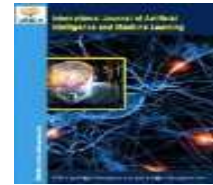




International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: <https://www.svedbergopen.com/>



Research Paper

Open Access

Automated Liver Lesion detection and classification using Deep learning and CT image

Trupti M. Kodinariya¹, Nikhil N. Gondaliya²

¹Research Scholar, Gujarat Technological University, Ahmedabad, Gujarat, India.

²Department of Information Technology, G H Patel College of Engineering and Technology, V. V. Nagar, Gujarat, India.

*Corresponding Author: nikhilgondaliya@gcet.ac.in

Abstract

Automated liver lesion analysis from computed tomography (CT) images can assist clinicians by providing objective information about lesion location and disease category. This paper presents a segmentation-based liver lesion classification framework developed using the Liver Tumor Segmentation Challenge 2017 (LiTS17) dataset. The proposed workflow follows a sequential strategy in which abdominal CT images are first preprocessed, the liver is segmented from the abdominal image, tumors are segmented from the extracted liver region, features are derived from the segmented tumor region, and the lesion is finally classified as benign or malignant. Several CT preprocessing methods are considered, including Hounsfield Unit (HU) windowing, resizing, Contrast Limited Adaptive Histogram Equalization (CLAHE), Min-Max normalization, Z-score normalization and median filtering. For the segmentation stage, multiple U-Net, LinkNet, PSPNet, FPN, DANet, CFNet, SpatialOCRNet and DeepLab configurations are investigated. The reported segmentation experiments achieve their strongest result with CFNet using a ResNet50V2 backbone, with 99.52% Dice and 99.05% IoU. For lesion classification, models are evaluated from scratch and with pretrained weights. Among the reported pretrained models, InceptionResNetV2 achieves the highest classification accuracy of 99%, with benign and malignant F1-scores of 98% and 99%, respectively. The study demonstrates the potential of combining explicit anatomical/lesion segmentation with deep feature learning for liver lesion classification.

Keywords: Liver lesion classification; LiTS17; CT image; liver segmentation; tumor segmentation; deep learning; benign tumor; malignant tumor; feature extraction.

1. Introduction

Liver cancer is a major cause of cancer death worldwide, and early detection and treatment planning rely strongly on imaging [1]. Contrast-enhanced CT is the most common modality for characterising hepatic lesions [2], but delineating the liver and its tumours by hand on hundreds of slices per patient is slow and varies between person. Automatic classification of liver lesions into clinically meaningful categories is challenging because lesion appearance can vary in size, shape, attenuation and surrounding anatomical context. A classification model trained directly on complete abdominal CT images may receive substantial irrelevant information from organs and tissues outside the lesion. Deep learning, and fully convolutional networks in particular, has become the standard way to automate this task [3], [4].

The Liver Tumor Segmentation Challenge (LiTS) provides the most widely used public benchmark, with expert-labelled liver and tumour masks for 130 training volumes [5]. The U-Net architecture introduced by Ronneberger et al. established an encoder-decoder framework with skip connections that has become a standard approach for biomedical image segmentation [6]. Subsequent architectures such as LinkNet, PSPNet and Feature Pyramid Networks introduced efficient encoding, multi-scale representations and contextual feature aggregation [11–13]. Attention-based models such as DANet further attempt to capture spatial and channel relationships that may improve segmentation of complex structures [14].

Two practical questions are often answered only implicitly: which preprocessing should be applied to the raw CT intensities, and which architecture and encoder should be chosen. Earlier studies showed that preprocessing can change U-Net liver segmentation quality noticeably [7], [8], but a combined view of preprocessing, window choice and a broad set of architectures on a single, consistently prepared LiTS17 dataset is rarely reported.

Segmentation quality matters beyond the mask itself: many liver tumour classification pipelines segment first and classify the segmented lesion [9], [10], so errors in the mask are carried into the diagnosis. This paper therefore also reports a benign versus malignant lesion classification study on the same data, using six models trained from scratch and ten ImageNet-pretrained models.

The contributions of this paper are:

- A reproducible slice-extraction and dataset-organisation pipeline for LiTS17 that links each CT slice to its mask and lesion type.
- An ablation of different preprocessing combinations with a U-Net, selected from a wider catalogue of techniques used in the medical literature.
- A benchmark of eight segmentation architectures and up to six pre-trained backbones for three-class liver and tumour segmentation, reported with Dice and IoU.
- An analysis of benchmark of six scratch-trained and ten pre-trained networks for benign versus malignant lesion classification on the same data.

