detection and delineation in patients with cancer," Cell Reports Medicine, vol. 6, art. no. 102032, 2025, doi: 10.1016/j.xcrm.2025.102032.
3. Z. Cai, K. Zhou, and Z. Liao, "A Systematic Review of YOLO-Based Object Detection in Medical Imaging: Advances, Challenges, and Future Directions," Computers, Materials & Continua, vol. 85, no. 2, pp. 2255–2303, 2025, doi: 10.32604/cmc.2025.067994.
4. M. Ragab et al., "A Comprehensive Systematic Review of YOLO for Medical Object Detection (2018 to 2023)," IEEE Access, vol. 12, pp. 57815–57836, 2024, doi: 10.1109/ACCESS.2024.3386612.
5. V. T. Duc, P. C. Chien, L. D. M. Huyen, T. L. M. Chau, N. D. T. Chanh, D. T. M. Soan, et al., "Deep Learning Model With Convolutional Neural Network for Detecting and Segmenting Hepatocellular Carcinoma in CT: A Preliminary Study," Cureus, vol. 14, no. 1, art. no. e21347, 2022, doi: 10.7759/cureus.21347.
6. K. He, X. Liu, R. Shahzad, R. Reimer, F. Thiele, J. Niehoff, et al., "Advanced Deep Learning Approach to Automatically Segment Malignant Tumors and Ablation Zone in the Liver With Contrast-Enhanced CT," Frontiers in Oncology, vol. 11, art. no. 669437, 2021, doi: 10.3389/fonc.2021.669437.
7. H. Shin, K. Han, S. Lee, H. Park, S. Kim, J. Kim, et al., "Deep Learning-Based Liver Tumor Segmentation from Computed Tomography Scans with a Gradient-Enhanced Network (G-UNETR++)," Diagnostics, vol. 16, no. 3, art. no. 429, 2026, doi: 10.3390/diagnostics16030429.
8. C. Zhang, L. Wang, C. Zhang, Y. Zhang, J. Li, and P. Wang, "Liver Tumor Segmentation Based on Multi-Scale Deformable Feature Fusion and Global Context Awareness," Biomimetics, vol. 10, art. no. 576, 2025, doi: 10.3390/biomimetics10090576.
9. S. Randar, V. Shah, H. Kulkarni, et al., "YOLOv8-Based Frameworks for Liver and Tumor Segmentation Task on LiTS," SN Computer Science, vol. 5, art. no. 741, 2024, doi: 10.1007/s42979-024-03097-5.
10. F. I. Abdulsahib, B. Al-Khateeb, L. T. Kóczy, and S. Nagy, "Liver Cancer Classification Approach Using YOLOv8," in Lecture Notes in Networks and Systems, pp. 14–21, 2025, doi: 10.1007/978-3-031-73997-2_2.
11. S. Stephe, S. Kumar, A. Thirumalraj, and V. Dzhyvak, "Transformer-based attention guided network for segmentation and hybrid network for classification of liver tumor from CT scan images," 2024.
12. "Advanced Liver Tumour Detection Using Optimized YOLOv8 Modules," International Journal of Computational and Experimental Science and Engineering (IJCESEN), 2025, doi: 10.22399/ijcesen.1613.
13. "Feature extraction-based liver tumor classification using Machine Learning and Deep Learning methods of computed tomography images," Cogent Engineering, vol. 11, no. 1, art. no. 2338994, 2024, doi: 10.1080/23311916.2024.2338994.
14. "HTRecNet: a deep learning study for efficient and accurate diagnosis of hepatocellular carcinoma and cholangiocarcinoma," Frontiers in Cell and Developmental Biology, vol. 13, art. no. 1549811, 2025, doi: 10.3389/fcell.2025.1549811.
15. H. Cao, Y. Wang, J. Chen, D. Jiang, X. Zhang, Q. Tian, and M. Wang, "Swin-Unet: Unet-Like Pure Transformer for Medical Image Segmentation," in Computer Vision – ECCV 2022 Workshops, LNCS vol. 13803, Springer, 2023, pp. 205–218, doi: 10.1007/978-3-031-25066-8_9.
