

Original Research

Received 25 June 2026

Revised 01 Aug 2026

Accepted 14 Aug 2026

Natural Resources for Human Health 2026, Vol. 6 (17s): 940–954

KEYWORDS:

liver lesion classification; LiTS'17; computed tomography; machine learning; deep learning; liver segmentation; tumour segmentation; handcrafted features; ensemble classifier; transfer learning.

Automated Liver Lesion Detection and Classification on LiTS'17

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

ABSTRACT: Automated detection and classification of liver lesions from computed tomography (CT) can give radiologists objective, reproducible information about lesion location and malignancy. This paper presents a five-phase framework for benign versus malignant liver lesion classification on the Liver Tumor Segmentation Challenge 2017 (LiTS'17) dataset and evaluates it through two routes: a machine learning (ML) route built on classical segmentation, handcrafted features and conventional classifiers, and a deep learning (DL) route built on advanced segmentation networks and convolutional neural network (CNN) classifiers. Both routes share a pre-processing phase of Hounsfield Unit (HU) windowing, resizing, normalisation, Contrast Limited Adaptive Histogram Equalization (CLAHE), and data augmentation. In the ML route, four segmentation pipelines (marker-controlled watershed with Gaussian Mixture Models, Fuzzy C-Means with Grey Wolf Optimization, FCM with region growing, and Kirsch edge detection with mean-shift smoothing and FCM) are combined with shape, intensity and texture features (FOS, GLCM, GLRLM, LBP, Gabor, SFTA), PCA-based reduction and an ensemble of KNN, MLP, SVM and Random Forest; the best pipeline reaches 96.97% accuracy, 96.5% precision and 96% recall. In the DL route, four advanced segmentation models (DANet, CFNet, SpatialOCRNet and DeepLab) are benchmarked with four backbones; CFNet and DANet give the most consistent masks, with about 99.5% Dice and 99.0% IoU. For classification, ImageNet-pretrained InceptionResNetV2 achieves 99% accuracy with benign and malignant F1-scores of 98% and 99%. The results show that the DL route outperforms the best ML route by about 2 percentage points in accuracy and removes the need for handcrafted feature design, while the ML route remains a competitive and interpretable baseline.

1. INTRODUCTION

Liver cancer is one of the leading causes of cancer death worldwide [34]. It may be primary, most often hepatocellular carcinoma (HCC), or secondary, arising from metastases of colorectal, breast, pancreatic and other cancers. Early detection, accurate characterisation and treatment planning depend strongly on imaging, and contrast-enhanced computed tomography (CT) is the most common modality for characterising focal hepatic lesions [35]. However, reading hundreds of slices per patient and delineating the liver and its lesions by hand is slow and suffers from inter-observer variability, particularly for small, low-contrast or multiple lesions.

Computer-aided diagnosis (CAD) systems aim to automate lesion detection, delineation and characterisation so that radiologists receive consistent, quantitative support. Automatic classification of liver lesions is challenging because lesions vary widely in size, shape, attenuation and anatomical context. A classifier trained directly on whole abdominal CT slices also receives a large amount of irrelevant information from the organs and tissues around the liver. Many CAD pipelines therefore segment the liver first, then the lesion inside it, and classify only the segmented lesion [15], [39], [40].

Research in this area has followed the broader evolution of medical image analysis [36]. Early systems relied on handcrafted texture and shape descriptors with conventional machine learning (ML) classifiers such as support vector machines (SVM) and random forests [2]–[12]. Since 2015, deep learning (DL), and in particular convolutional neural networks (CNNs) and encoder–decoder networks such as U-Net [13], has displaced handcrafted pipelines and now dominates public benchmarks [1], [14]–[21]. Attention modules and Vision Transformers have more recently been introduced to capture long-range context [30]–[33].

The Liver Tumor Segmentation (LiTS) benchmark [1] provides the most widely used public CT dataset for this task, with expert-labelled liver and tumour masks. LiTS does not, however, include lesion-type labels, and most studies evaluate either an ML pipeline or a DL pipeline in isolation, with different pre-processing, splits and metrics. A side-by-side evaluation of both paradigms on the same LiTS'17 data, under one shared pre-processing phase, is rarely reported.

This paper addresses that gap with a unified framework for automated liver lesion detection and classification on LiTS'17, using benign and malignant labels assigned to the LiTS training volumes by Alawneh et al. [39]. The contributions are:

- A five-phase framework (pre-processing, segmentation, feature extraction, feature selection and classification) that supports both an ML route and a DL route on the same data.
- An evaluation of four classical segmentation pipelines for the ML route, combined with shape, intensity and texture features, PCA-based reduction and SVM or ensemble classifiers.
- A benchmark of four advanced segmentation architectures (DANet, CFNet, SpatialOCRNet and DeepLab) with four pre-trained backbones for three-class liver and tumour segmentation.
- A benchmark of six scratch-trained and ten ImageNet-pretrained CNNs for benign versus malignant lesion classification, and a direct comparison of the ML and DL routes.

Section 2 reviews related work. Section 3 describes the proposed framework and both routes. Section 4 reports and discusses the experimental results, and Section 5 concludes the paper.

2. LITERATURE REVIEW

2.1 The LiTS'17 dataset

LiTS was organised with ISBI 2017 and MICCAI 2017/2018 [1]. It contains 131 annotated training CT volumes and 70 test volumes from seven hospitals and research institutions, with heterogeneous scanners, slice thicknesses (about 0.7–5 mm) and contrast phases. The volumes include primary and secondary tumours of varying size, number and appearance, with both hyper- and hypo-dense lesions. Across 75 submitted algorithms, the best liver Dice was 0.963, whereas the best tumour Dice was 0.674, 0.702 and 0.739 in the three rounds, and the best lesion-wise recall was only 0.554 [1].

LiTS provides voxel-wise labels for liver and tumour only, not histopathological lesion types. Classification studies on LiTS therefore address slice- or patch-level tumour versus non-tumour classification, detection derived from segmentation, or benign/malignant labels obtained from additional annotation [39]. Multi-class lesion-type classification is usually reported on private clinical datasets [6], [7], [10], [22]–[25], which complicates comparison.