Section 2 reviews related work, Section 3 describes the proposed work, Section 4 reports and discusses the experiments, and Section 5 concludes.

2. Related Work

2.1 Deep learning for liver and tumour segmentation

CNN-based methods learn hierarchical features directly from images and replaced hand-crafted features across medical image analysis [3]. For the liver, cascaded fully convolutional networks first segment the liver and then the tumours inside it [4]. The LiTS benchmark [5] standardised evaluation with multi-centre CT volumes acquired with varying scanners, slice thickness and protocols, and U-Net [6], an encoder-decoder with skip connections, became the common reference architecture.

2.2 Encoder-decoder and pyramid architectures

Besides U-Net, LinkNet connects encoder and decoder blocks directly, giving an efficient network [11]. The Feature Pyramid Network (FPN) fuses multi-scale features through a top-down pathway with lateral connections [12], and PSPNet adds a pyramid pooling module to aggregate context from several region sizes [13]. All four can use ImageNet-pretrained encoders, which improves convergence on limited medical data.

2.3 Context and attention based architectures

For larger context, the Dual Attention Network (DANet) models spatial and channel dependencies with two attention modules [14]. CFNet exploits co-occurrent features between classes [15], SpatialOCRNet (object-contextual representations) aggregates pixel features according to object-region representations [16], and DeepLab uses atrous (dilated) convolution with an encoder-decoder design [17]. These models are designed for natural-image scene parsing, and their behaviour on small, low-contrast liver tumours is less documented.

2.4 Backbones

The encoders used in this work cover residual networks (ResNet50 [18] and ResNet50V2 [19]), grouped and squeeze-and-excitation variants (SE-ResNeXt50 [20], [21]), the Inception family (InceptionV3 [22] and InceptionResNetV2 [23]), lightweight and compound-scaled networks (MobileNetV2 [24] and EfficientNetB3 [25]) and depthwise separable convolutions (Xception [26]).

2.5 Preprocessing of CT images

CT intensities are expressed in Hounsfield Units (HU); windowing selects the HU range that is mapped to grey levels, which increases contrast of the organ of interest. Phan et al. showed that exploiting HU attenuation information improves liver lesion analysis [29]. Contrast-limited adaptive histogram equalisation (CLAHE) [27] enhances local contrast, and denoising by median, Gaussian or block-matching 3D (BM3D) filtering [28] is also common. Islam et al. [7] and Kaczor et al. [8] compared preprocessing techniques for U-Net liver segmentation and found that the choice affects accuracy, which motivates the ablation in this paper.

2.6 Research gap

Existing studies typically evaluate one or two architectures with a single preprocessing choice. A consistent comparison of preprocessing techniques, HU windows and a broad range of architecture-backbone pairs, using the same data preparation and metrics on LiTS17, is the gap addressed here.

3. Proposed Work

The work has four steps: (1) data preparation and preprocessing of CT slices and masks, (2) selection of a preprocessing technique through a U-Net ablation, and (3) multi-class liver and tumour segmentation with traditional and advanced architectures, each trained with several pre-trained backbones and evaluated with Dice and IoU, and (4) benign versus malignant lesion classification with scratch-trained and pre-trained CNNs. The architecture flow of proposed work mentioned in Figure 1.

3.1 Data set

We employ in this work the Liver Tumor Segmentation Challenge 2017 (LiTS17) [19] dataset, a benchmark dataset provided for liver lesion segmentation from CT scans. The Dataset Consists of a total of 130 and 70 contrast-enhanced 3D abdominal CT scans from several clinical centers with different scanners and protocols for training and testing, respectively. Accordingly, the datasets exhibit highly varying spatial resolution and fields of view. The in-plane resolution ranges from 0.60 mm to 0.98 mm, and the slice spacing ranges from 0.45 mm to 5.0 mm. All the

scan axial slices are the same size, 512 x 512, but the number of slices varies significantly and ranges from 42 to 1026. The publicly available LiTS17 dataset includes only the ground truth (GT) of liver and lesion segmentation for the training split. A GT lesion type is not available. Khaled et al [30] published a paper where lesions are divided into three categories: normal, benign, and malignant (4 Normal, 41 Benign, and 85 Malignant) for the LiTS17 training dataset. After extracting the slices that contain liver and tumour information, the working dataset has 6,930 images with 2,602 benign and 4,766 malignant images.