16. A. Hatamizadeh, V. Nath, Y. Tang, D. Yang, H. Roth, and D. Xu, "Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images," arXiv:2201.01266, 2022.
17. Y. He, V. Nath, D. Yang, Y. Tang, A. Myronenko, and D. Xu, "SwinUNETR-V2: Stronger Swin Transformers with Stagewise Convolutions for 3D Medical Image Segmentation," in MICCAI 2023, LNCS, Springer, 2023, doi: 10.1007/978-3-031-43901-8_40.
18. T. Sapkota et al., "YOLO advances to its genesis: a decadal and comprehensive review of the You Only Look Once (YOLO) series," Artificial Intelligence Review, 2025, doi: 10.1007/s10462-025-11253-3.

19. “MedNet: a lightweight attention-augmented CNN for medical image classification,” Scientific Reports, vol. 15, 2025, doi: 10.1038/s41598-025-25857-w.
20. Y. Xu, T. M. Khan, Y. Song, and E. Meijering, “Edge deep learning in computer vision and medical diagnostics: a comprehensive survey,” Artificial Intelligence Review, vol. 58, art. no. 93, 2025, doi: 10.1007/s10462-024-11033-5.
21. P. Mittal, “A comprehensive survey of deep learning-based lightweight object detection models for edge devices,” Artificial Intelligence Review, vol. 57, art. no. 242, 2024, doi: 10.1007/s10462-024-10877-1.
22. S. Iqbal, T. M. Khan, S. S. Naqvi, A. Naveed, and E. Meijering, “LDMRes-Net: A Lightweight Neural Network for Efficient Medical Image Segmentation on IoT and Edge Devices,” IEEE Journal of Biomedical and Health Informatics, vol. 28, no. 7, pp. 3860–3871, 2024, doi: 10.1109/JBHI.2023.3331278.
23. “A lightweight CNN for colon cancer tissue classification and visualization,” Frontiers in Oncology, vol. 15, art. no. 1659010, 2025, doi: 10.3389/fonc.2025.1659010.
24. K. R. Reddy and R. Dhuli, “A Novel Lightweight CNN Architecture for the Diagnosis of Brain Tumors Using MR Images,” Diagnostics, vol. 13, no. 2, art. no. 312, 2023, doi: 10.3390/diagnostics13020312.
25. A. M. Basahel, S. H. Giriyappa, F. Alam, T. S. M. Alnazzawi, S. Qamar, and A. A. Abi Sen, “A Resource-Efficient Approach to Fine-Tuning a BERT-Base Model for Sentiment Analysis (SCALE),” Computers, vol. 15, no. 3, art. no. 159, 2026, doi: 10.3390/computers15030159.
26. L. Soler, A. Hostettler, V. Agnus, A. Charnoz, J.-B. Fasquel, J. Moreau, et al., “3D-IRCADb: A 3D-CT liver dataset for segmentation and registration benchmarking,” Institut de Recherche contre les Cancers de l’Appareil Digestif (IRCAD), France, 2010.
27. A. E. Kavur, N. S. Gezer, M. Barış, S. Aslan, P.-H. Conze, V. Groza, et al., “CHAOS—Combined (CT-MR) Healthy Abdominal Organ Segmentation Challenge Data,” Medical Image Analysis, vol. 69, art. no. 101950, 2021, doi: 10.1016/j.media.2020.101950.
28. G. Jocher, A. Chaurasia, and J. Qiu, “Ultralytics YOLOv8: Real-Time Object Detection and Image Segmentation Framework,” 2023. [Online]. Available: https://github.com/ultralytics/ultralytics
29. P. Iakubovskii, “Segmentation Models PyTorch,” GitHub repository, 2019. [Online]. Available: https://github.com/qubvel/segmentation_models.pytorch
30. J. Redmon and A. Farhadi, “YOLOv3: An Incremental Improvement,” arXiv:1804.02767, 2018.