2.2 Machine learning approaches

Conventional ML pipelines follow a fixed sequence: pre-processing (HU windowing, denoising), liver and lesion segmentation (thresholding, region growing, watershed, level sets or graph cuts), handcrafted feature extraction, feature selection and classification. Texture is the most widely used cue because lesions differ from normal parenchyma in heterogeneity and enhancement pattern. Grey-level co-occurrence matrix (GLCM) statistics [2] remain the most popular descriptors, complemented by first-order statistics, run-length features, Gabor and wavelet responses and local binary patterns (LBP) [3]. Shape descriptors and, in multiphase studies, kinetic enhancement features are often added [12]. The features are fed to SVMs [4], random forests [5], k-nearest neighbours, logistic regression, shallow neural networks or extreme learning machines. Gletsos et al. [6] built one of the first CT focal liver lesion CAD systems, using GLCM features and a cascade of neural networks to separate normal liver, cysts, hemangiomas and HCC. Mougiakakou et al. [7] added genetic-algorithm feature selection and an ensemble of classifiers, reaching about 85% accuracy and showing that ensembles outperform single classifiers. Kumar et al. [8] segmented the liver and lesions automatically before extracting contourlet texture features for HCC versus hemangioma classification. Huang et al. [9] used an ensemble of extreme learning machines on random feature subspaces, and Diamant et al. [10] showed that a bag-of-visual-words model learnt separately for the lesion interior and boundary improves SVM

classification. Conze et al. [11] used multi-scale supervoxels and random forests on dynamic contrast-enhanced CT, and Chang et al. [12] combined texture, shape and kinetic features for benign/malignant classification. Table 1 summarises these studies.

Table 1. Machine learning approaches for liver lesion detection and classification

Author [Ref.]	Features / pre-processing	Classifier	Data	Task	Key finding
Gletsos et al. [6]	First-order and GLCM texture from manual lesion ROIs	Cascade of feed-forward NNs	Clinical non-contrast CT	Normal, cyst, hemangioma, HCC	Texture alone separates the main lesion types
Mougiakakou et al. [7]	Multiple texture sets with GA feature selection	Ensemble of NN, kNN and others	Clinical non-contrast CT	Normal, cyst, hemangioma, HCC	Ensemble reached $\approx 84.96\%$ accuracy, above single classifiers
Kumar et al. [8]	Automatic segmentation + contourlet co-occurrence texture	Probabilistic NN	Clinical CT	HCC vs hemangioma	Fully automatic segment-then-classify CAD
Huang et al. [9]	Voxel intensity/texture over random subspaces	Ensemble ELM	Contrast CT	Tumour detection and segmentation	Subspace ensemble more robust than a single ELM
Diamant et al. [10]	Bag-of-visual-words for interior and boundary	SVM	Portal-venous CT	Cyst, metastasis, hemangioma	Separate interior/boundary modelling improved accuracy
Conze et al. [11]	Multi-scale supervoxel texture, multiphase CT	Random forest	Dynamic contrast CT	Tumour vs parenchyma	Supervoxels cut computation; multiphase data helped
Chang et al. [12]	Texture, shape and kinetic features	Logistic regression	Contrast CT	Benign vs malignant	Combined cues beat any single feature group

Results are as reported in the original studies, which use different datasets and protocols and are not directly comparable.

ML approaches have four main limitations. Feature values depend on accurate lesion boundaries, so segmentation errors propagate into classification. Handcrafted descriptors struggle with the variability of multi-centre data such as LiTS [1]. Feature selection and parameter tuning need expert effort and transfer poorly across scanners. Most ML studies were also validated on small, single-centre datasets.

2.3 Deep learning approaches

Fully convolutional networks and U-Net [13] enabled end-to-end pixel-wise prediction. Ben-Cohen et al. [14] showed that an FCN can segment the liver and detect lesions with few false positives. Christ et al. [15] introduced the cascaded approach, in which one FCN segments the liver, a second segments lesions inside it and a 3D conditional random field refines the result;

this became the standard LiTS template. H-DenseUNet [16] fused 2D intra-slice and 3D inter-slice features and reached a tumour Dice of 0.722 per case and 0.824 globally on the LiTS test set. Kaluva et al. [17] used a lightweight 2D dense FCN cascade, Vorontsov et al. [18] trained liver and lesion segmentation jointly, Seo et al. [19] added residual paths to U-Net skip connections for small tumours, Ayalew et al. [21] addressed class imbalance, and nnU-Net [20] showed that a self-configured U-Net is a very strong baseline.

A parallel line of work classifies lesion patches or slices directly with CNNs. Yasaka et al. [22] classified five categories of liver masses on three-phase CT with 0.84 accuracy, and Frid-Adar et al. [23] used GAN-generated lesions to raise sensitivity from 78.6% to 85.7%. Liang et al. [24] combined a residual CNN with a bidirectional LSTM for multiphase CT, and Hamm et al. [25] reported 92% accuracy for six lesion classes on multiphase MRI. ImageNet-pretrained backbones such as ResNet [28] and DenseNet [29] are widely used as feature extractors. Hybrid systems combine classical and deep components: Das et al. [26] used watershed and Gaussian mixture model segmentation with a deep neural network classifier, and Ghoniem [27] combined SegNet/U-Net segmentation with a LeNet-5 classifier tuned by an artificial bee colony algorithm. Alawneh et al. [39] proposed LiverNet for benign/malignant diagnosis on LiTS, and Phan et al. [41] showed that HU attenuation information improves lesion classification.

CNNs model local patterns well but have a limited receptive field. Attention modules such as CBAM [30] reweight channel and spatial features at little cost, and context-aggregation segmentation networks such as DANet [59], CFNet [60], OCRNet [61] and DeepLab [62] capture longer-range dependencies. Vision Transformers [31] model global relationships but need large datasets, so hybrid designs such as TransUNet [32] and UNETR [33] pair CNNs with Transformers. Table 2 summarises the main DL studies.