3.2 Data preparation pipeline

1. CT volume processing. Each volume (Volume-1 to Volume-130) is read from its .nii file and rotated by 90°. A folder is created for every patient, the Hounsfield window is applied to each slice, and the slice is saved as a .jpg image (0.jpg, 1.jpg, ...).
2. Mask processing. Each segmentation file (Segmentation-1 to Segmentation-130) is read and rotated in the same way, and each mask slice is saved in a patient folder as a .png file, with labels 0 (background), 1 (liver) and 2 (tumour).
3. Slice selection and indexing. Slices that contain liver and tumour information are listed in a .csv file with the path of the volume slice, the path of the corresponding mask slice and the lesion type. The final dataset therefore consists of a Volume folder and a Mask folder per patient, and the CSV file.

From CT volumes to a benign or malignant label in five stages

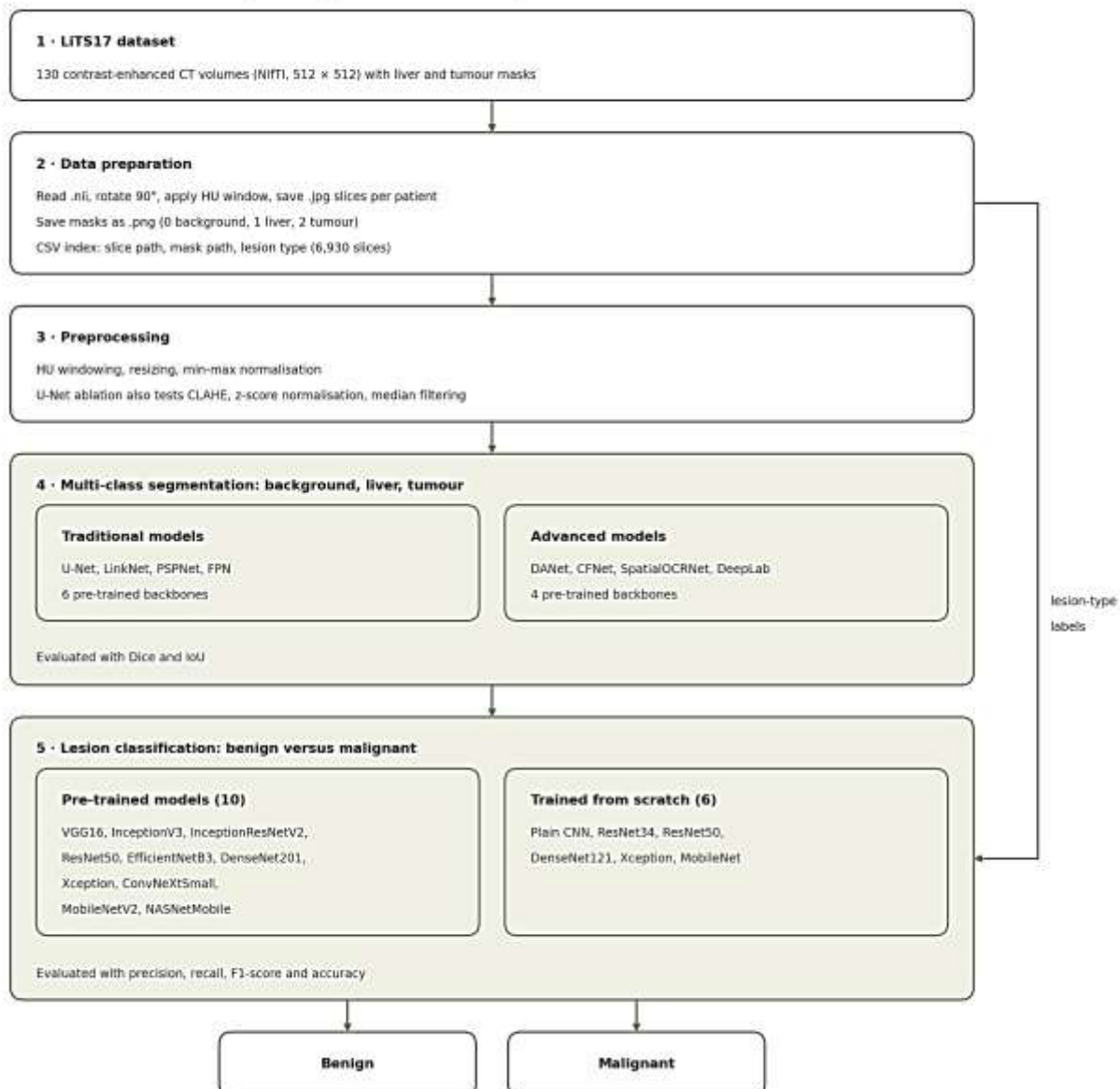


Fig. 1: Architecture flow of proposed work

3.3 Preprocessing techniques considered

The following techniques, which are applied to enhanced CT images in the medical literature, were considered:

- **HU Windowing:** HU values are mapped to grey levels within a window; values below the lower bound appear black and values above the upper bound appear white. The window width defines the range of CT numbers in the image. A large width assigns many HU values to each grey level, produces a long transition from white to black and obscures small differences in tissue attenuation. Different Window size is examined as shown in Figure 2 Among that window size [-30, 150] selected which emphasize the organ of interest, like the liver parenchyma and lesions. It serves to suppress irrelevant structures (e.g., bones or air), enhance contrast, and better model feature extraction.

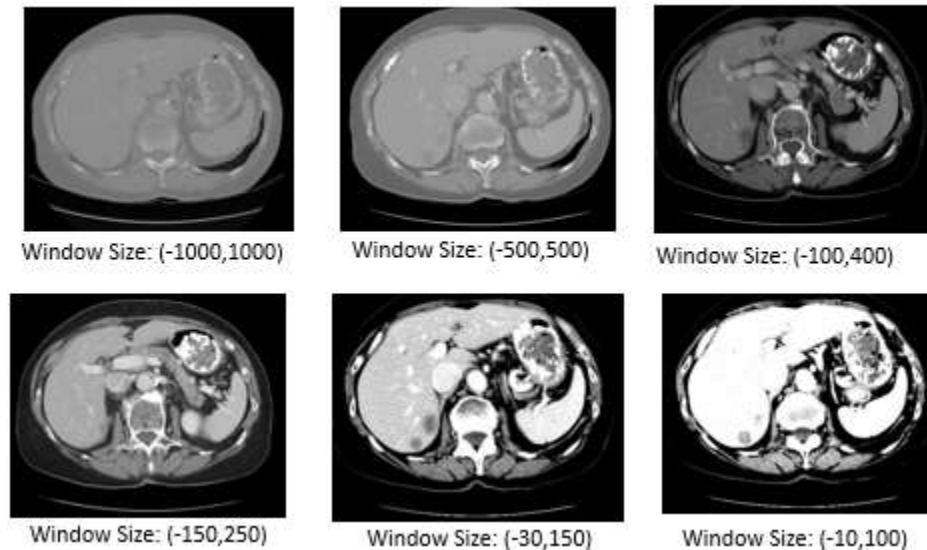


Fig. 2: Effect of Different HU window size selection on CT image

- **Image Resizing and Normalization:** Resizing is applied before segmentation to standardize the input dimensions. The conventional segmentation models use input size 224×224 , PSPNet uses 384×384 as input size, and advanced segmentation models use 320×320 as input size. The preprocessing study also considers Min-Max normalization and Z-score normalization.
- **Contrast Enhancement using CLAHE:** It is applied to improve the local contrast of CT images. When compared to global histogram equalization, CLAHE's 7×7 image tile diffusion effectively eliminates noise while revealing fine features. It comes in use for medical imaging, in particular when there is poor contrast between different types of tissue and a crucial feature (like a tumor) that needs to be examined. The approach additionally made use of a clip limit (0.2) to prevent the histogram equalization inside each tile from amplifying noise to an unreasonable degree. At last, the produced tiles are combined using bilinear interpolation.
- **Noise Removal:** Noise are removed using median filtering
- **Data Augmentation:** The dataset contains 6930 CT images, with 2,602 benign cases and 4,766 malignant instances, resulting in a class imbalance, which can be addressed through data augmentation. Classical image manipulation techniques were used to enhance data from the minority (benign) class in order to increase the model's generalizability and address this problem. The classical augmentation methods (Horizontal Flip, Vertical Flip, Rotations (90° , 180° , and 270°), and Random Shifting by 10-15 pixels in horizontal and vertical directions) were employed: These transformations were applied to synthetically increase the number of benign samples and ensure a more balanced distribution between classes, thereby reducing bias during model training
- **Hybrid combinations:** Different combinations were compared with a U-Net like HU windowing + min-max normalisation, HU windowing + CLAHE, HU windowing + z-score normalisation and HU windowing + median filtering and HU windowing + min_max normalization + CLAHE.