Table 2. Deep learning approaches for liver lesion segmentation, detection and classification

Author [Ref.]	Architecture	Data	Task	Reported result / finding
Ben-Cohen et al. [14]	FCN (VGG-based)	Clinical CT	Liver segmentation + lesion detection	Lesion detection with few false positives
Christ et al. [15]	Cascaded 2D FCNs + 3D CRF	3DIRCADb	Liver and lesion segmentation	Liver Dice above 0.94; became the LiTS template
Li et al. [16]	H-DenseUNet	LiTS	Liver and tumour segmentation	Tumour Dice 0.722 per case, 0.824 global
Kaluva et al. [17]	2D dense FCN cascade	LiTS	Liver and tumour segmentation	Competitive at low computational cost
Vorontsov et al. [18]	Joint liver–lesion FCN	LiTS	Lesion segmentation	Liver context improved lesion segmentation
Seo et al. [19]	mU-Net	LiTS, 3DIRCADb	Liver and tumour segmentation	Better small-tumour handling than U-Net
Isensee et al. [20]	nnU-Net	MSD Liver	Liver and tumour segmentation	Strong, reproducible baseline
Bilic et al. [1]	LiTS benchmark, 75 submissions	LiTS	Segmentation and detection	Best tumour Dice 0.739; lesion recall 0.554
Yasaka et al. [22]	CNN	3-phase CT	5-category classification	Accuracy 0.84; AUC 0.92

Author [Ref.]	Architecture	Data	Task	Reported result / finding
Frid-Adar et al. [23]	CNN + GAN augmentation	182 CT lesions	Cyst, metastasis, hemangioma	Sensitivity 78.6% → 85.7%
Liang et al. [24]	Residual CNN + BiLSTM	Multiphase CT	Four lesion types	Accuracy ≈90.9%
Hamm et al. [25]	3D CNN	Multiphase MRI	Six lesion classes	Accuracy 92%
Das et al. [26]	Watershed + GMM + DNN	Clinical CT	HCC, hemangioma, metastasis	Accuracy 99.38% (small dataset)
Ghoniem [27]	SegNet/U-Net + LeNet-5	Public CT	Benign vs malignant	Beat feature-engineering baselines
Chen et al. [32]	TransUNet	Synapse CT	Organ segmentation	CNN-Transformer beat pure CNNs
Hatamizadeh et al. [33]	UNETR	BTCV, MSD	3D segmentation	Strong long-range context modelling

3. AUTOMATED LIVER LESION DETECTION AND CLASSIFICATION FRAMEWORK

The proposed framework processes abdominal CT images in five phases, shown in Fig. 1. Phase 1 pre-processes the CT slices, and Phase 2 segments the liver and then the lesion inside it. After segmentation the framework branches. The ML route extracts handcrafted shape, texture and intensity features (Phase 3), selects or reduces them (Phase 4) and classifies them with conventional classifiers (Phase 5). The DL route passes the segmented region directly to CNN-based, Transformer or hybrid models (Phase 5), which learn their own features.

3.1 Dataset

This work uses the Liver Tumor Segmentation Challenge 2017 (LiTS'17) dataset [1], which consists of 130 training and 70 test contrast-enhanced 3D abdominal CT scans from several clinical centres with different scanners and protocols. The in-plane resolution ranges from 0.60 mm to 0.98 mm and the slice spacing from 0.45 mm to 5.0 mm. Every axial slice is 512×512 pixels, while the number of slices per volume ranges from 42 to 1,026. Ground-truth liver and lesion masks are public only for the training split, and no lesion type is provided. Alawneh et al. [39] assigned each training volume to one of three categories: 4 normal, 41 benign and 85 malignant. After extracting the slices that contain liver and tumour, the working dataset contains 6,930 images, with 2,602 benign and 4,766 malignant images.

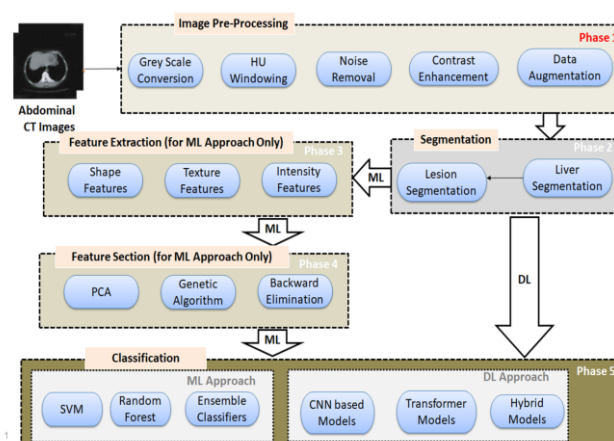


Fig. 1. Proposed five-phase framework for automated liver lesion detection and classification. Phases 3 and 4 apply to the ML route only.

The data are prepared in three steps:

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 HU window is applied to each slice and the slice is saved as a .jpg image.
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 as a .png file with labels 0 (background), 1 (liver) and 2 (tumour).
3. **Slice selection and indexing.** Slices that contain liver and tumour are listed in a .csv file with the path of the CT slice, the path of its mask and the lesion type, so that every image is linked to its mask and label.

3.2 Phase 1: Image pre-processing

Both routes share the following pre-processing techniques, since the treatment of raw CT intensities noticeably affects segmentation quality [37], [38].

Grey-scale conversion and HU windowing. CT intensities are expressed in Hounsfield Units. Windowing maps a chosen HU range $[L, U]$ to grey levels: values below L appear black, values above U appear white, and values in between are scaled linearly. A wide window assigns many HU values to each grey level and hides small attenuation differences, while a narrow window around soft tissue emphasises the liver parenchyma and lesions and suppresses bone and air [41]. Several window sizes were examined, and the window $[-30, 150]$ HU was selected.

Resizing. Images are resized to a fixed input size before segmentation and classification based on Model like for the advanced segmentation models use 320×320 inputs.

Normalisation. Min-max normalisation rescales intensities to $[0, 1]$ applied

Contrast enhancement with CLAHE. Contrast Limited Adaptive Histogram Equalization [42] equalises the histogram of small tiles rather than the whole image, which reveals fine structures where the contrast between tissues is poor. A 7×7 tile grid and a clip limit of 0.2 are used so that noise is not over-amplified, and neighbouring tiles are merged by bilinear interpolation.