3.3 Liver and Tumor Segmentation

The segmentation stage is a multi-class problem with three mask classes: background, liver and tumor. Two groups of models were trained for multi-class segmentation of liver and tumour.

- **Traditional models** (U-Net [6], LinkNet [11], PSPNet [13], FPN [12]) with six pre-trained backbones: ResNet50, SE-ResNeXt50, InceptionV3, MobileNetV2, InceptionResNetV2 and EfficientNetB3, using the segmentation_models library [30].
- **Advanced models** (DANet [14], CFNet [15], SpatialOCRNet [16], DeepLab [17]) with four backbones: EfficientNetB3, Xception, ResNet50V2 and MobileNetV2, using the TensorFlow Advanced Segmentation Models library [31].

3.4 Lesion Classification

The classification stage addresses binary classification of liver lesions into benign and malignant categories.

Two groups of models were compared. Ten ImageNet-pretrained models were fine-tuned: VGG16 [35], InceptionV3 [22], InceptionResNetV2 [23], ResNet50 [18], EfficientNetB3 [25], DenseNet201 [36], Xception [26], ConvNeXtSmall [37], MobileNetV2 [24] and NASNetMobile [38]. Six models were trained from scratch: a plain CNN, ResNet34, ResNet50, DenseNet121, Xception and MobileNet [39]. All models used the Adam optimiser with a learning rate of 0.0001 and batch sizes of 8 or 16.

4. Experimental Result

This section demonstrates the outcome attained by the different segmentation and classification model. In addition, a comparative analysis of the respective model with existing mechanisms is illustrated.

4.1 Experimental setup

All of the tests were done in Google Colab Pro using Accelerated for NVIDIA GPU. The LiTS17 dataset, which is publicly available, was used as the source of the data. The total number of abdominal CT slices was then randomly divided into a training subset, validation subset, and test subset in a ratio of 70%, 10% and 20%, respectively. All models were trained with the Adam optimiser [32] and a learning rate of 0.0001, for 50 or 100 epochs and batch sizes of 8 or 16. The traditional models used 224×224 inputs, except PSPNet (384×384), and the advanced models used 320×320 inputs.

4.2 Effects of Pre-processing Techniques

To evaluate the effect of preprocessing on liver and tumor segmentation, a U-Net model was trained and tested with five different preprocessing pipelines, all built on Hounsfield Unit (HU) windowing. Segmentation performance was measured with the Dice similarity coefficient (DSC, reported in %) on the training, validation, and testing sets, as summarized in the table 1.

Table 1: Dice score (%) of U-Net for different preprocessing techniques

Preprocessing technique	Training	Validation	Testing
HU windowing + min-max normalization	99.01	98.19	97.04
HU windowing + CLAHE	98.58	98.01	97.03
HU windowing + z-score normalization	98.20	95.23	94.14
HU windowing + median filtering	97.57	92.78	89.92
HU windowing + min-max normalization + CLAHE	99.42	98.39	97.21

Among the single-technique pipelines, HU windowing combined with min-max normalization performed best, with Dice scores of 99.01%, 98.19%, and 97.04% on the training, validation, and testing sets respectively. HU windowing with CLAHE (Contrast Limited Adaptive Histogram Equalization) was nearly identical, reaching Dice scores of 98.58% (training), 98.01% (validation), and 97.03% (testing). This suggests that enhancing local contrast helps the network separate liver parenchyma from tumor regions, where intensity differences are subtle.

In contrast, HU windowing with z-score normalization generalized noticeably worse, with a training Dice of 98.20% but only 95.23% on validation and 94.14% on testing. HU windowing with median filtering performed worst, reaching a training Dice of 97.57% but dropping to 92.78% on validation and 89.92% on testing. This is likely because median filtering smooths the image and blurs fine boundary details and small tumor edges that are important for accurate segmentation. The widening gap between training and testing Dice for these two methods (about 4.06% for z-score and 7.65% for median filtering) also points to weaker generalization compared with the other pipelines.

The best overall performance came from combining HU windowing, min-max normalization, and CLAHE, which achieved Dice scores of 99.42% (training), 98.39% (validation), and 97.21% (testing). The small gap between training and testing Dice (about 2.21%) indicates stable learning with little overfitting. These results show that min-max normalization and CLAHE complement each other: normalization provides a consistent intensity range for stable training, while CLAHE improves local contrast and the visibility of tumor boundaries. This combined pipeline was therefore selected as the preferred preprocessing strategy for liver and tumor segmentation with U-Net.