Data augmentation. The dataset is imbalanced, with 2,602 benign and 4,766 malignant images. The minority benign class is enlarged with classical transformations: horizontal and vertical flips, rotations by 90°, 180° and 270°, and random shifts of 10–15 pixels horizontally and vertically. This balances the classes and reduces bias during training.

3.3 Machine learning route

The ML route follows Phases 2 to 5 of Fig. 1 with classical, unsupervised segmentation and handcrafted features. Four pipelines were designed; their components are summarised below and their results in Table 3.

3.3.1 Phase 2: Liver and lesion segmentation

Pipeline 1: Marker-controlled watershed with GMM. The liver is segmented with a marker-controlled watershed transform [44]. Otsu thresholding [43] first separates foreground from background and provides the markers, the watershed transform splits touching regions along intensity ridges, and morphological operations remove small spurious regions and fill holes. Lesions are then segmented inside the liver with a Gaussian Mixture Model (GMM), whose parameters are estimated by expectation–maximisation [45] and which assigns each pixel to the intensity component of liver or lesion tissue.

Pipeline 2: FCM optimised by Grey Wolf Optimization, with Fast FCM. The liver is segmented with Fuzzy C-Means (FCM) clustering [46], which gives each pixel a degree of membership in every cluster by minimising the objective below, where u_{ij} is the membership of pixel x_i in cluster j , c_j the cluster centre and $m > 1$ the fuzzifier. Because FCM is sensitive to its initial centres, Grey Wolf Optimization (GWO) [47], a swarm-based metaheuristic modelled on the hunting hierarchy of grey wolves, searches for the cluster centres that minimise this objective. Lesions are then segmented inside the liver with Fast FCM [48], which clusters the grey-level histogram rather than every pixel and is therefore much faster.

Pipeline 3: FCM with region growing. The same procedure is applied to both liver and lesion. FCM clustering produces an initial coarse partition, seeded region growing [49] refines it by adding neighbouring pixels whose intensity is close to the seed region, and morphological operations smooth the boundary and remove small components.

Pipeline 4: Kirsch edges and mean-shift smoothing with FCM. For both liver and lesion, Kirsch compass edge detection [50] highlights organ and lesion boundaries in eight directions, and mean-shift smoothing [51] flattens texture inside homogeneous regions while preserving those edges. FCM clustering then refines the partition, and morphological operations clean the final mask.

3.3.2 Phase 3: Feature extraction

Three groups of features are extracted from each segmented lesion region:

- **Intensity features:** first-order statistics (FOS) of the grey-level histogram, such as mean, variance, skewness, kurtosis, energy and entropy.
- **Texture features:** GLCM features [2] (contrast, correlation, energy, homogeneity), grey-level run-length matrix (GLRLM) features [52], local binary patterns (LBP) [3], Gabor filter responses at several scales and orientations [53], and segmentation-based fractal texture analysis (SFTA) [54].
- **Shape features:** area, perimeter, circularity, compactness, eccentricity and solidity of the lesion region.

Pipeline 1 uses GLCM, LBP and Gabor features. Pipelines 2 and 4 use FOS, GLCM, LBP, SFTA and shape features, and Pipeline 3 uses FOS, GLCM, GLRLM, LBP, Gabor and SFTA features.

3.3.3 Phase 4: Feature normalisation and selection

In Pipelines 2 to 4, all features are scaled to [0, 1] with min-max scaling so that no feature dominates because of its range. The framework supports principal component analysis (PCA), genetic-algorithm selection and backward elimination; in these experiments PCA [55] projects the features onto the principal components that retain most of the variance, which removes redundancy between correlated texture descriptors.

3.3.4 Phase 5: Classification

Pipeline 1 uses a single SVM [4], which finds the maximum-margin hyperplane between benign and malignant feature vectors. Pipelines 2 to 4 use an ensemble [58] of four classifiers: k-nearest neighbours (KNN) [56], a multilayer perceptron (MLP) [57], an SVM [4] and a Random Forest (RF) [5]. The ensemble combines the complementary decision boundaries of instance-based, neural, margin-based and tree-based learners, which is more robust than any one of them.

3.4 Deep learning route

The DL route applies HU windowing [−30, 150], min-max normalisation and CLAHE, then uses advanced segmentation networks for Phase 2 and CNN classifiers for Phase 5.

3.4.1 Phase 2: Liver and tumour segmentation with advanced models

Segmentation is a three-class problem: background, liver and tumour. Four advanced, context-aggregation architectures were trained with four ImageNet-pretrained backbones (EfficientNetB3 [65], Xception [64], ResNet50V2 [63] and MobileNetV2 [66]) using the TensorFlow Advanced Segmentation Models library [67]:

- **DANet** [59] adds two parallel attention modules to a dilated backbone. The position attention module relates every pixel to every other pixel, and the channel attention module models dependencies between feature maps, so that features of the same structure reinforce each other across the image.
- **CFNet** [60] learns co-occurrent features: an aggregated co-occurrent feature module estimates which semantic classes tend to appear together and uses this context to refine each pixel's representation, which suits the fixed liver–tumour spatial relationship.
- **SpatialOCRNet** [61] uses object-contextual representations. A coarse segmentation forms a soft representation of each object region, and each pixel is then augmented with the representations of the regions it is related to.
- **DeepLab** [62] uses atrous (dilated) separable convolutions and atrous spatial pyramid pooling to capture multi-scale context, with an encoder–decoder structure that recovers object boundaries.

The segmented liver and tumour regions are passed to the classification phase, so the classifier sees the lesion with little interference from surrounding organs.

3.4.2 Phase 5: Lesion classification

The classification phase separates benign and malignant lesions. Two groups of models were compared:

- **Transfer learning:** ten ImageNet-pretrained models were fine-tuned: VGG16 [68], InceptionV3 [69], InceptionResNetV2 [70], ResNet50 [28], EfficientNetB3 [65], DenseNet201 [29], Xception [64], ConvNeXtSmall [71], MobileNetV2 [66] and NASNetMobile [72].
- **Training from scratch:** six models were trained with random initialisation: a plain CNN, ResNet34, ResNet50 [28], DenseNet121 [29], Xception [64] and MobileNet [73].