4.3 Traditional Segmentation Models

Table 2 gives Dice and IoU for U-Net, LinkNet, PSPNet and FPN with the six backbones under the [-30, 150] HU window. All 24 combinations reached a Dice score of 99.76% or higher, so Dice alone hardly separates them. IoU is more discriminating. U-Net with SE-ResNeXt50 gave the highest IoU of the group (95.65%), followed by U-Net with ResNet50 (95.54%) and InceptionResNetV2 (95.53%). LinkNet was similar (up to 95.20% with EfficientNetB3), and PSPNet, which was run at the higher 384×384 resolution, was lower (92.28–93.84%). MobileNetV2 was the weakest backbone for every architecture, which is consistent with its small capacity, but the gap to the best backbone is small (about 0.7 IoU points for U-Net). FPN reached the highest Dice values (up to 99.87% with ResNet50), but its

IoU was only 62.04–62.51%. Such a large Dice–IoU gap suggests that the FPN predictions overlap well on the dominant liver region but poorly on the small tumour class, or that the two metrics were averaged differently for this model; this should be checked before using FPN for tumour delineation.

Table 2: Experimental Results using traditional Segmentation Model

BACKBONE/MODEL	UNET		LINKNET		PSPNET (384 x 384)		FPN	
	DICE	IOU	DICE	IOU	DICE	IOU	DICE	IOU
RESNET50	99.86	95.54	99.85	95.05	99.81	93.79	99.87	62.51
seresnext50	99.86	95.65	99.83	94.37	99.81	93.77	99.85	62.39
inceptionv3	99.86	95.36	99.84	95.01	99.80	93.34	99.85	62.20
mobilenetv2	99.84	94.97	99.82	94.22	99.76	92.28	99.82	62.04
inceptionresnetv2	99.86	95.53	99.84	95.16	99.80	93.72	99.86	62.37
efficientnetb3	99.86	95.44	99.85	95.20	99.81	93.84	99.84	62.29

4.4 Advanced Segmentation Models

Table 3 reports DANet, CFNet, SpatialOCRNet and DeepLab at 320×320 resolution under the $[-30, 150]$ HU window. Their Dice scores (about 99.4–99.5%) are slightly below those of the traditional models, but their IoU is clearly higher (about 98.8–99.0%), which indicates more precise overlap on both liver and tumour. DANet and CFNet were the most consistent: CFNet gave 99.47–99.49% Dice and 98.96–98.99% IoU with all four backbones, and DANet reached 99.49% Dice and 98.99% IoU with Xception. SpatialOCRNet depended strongly on the backbone: it was close to CFNet with ResNet50V2 (99.48% Dice, 98.97% IoU) but fell to 95.94% Dice and 92.47% IoU with EfficientNetB3. DeepLab was the weakest model for all backbones, with Dice between 76.30% and 84.64% and IoU between 61.92% and 73.37%, so it is not suitable under this training configuration.

Table 3: Experimental Results using traditional Segmentation Model

BACKBONE/MODEL	DANet		CFNet		SpatialOCRNet		DeepLab	
	DICE	IOU	DICE	IOU	DICE	IOU	DICE	IOU
efficientnetb3	99.38	98.77	99.49	98.99	95.94	92.47	84.64	73.37
Xception	99.49	98.99	99.47	98.96	98.88	97.80	83.29	64.32
resnet50v2	99.49	98.98	99.49	98.99	99.48	98.97	77.29	69.21
mobilenetv2	99.37	98.76	99.48	98.97	98.81	97.67	76.30	61.92

4.5 Lesion classification result

Table 4 gives the results of the models trained from scratch. (without transfer learning). ResNet34 and ResNet50 were best (98% accuracy, with benign recall of 97% and 98%) and the plain CNN reached 96.97% (benign recall 93%). Xception (91%) and MobileNet (82%) detected the minority benign class poorly (benign recall 79% and 55%), and DenseNet121 was unstable (72% accuracy, malignant recall 61% and benign precision 53%)