All classifiers used the Adam optimiser [74] with a learning rate of 0.0001 and batch sizes of 8 or 16.

4. EXPERIMENTAL RESULTS

The best DL route (CFNet or DANet segmentation followed by pretrained InceptionResNetV2) reaches 99% classification accuracy, against 96.97% for the best ML route (FCM + GWO segmentation with an ensemble classifier). This section describes the setup and metrics, then reports the ML route, the advanced segmentation models and the lesion classifiers, and closes with a comparison of the two routes.

4.1 Experimental setup

All experiments were run in Google Colab Pro on an NVIDIA GPU. The abdominal CT slices of LiTS'17 were randomly divided into training, validation and test subsets in the ratio 70% : 10% : 20%. The DL models were trained with the Adam optimiser [74], a learning rate of 0.0001, 50 or 100 epochs and batch sizes of 8 or 16. The advanced segmentation models used 320×320 inputs, and all models used the $[-30, 150]$ HU window.

4.2 Evaluation metrics

Classification is evaluated with accuracy, precision, recall and F1-score, computed from true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN). Precision, recall and F1-score are reported per class (benign and malignant) for the DL classifiers.

Segmentation is evaluated with the Dice similarity coefficient [75] and the Intersection over Union (IoU, Jaccard index) [76] between the predicted mask A and the ground-truth mask B.

4.3 Results of the machine learning route

Table 3 reports the four ML pipelines. Segmentation quality and feature richness decide the outcome: the pipeline with the weakest segmentation and the smallest feature set performs poorly, while the pipelines that pair FCM-based segmentation with a broad feature set and an ensemble exceed 86% accuracy.

Table 3. Experimental results of the traditional (ML) techniques

Pipeline	Segmentation	Features and reduction	Classifier	Accuracy (%)	Precision (%)	Recall (%)
P1	Liver: marker-controlled watershed (Otsu + watershed + morphology); lesion: GMM	GLCM, LBP, Gabor	SVM	59.00	61.00	59.00
P2	Liver: FCM + GWO; lesion: Fast FCM	FOS, GLCM, LBP, SFTA, shape; min-max; PCA	Ensemble (KNN, MLP, SVM, RF)	96.97	96.50	96.00

Pipeline	Segmentation	Features and reduction	Classifier	Accuracy (%)	Precision (%)	Recall (%)
P3	Liver and lesion: FCM pre-processing, region-growing refinement, morphological post-processing	FOS, GLCM, GLRLM, LBP, Gabor, SFTA; min-max; PCA	Ensemble (KNN, MLP, SVM, RF)	86.76	82.00	80.00
P4	Liver and lesion: Kirsch edges + mean-shift smoothing, FCM refinement, morphological post-processing	FOS, GLCM, LBP, SFTA, shape; min-max; PCA	Ensemble (KNN, MLP, SVM, RF)	95.74	96.00	94.00

Pipeline P1 reached only 59% accuracy, 61% precision and 59% recall. Watershed segmentation tends to over-segment the liver when organ boundaries are weak, and the GMM struggles when lesion and parenchyma intensities overlap, so the extracted regions are inaccurate. The small texture-only feature set and a single SVM cannot compensate for these errors.

Pipeline P2 gave the best ML result, with 96.97% accuracy, 96.5% precision and 96% recall. GWO avoids the poor local minima that a randomly initialised FCM can reach, which yields more accurate liver masks, and Fast FCM then isolates the lesion reliably. Adding shape features to the intensity and texture descriptors, with PCA removing redundancy, gives the ensemble a compact and discriminative input.

Pipeline P4 was close behind, at 95.74% accuracy, 96% precision and 94% recall. Kirsch edge detection and mean-shift smoothing sharpen organ boundaries before FCM refinement, and the pipeline uses the same feature set as P2. Pipeline P3 reached 86.76% accuracy but only 82% precision and 80% recall. Region growing is sensitive to seed placement and can leak into adjacent tissue of similar intensity, and P3 omits shape features, which carry much of the benign/malignant distinction.

4.4 Results of the advanced segmentation models

Fig 2 reports Dice and IoU for DANet, CFNet, SpatialOCRNet and DeepLab with four backbones at 320×320 resolution under the $[-30, 150]$ HU window.

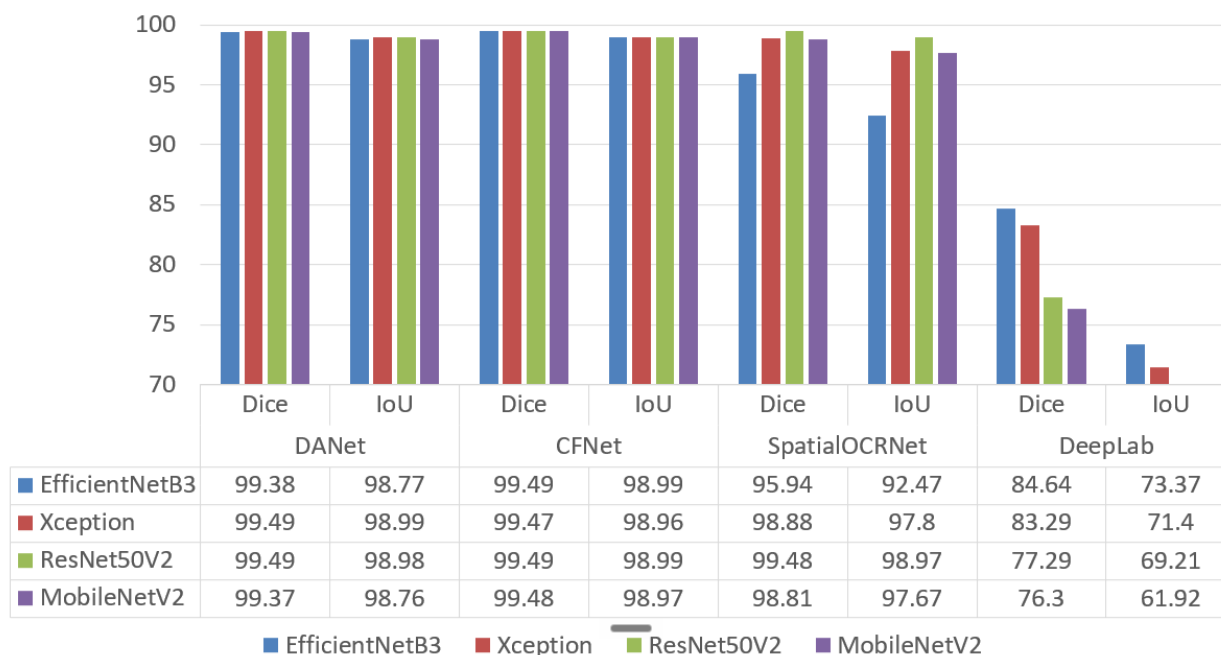


Fig. 2: Dice and IoU (%) of the advanced segmentation models

CFNet was the most consistent model, with 99.47–99.49% Dice and 98.96–98.99% IoU across all four backbones, so its performance hardly depends on the encoder. Its co-occurrent feature module exploits the fixed relationship between liver and tumour, since a tumour label is only plausible inside the liver. DANet matched CFNet's best result with the Xception backbone (99.49% Dice, 98.99% IoU) and stayed above 99.3% Dice with every backbone, which shows the benefit of modelling position and channel dependencies jointly. SpatialOCRNet depended strongly on the backbone. With ResNet50V2 it was close to CFNet (99.48% Dice, 98.97% IoU), but with EfficientNetB3 it fell to 95.94% Dice and 92.47% IoU, because its object-region representations rely on a good coarse segmentation from the backbone. DeepLab was the weakest model with every backbone, with Dice between 76.30% and 84.64% and IoU between 61.92% and 73.37%. Its atrous convolutions enlarge the receptive field but lose the fine spatial detail needed for small tumours, so it is not suitable under this training configuration. CFNet and DANet were therefore selected to produce the lesion regions for classification.

4.5 Lesion classification results

Table 5 reports the models trained from scratch, without transfer learning.

Table 5. Lesion classification results of models trained from scratch

Model	Precision		Recall		F1-Score		Accuracy
	Benign	Malignant	Benign	Malignant	Benign	Malignant	
CNN	99.38	98.77	99.49	98.99	95.94	92.47	84.64
ResNet34	99.49	98.99	99.47	98.96	98.88	97.80	64.32
ResNet50	99.49	98.98	99.49	98.99	99.48	98.97	77.29
DenseNet121	99.37	98.76	99.48	98.97	98.81	97.67	76.30
Xception	99.38	98.77	99.49	98.99	95.94	92.47	84.64
MobileNet	99.49	98.99	99.47	98.96	98.88	97.80	64.32

ResNet34 and ResNet50 were the best scratch-trained models, with 98% accuracy and benign recall of 97% and 98%; their residual connections let them train well on the moderate dataset. The plain CNN reached 96.97% accuracy but a lower benign recall of 93%. Xception (91%) and MobileNet (82%) detected the minority benign class poorly, with benign recall of 79% and 55%. DenseNet121 was unstable, with 72% accuracy, 61% malignant recall and 53% benign precision.

Table 6 reports the ImageNet-pretrained models.

Table 6. Lesion classification results of pretrained models

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%	99%	99%	98.%
<i>ConvNeXtSmall</i>	0%	69%	0%	100%	0%	82%	69%
MobileNetV2	92%	98%	95%	96%	94%	97%	96%
<i>NASNetMobile</i>	0%	69%	0%	100%	0%	82%	69%

InceptionResNetV2 was the best classifier, with 98% / 99% precision, recall and F1-score for the benign / malignant classes and 99% accuracy. Its combination of multi-scale Inception blocks and residual connections captures both fine lesion texture and wider structure. VGG16, ResNet50, DenseNet201 and Xception followed at 98%, InceptionV3 and EfficientNetB3 at 97% and MobileNetV2 at 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. These are treated as failed runs rather than evidence about the architectures, and their causes are left for future investigation.

4.6 Comparison of the ML and DL routes

Table 7 compares the best configuration of each route.

Table 7. Best ML route versus best DL route

Route	Segmentation	Features	Classifier	Accuracy (%)	Precision (%)	Recall (%)
ML (P2)	FCM + GWO (liver), Fast FCM (lesion)	Handcrafted FOS, GLCM, LBP, SFTA, shape + PCA	Ensemble (KNN, MLP, SVM, RF)	96.97	96.50	96.00
DL	CFNet / DANet ($\approx 99.5\%$ Dice, $\approx 99.0\%$ IoU)	Learnt by the network	Pretrained InceptionResNetV2	99.00	98–99	98–99

The DL route outperforms the best ML route by about 2 percentage points in accuracy and 2–3 points in precision and recall. Three factors explain the gap. First, the advanced segmentation networks produce more precise lesion masks than clustering-based segmentation, so fewer errors propagate to classification. Second, CNN features are learnt from data and adapt to the variability of the multi-centre LiTS'17 images, whereas handcrafted descriptors are fixed in advance. Third, transfer learning from ImageNet gives the DL classifiers robust low-level filters despite the moderate dataset size.

The ML route nevertheless remains competitive. Its best pipeline exceeds 96% on all three metrics, it requires far less computation and no GPU, and each of its features has a direct physical meaning, which makes its decisions easier to interpret. The results also confirm that, in both routes, the quality of the segmentation stage largely determines the final classification performance: P1 with weak segmentation reached 59%, and DeepLab's poor masks would make it unsuitable for the DL route.

These results should be interpreted with two limitations in mind. The split is made at slice level, so slices from the same patient may appear in both training and test sets, which can make the results optimistic. The benign/malignant labels are also assigned per volume [39], not per lesion.

5. CONCLUSION

This paper presented a five-phase framework for automated liver lesion detection and benign versus malignant classification on the LiTS'17 CT dataset, and evaluated it through a machine learning route and a deep learning route with a shared pre-processing phase. In the ML route, the quality of classical segmentation decided the outcome: FCM optimised by Grey Wolf Optimization with Fast FCM lesion segmentation, handcrafted FOS, GLCM, LBP, SFTA and shape features, PCA and an ensemble of KNN, MLP, SVM and RF gave the best result, with 96.97% accuracy, 96.5% precision and 96% recall, while watershed and GMM segmentation with an SVM reached only 59%.