Table 4. Lesion classification result using Models trained from scratch

Model	Precision		Recall		F1-Score		Accuracy
	Benign	Malignant	Benign	Malignant	Benign	Malignant	
CNN	98%	97%	93%	99%	95%	98%	96.97%
ResNet34	97%	99%	97%	99%	97%	99%	98%
ResNet50	96%	99%	98%	98%	97%	99%	98%
DenseNet121	53%	97%	95%	61%	68%	75%	72%
Xception	90%	91%	79%	96%	84%	93%	91%
MobileNet	83%	82%	55%	95%	66%	88%	82%

Table 5 gives the pre-trained results. InceptionResNetV2 was the best model, with 98%, 99% and 99% precision, recall and F1-score respectively for the benign / malignant classes and 99% accuracy. VGG16, ResNet50, DenseNet201 and Xception followed with 98%, InceptionV3 and EfficientNetB3 with 97% and MobileNetV2 with 96%. Pre-training clearly helped the models that were weak from scratch: Xception rose from 91% to 98% accuracy and from 79% to 98% benign recall. ConvNeXtSmall and NASNetMobile did not learn the task: both predicted every slice as malignant (benign recall 0%, malignant recall 100%), giving 69% accuracy, which equals the proportion of malignant slices. They should be regarded as failed runs rather than as evidence about the architectures themselves, and their causes were not investigated.

Table 5: Lesion classification result using pre-trained model

Model	Precision		Recall		F1-Score		Accuracy
	Benign	Malignant	Benign	Malignant	Benign	Malignant	
VGG16	97%	99%	97%	99%	97%	99%	98%
InceptionV3	96%	98%	95%	98%	96%	98%	97%
InceptionResNetV2	98%	99%	98%	99%	98%	99%	99%
ResNet50	96%	99%	98%	98%	97%	99%	98%
EfficientNetB3	95%	98%	96%	98%	96%	98%	97%
DenseNet201	96%	99%	97%	98%	96%	98%	98%
Xception	97%	99%	98%	99%	97%	99%	98%
ConvNeXtSmall	0%	69%	0%	100%	0%	82%	69%
MobileNetv2	92%	98%	95%	96%	94%	97%	96%

NASNetMobile	0%	69%	0%	100%	0%	82%	69%
--------------	----	-----	----	------	----	-----	-----

5. Conclusion

This paper presents a segmentation-based liver lesion classification framework using the LiTS17 CT dataset. The framework first preprocesses abdominal CT images, then performs liver and tumor segmentation, extracts lesion-focused representations and finally classifies the lesion as benign or malignant. This paper benchmarked preprocessing and architectures for liver and tumour segmentation on LiTS17 using one consistent data-preparation pipeline. HU windowing combined with min-max normalization and CLAHE gave the best U-Net generalisation, whereas z-score normalisation and median filtering reduced it. Of the eight architectures, DANet and CFNet with Xception or ResNet50V2 backbones gave the most accurate and consistent segmentation (about 99.5% Dice and 99% IoU, best 99.52% and 99.05% for CFNet with ResNet50V2 under the [-150, 250] HU window), the traditional U-Net family reached Dice scores of 99.76–99.87% with IoU up to 95.65%, and DeepLab was not competitive. For benign versus malignant lesion classification, pre-trained InceptionResNetV2 was best (99% accuracy; 98% benign and 99% malignant F1-score), VGG16, ResNet50, Xception and DenseNet201 followed at 98%, and ConvNeXtSmall and NASNetMobile collapsed to the majority malignant class (69%). Future work will study further U-Net variants, compare loss functions for segmentation, add Gaussian and BM3D filtering