In the DL route, among the advanced segmentation models, CFNet and DANet produced the most accurate and consistent masks, at about 99.5% Dice and 99.0% IoU, SpatialOCRNet depended strongly on its backbone and DeepLab was not competitive. For classification, pretrained InceptionResNetV2 was best, with 99% accuracy and benign and malignant F1-scores of 98% and 99%, followed by VGG16, ResNet50, DenseNet201 and Xception at 98%. Transfer learning markedly improved models that were weak from scratch, while ConvNeXtSmall and NASNetMobile collapsed to the majority malignant class.

Overall, the DL route outperformed the best ML route by about 2 percentage points in accuracy and removed the need for handcrafted feature design. The ML route remains a lightweight and interpretable baseline that exceeds 96% on all metrics. Future work will use patient-level splits and cross-validation with external validation on 3DIRCADb, extend the classification to lesion types, evaluate Gaussian and BM3D [77] denoising in the pre-processing phase, and explore segmentation-free CNN–Transformer hybrids with attention mechanisms to improve small-lesion sensitivity and explainability.

REFERENCES

- [1] P. Bilic et al., "The Liver Tumor Segmentation Benchmark (LiTS)," *Medical Image Analysis*, vol. 84, p. 102680, 2023.
- [2] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Systems, Man, and Cybernetics*, vol. SMC-3, no. 6, pp. 610–621, 1973.
- [3] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, 2002.
- [4] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [5] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [6] M. Gletsos, S. G. Mougiakakou, G. K. Matsopoulos, K. S. Nikita, A. S. Nikita, and D. Kelekis, "A computer-aided diagnostic system to characterize CT focal liver lesions: Design and optimization of a neural network classifier," *IEEE Trans. Information Technology in Biomedicine*, vol. 7, no. 3, pp. 153–162, 2003.
- [7] S. G. Mougiakakou, I. K. Valavanis, A. Nikita, and K. S. Nikita, "Differential diagnosis of CT focal liver lesions using texture features, feature selection and ensemble driven classifiers," *Artificial Intelligence in Medicine*, vol. 41, no. 1, pp. 25–37, 2007.
- [8] S. S. Kumar, R. S. Moni, and J. Rajeesh, "An automatic computer-aided diagnosis system for liver tumours on computed tomography images," *Computers & Electrical Engineering*, vol. 39, no. 5, pp. 1516–1526, 2013.

- [9] W. Huang et al., "Random feature subspace ensemble based extreme learning machine for liver tumor detection and segmentation," in Proc. 36th Annu. Int. Conf. IEEE Engineering in Medicine and Biology Society (EMBC), 2014, pp. 4675–4678.
- [10] Diamant et al., "Improved patch-based automated liver lesion classification by separate analysis of the interior and boundary regions," IEEE J. Biomedical and Health Informatics, vol. 20, no. 6, pp. 1585–1594, 2016.
- [11] P.-H. Conze et al., "Scale-adaptive supervoxel-based random forests for liver tumor segmentation in dynamic contrast-enhanced CT scans," Int. J. Computer Assisted Radiology and Surgery, vol. 12, no. 2, pp. 223–233, 2017.
- [12] C.-C. Chang et al., "Computer-aided diagnosis of liver tumors on computed tomography images," Computer Methods and Programs in Biomedicine, vol. 145, pp. 45–51, 2017.
- [13] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in Proc. MICCAI 2015, LNCS, vol. 9351, 2015, pp. 234–241.
- [14] Ben-Cohen, I. Diamant, E. Klang, M. Amitai, and H. Greenspan, "Fully convolutional network for liver segmentation and lesions detection," in Deep Learning and Data Labeling for Medical Applications (DLMIA 2016), LNCS, vol. 10008, 2016, pp. 77–85.
- [15] P. F. Christ et al., "Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields," in Proc. MICCAI 2016, LNCS, vol. 9901, 2016, pp. 415–423.
- [16] X. Li, H. Chen, X. Qi, Q. Dou, C.-W. Fu, and P.-A. Heng, "H-DenseUNet: Hybrid densely connected UNet for liver and tumor segmentation from CT volumes," IEEE Trans. Medical Imaging, vol. 37, no. 12, pp. 2663–2674, 2018.
- [17] K. C. Kaluva, M. Khened, A. Kori, and G. Krishnamurthi, "2D-densely connected convolution neural networks for automatic liver and tumor segmentation," arXiv:1802.02182, 2018.
- [18] E. Vorontsov, A. Tang, C. Pal, and S. Kadoury, "Liver lesion segmentation informed by joint liver segmentation," in Proc. IEEE 15th Int. Symp. Biomedical Imaging (ISBI), 2018, pp. 1332–1335.
- [19] H. Seo, C. Huang, M. Bassenne, R. Xiao, and L. Xing, "Modified U-Net (mU-Net) with incorporation of object-dependent high level features for improved liver and liver-tumor segmentation in CT images," IEEE Trans. Medical Imaging, vol. 39, no. 5, pp. 1316–1325, 2020.
- [20] F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein, "nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation," Nature Methods, vol. 18, no. 2, pp. 203–211, 2021.
- [21] Y. A. Ayalew, K. A. Fante, and M. A. Mohammed, "Modified U-Net for liver cancer segmentation from computed tomography images with a new class balancing method," BMC Biomedical Engineering, vol. 3, art. no. 4, 2021.
- [22] K. Yasaka, H. Akai, O. Abe, and S. Kiryu, "Deep learning with convolutional neural network for differentiation of liver masses at dynamic contrast-enhanced CT: A preliminary study," Radiology, vol. 286, no. 3, pp. 887–896, 2018.
- [23] M. Frid-Adar, I. Diamant, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, "GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification," Neurocomputing, vol. 321, pp. 321–331, 2018.
- [24] D. Liang et al., "Combining convolutional and recurrent neural networks for classification of focal liver lesions in multi-phase CT images," in Proc. MICCAI 2018, LNCS, vol. 11071, 2018, pp. 666–675.
- [25] C. A. Hamm et al., "Deep learning for liver tumor diagnosis part I: Development of a convolutional neural network classifier for multi-phasic MRI," European Radiology, vol. 29, no. 7, pp. 3338–3347, 2019.
- [26] Das, U. R. Acharya, S. S. Panda, and S. Sabut, "Deep learning based liver cancer detection using watershed transform and Gaussian mixture model techniques," Cognitive Systems Research, vol. 54, pp. 165–175, 2019.
- [27] R. M. Ghoniem, "A novel bio-inspired deep learning approach for liver cancer diagnosis," Information, vol. 11, no. 2, p. 80, 2020.
- [28] 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.
- [29] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in Proc. IEEE CVPR, 2017, pp. 4700–4708.
- [30] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in Proc. European Conf. Computer Vision (ECCV), LNCS, vol. 11211, 2018, pp. 3–19.
- [31] Dosovitskiy et al., "An image is worth 16x16 words: Transformers for image recognition at scale," in Proc. Int. Conf. Learning Representations (ICLR), 2021.

- [32] J. Chen et al., "TransUNet: Transformers make strong encoders for medical image segmentation," arXiv:2102.04306, 2021.
- [33] Hatamizadeh et al., "UNETR: Transformers for 3D medical image segmentation," in Proc. IEEE/CVF Winter Conf. Applications of Computer Vision (WACV), 2022, pp. 574–584.
- [34] 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.
- [35] 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.
- [36] G. Litjens et al., "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.
- [37] 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.
- [38] 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.
- [39] K. Alawneh et al., "LiverNet: Diagnosis of liver tumors in human CT images," *Applied Sciences*, vol. 12, no. 11, p. 5501, 2022.
- [40] 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.
- [41] 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.
- [42] K. Zuiderveld, "Contrast limited adaptive histogram equalization," in *Graphics Gems IV*, P. S. Heckbert, Ed. Academic Press, 1994, pp. 474–485.
- [43] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.
- [44] S. Beucher and F. Meyer, "The morphological approach to segmentation: The watershed transformation," in *Mathematical Morphology in Image Processing*, E. R. Dougherty, Ed. New York: Marcel Dekker, 1993, pp. 433–481.
- [45] P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Royal Statistical Society: Series B*, vol. 39, no. 1, pp. 1–38, 1977.
- [46] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York: Plenum Press, 1981.
- [47] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Advances in Engineering Software*, vol. 69, pp. 46–61, 2014.
- [48] W. Cai, S. Chen, and D. Zhang, "Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation," *Pattern Recognition*, vol. 40, no. 3, pp. 825–838, 2007.
- [49] R. Adams and L. Bischof, "Seeded region growing," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 16, no. 6, pp. 641–647, 1994.
- [50] R. A. Kirsch, "Computer determination of the constituent structure of biological images," *Computers and Biomedical Research*, vol. 4, no. 3, pp. 315–328, 1971.
- [51] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, 2002.
- [52] M. M. Galloway, "Texture analysis using gray level run lengths," *Computer Graphics and Image Processing*, vol. 4, no. 2, pp. 172–179, 1975.
- [53] K. Jain and F. Farrokhnia, "Unsupervised texture segmentation using Gabor filters," *Pattern Recognition*, vol. 24, no. 12, pp. 1167–1186, 1991.
- [54] F. Costa, G. Humpire-Mamani, and A. J. M. Traina, "An efficient algorithm for fractal analysis of textures," in Proc. 25th SIBGRAPI Conf. Graphics, Patterns and Images, 2012, pp. 39–46.
- [55] T. Jolliffe, *Principal Component Analysis*, 2nd ed. New York: Springer, 2002.
- [56] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Information Theory*, vol. 13, no. 1, pp. 21–27, 1967.
- [57] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, pp. 533–536, 1986.

- [58] T. G. Dietterich, "Ensemble methods in machine learning," in Proc. Int. Workshop Multiple Classifier Systems (MCS), LNCS, vol. 1857, 2000, pp. 1–15.
- [59] J. Fu et al., "Dual attention network for scene segmentation," in Proc. IEEE/CVF CVPR, 2019, pp. 3146–3154.
- [60] H. Zhang et al., "Co-occurrent features in semantic segmentation," in Proc. IEEE/CVF CVPR, 2019, pp. 548–557.
- [61] Y. Yuan, X. Chen, and J. Wang, "Object-contextual representations for semantic segmentation," in Proc. ECCV, 2020, pp. 173–190.
- [62] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in Proc. ECCV, 2018, pp. 801–818.
- [63] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in Proc. ECCV, 2016, pp. 630–645.
- [64] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in Proc. IEEE CVPR, Honolulu, HI, USA, 2017, pp. 1800–1807.
- [65] 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.
- [66] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in Proc. IEEE/CVF CVPR, 2018, pp. 4510–4520.
- [67] J. M. Kezmann, "TensorFlow Advanced Segmentation Models," GitHub repository. [Online]. Available: <https://github.com/JanMarcelKezmann/TensorFlow-Advanced-Segmentation-Models>
- [68] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in Proc. ICLR, 2015.
- [69] Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception architecture for computer vision," in Proc. IEEE CVPR, 2016, pp. 2818–2826.
- [70] 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.
- [71] Z. Liu et al., "A ConvNet for the 2020s," in Proc. IEEE/CVF CVPR, 2022, pp. 11976–11986.
- [72] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le, "Learning transferable architectures for scalable image recognition," in Proc. IEEE/CVF CVPR, 2018, pp. 8697–8710.
- [73] G. Howard et al., "MobileNets: Efficient convolutional neural networks for mobile vision applications," arXiv:1704.04861, 2017.
- [74] P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in Proc. ICLR, 2015.
- [75] L. R. Dice, "Measures of the amount of ecologic association between species," Ecology, vol. 26, no. 3, pp. 297–302, 1945.
- [76] P. Jaccard, "The distribution of the flora in the alpine zone," New Phytologist, vol. 11, no. 2, pp. 37–50, 1912.
- [77] 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.