References

- [1] H. Sung et al., "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," *CA: A Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209–249, 2021.
- [2] K. Y. Elbanna and A. Z. Kielar, "Computed Tomography Versus Magnetic Resonance Imaging for Hepatic Lesion Characterization/Diagnosis," *Clinical Liver Disease*, vol. 17, no. 3, pp. 159–164, 2021.
- [3] G. Litjens et al., "A Survey on Deep Learning in Medical Image Analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.
- [4] P. F. Christ et al., "Automatic Liver and Tumor Segmentation of CT and MRI Volumes using Cascaded Fully Convolutional Neural Networks," *arXiv preprint*, 2017.
- [5] P. Bilic et al., "The Liver Tumor Segmentation Benchmark (LiTS)," *Medical Image Analysis*, vol. 84, pp. 1–24, 2023.
- [6] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Proc. MICCAI*, 2015, pp. 234–241.
- [7] M. Islam, K. N. Khan, and M. S. Khan, "Evaluation of Preprocessing Techniques for U-Net Based Automated Liver Segmentation," in *Proc. Int. Conf. Artificial Intelligence (ICAI)*, Islamabad, Pakistan, 2021, pp. 187–192.
- [8] K. Kaczor et al., "Comparison of Image Pre-processing Methods in Liver Segmentation Task," in *Proc. 15th Int. Conf. Human System Interaction (HSI)*, Melbourne, Australia, 2022, pp. 1–5.
- [9] K. Alawneh et al., "LiverNet: Diagnosis of Liver Tumors in Human CT Images," *Applied Sciences*, vol. 12, no. 11, p. 5501, 2022.
- [10] P. R. Aparna and T. M. Libish, "Automatic Segmentation and Classification of the Liver Tumor using Deep Learning Algorithms," in *Proc. 3rd Int. Conf. Advances in Computing, Communication, Embedded and Secure Systems (ACCESS)*, Kalady, India, 2023, pp. 334–339.
- [11] A. Chaurasia and E. Culurciello, "LinkNet: Exploiting Encoder Representations for Efficient Semantic Segmentation," in *Proc. IEEE Visual Communications and Image Processing (VCIP)*, 2017, pp. 1–4.
- [12] T.-Y. Lin et al., "Feature Pyramid Networks for Object Detection," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2117–2125.
- [13] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid Scene Parsing Network," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2881–2890.
- [14] J. Fu et al., "Dual Attention Network for Scene Segmentation," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 3146–3154.
- [15] H. Zhang et al., "Co-occurrent Features in Semantic Segmentation," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 548–557.
- [16] Y. Yuan, X. Chen, and J. Wang, "Object-Contextual Representations for Semantic Segmentation," in *Proc. European Conf. Computer Vision (ECCV)*, 2020, pp. 173–190.
- [17] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation," in *Proc. European Conf. Computer Vision (ECCV)*, 2018, pp. 801–818.
- [18] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.

- [19] K. He, X. Zhang, S. Ren, and J. Sun, "Identity Mappings in Deep Residual Networks," in Proc. European Conf. Computer Vision (ECCV), 2016, pp. 630–645.
- [20] S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, "Aggregated Residual Transformations for Deep Neural Networks," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2017, pp. 1492–1500.
- [21] J. Hu, L. Shen, and G. Sun, "Squeeze-and-Excitation Networks," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2018, pp. 7132–7141.
- [22] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2016, pp. 2818–2826.
- [23] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning," in Proc. AAAI Conf. Artificial Intelligence, vol. 31, no. 1, 2017, pp. 4278–4284.
- [24] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," in Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR), 2018, pp. 4510–4520.
- [25] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in Proc. Int. Conf. Machine Learning (ICML), 2019, pp. 6105–6114.
- [26] F. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 1800–1807.
- [27] K. Zuiderveld, "Contrast Limited Adaptive Histogram Equalization," in Graphics Gems IV, P. S. Heckbert, Ed. Academic Press, 1994, pp. 474–485.
- [28] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering," IEEE Trans. Image Processing, vol. 16, no. 8, pp. 2080–2095, 2007.
- [29] A.-C. Phan et al., "Improving Liver Lesions Classification on CT/MRI Images Based on Hounsfield Units Attenuation and Deep Learning," Gene Expression Patterns, vol. 47, 2023.
- [30] P. Yakubovskiy, "Segmentation Models," GitHub repository, https://github.com/qubvel/segmentation_models.
- [31] J. M. Kezmann, "TensorFlow Advanced Segmentation Models," GitHub repository, <https://github.com/JanMarcelKezmann/TensorFlow-Advanced-Segmentation-Models>.
- [32] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," in Proc. Int. Conf. Learning Representations (ICLR), 2015.
- [33] L. R. Dice, "Measures of the Amount of Ecologic Association Between Species," Ecology, vol. 26, no. 3, pp. 297–302, 1945.
- [34] P. Jaccard, "The Distribution of the Flora in the Alpine Zone," New Phytologist, vol. 11, no. 2, pp. 37–50, 1912.
- [35] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," in Proc. Int. Conf. Learning Representations (ICLR), 2015.
- [36] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2017, pp. 4700–4708.
- [37] Z. Liu et al., "A ConvNet for the 2020s," in Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR), 2022, pp. 11976–11986.
- [38] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le, "Learning Transferable Architectures for Scalable Image Recognition," in Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR), 2018, pp. 8697–8710.
- [39] A. G. Howard et al., "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," arXiv preprint arXiv:1704.04861, 2017.