SJIF Impact Factor (2023): 8.574| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 8 | Issue: 7 | July 2023 - Peer Reviewed Journal

AUTOMATED LIVER TUMOR SEGMENTATION USING DEEP TRANSFER LEARNING AND ATTENTION MECHANISMS

J. Jenisha[1], A. Joel Dickson[2]
[1*]*Communication Systems, Bethlahem Institute of Engineering, Anna University.*
[2]*Electronics and Communication Engineering, Bethlahem Institute of Engineering, Anna University, Chennai.*

***Corresponding Author**: J. Jenisha*

ABSTRACT
Accurate and efficient segmentation of liver tumors is essential for precise diagnosis, treatment planning, and monitoring of patients. To address these limitations, we proposed a novel framework called Deep Transfer Attention Network (DTAN) that integrates transfer learning and attention mechanisms for automated liver tumor segmentation. As a feature extractor, the DTAN model leverages a pre-trained convolutional neural network (CNN) to learn high-level representations from liver MRI images. To capture fine-grained spatial dependencies and emphasize tumor regions of interest, we introduce an attention mechanism that adaptively weights local features based on their relevance to the liver tumor segmentation task. We evaluate the model using metrics such as Hausdorff distance, Dice coefficient, specificity and sensitivity. The combination of transfer learning and attention mechanisms enables the extraction of discriminative features and enhanced understanding of spatial context, leading to more accurate and reliable liver tumor segmentation results. The proposed framework holds significant potential in supporting radiologists and clinicians in making timely and informed decisions for liver tumor diagnosis, treatment planning, and patient management.
KEYWORDS: *CNN, DTAN, Liver Tumor Segmentation, transfer learning, attention mechanis.*

1. INTRODUCTION

With liver cancer, liver tumorposes a significant health challenge worldwidebeing the fourth important cause of cancer andthe sixth most mutual cancerrelated deaths globally. Accurate and efficient segmentation of liver tumorsshows a crucial partin precise diagnosis, monitoring of patients and treatment planning [1]. Medical imaging techniques, such as magnetic resonance imaging (MRI), have proven instrumental invisualizing liver tumors and providing vital information for clinicians.

Deep learning models have demonstrated remarkable success in severalfields, including medical image analysis. These models can automatically learn and extract meaningful representations from large-scale datasets, enabling them to capture intricate patterns and make accurate predictions [2]. However, when it comes to liver tumorsegmentation, existing approaches encounter difficulties in capturing the intricate features and spatial relationships across different tumor regions within the liver [3].

Capturing the intricate features within liver tumors is crucial because tumors can exhibit heterogeneous characteristics. For instance, tumorsmay contain areas with varying degrees of malignancy, different tumorsubtypes, or regions of necrosis [4]. Accurate segmentation of these regions is essential for guiding treatment decisions, such as surgical resection or radiation therapy, and assessing treatment response over time.

Moreover, spatial relationships between different tumor regions play a critical role in accurately delineating tumor boundaries. Tumors can have irregular shapes, invade adjacentstructures, or exhibit complex internal structures [5]. Existing approaches often struggle to capture these spatial dependencies, leading to incomplete or inaccurate tumorsegmentations.

The proposed Deep Transfer Attention Network (DTAN) for automated liver tumor segmentation makes several significant contributions:

- The DTAN model leverages transfer learning by using a pre-trained CNN as a feature extractor, enabling the learning of high-level representations from liver MRI images and capturing relevant features for accurate liver tumor segmentation, even with limited training data.
- The DTAN model integrates an attention mechanism to adaptively weigh local features, allowing for fine-grained spatial dependencies and accurate segmentation of liver tumors by emphasizing relevant tumor regions.

The paper's remaining sections are as follows: Section 2 introduces the proposed DTAN model, integrating deep learning, attention mechanisms, and transfer learning for advanced liver tumor diagnosis. Section 3 details the utilization of transfer learning and attention mechanisms in the DTAN model for feature extraction and spatial dependency capture from liver MRI images. Section 4 presents the evaluation of the DTAN model's performance using metrics such as Dice coefficient and Hausdorff distance, confirming the effectiveness of transfer learning and attention mechanisms in extracting discriminative features. Finally, Section 5 concludes by emphasizing the potential of the DTAN framework to assist radiologists and clinicians in accurate liver tumor segmentation for informed decision-making.

2. RELATED WORKS

Some of the papers based on the Liver tumor segmentation are reviewed below,
Amin *et al*. [6] proposed method involves enhancing image quality using a local Laplacian filter and utilizing a semantic segmentation model with features extracted from a pre-trained VGG16 model and passed through a U-shaped network.

Li *et al*. [7] selectedUNet++ structure as the baseline, incorporating a context-aware feature encoder to enhance deep network degradation and an efficient attention module to effectively combine spatial information for improved feature map depth. Furthermore, with Dice Loss the cross-entropy loss function was replaced by optimizing parameters of network.

Balasubramanian *et al*. [8] suggested a novel deep learning model for segmentation and classification of liver tumor, involving a three-stage process: pre-processing of CT images, liver segmentation using Mask R-CNN, and tumor classification, including contrast improvement, noise reduction, and liver separation using the Mask R-CNN model on CT abdominal images.

Tummala and Barpanda [9] proposed a curriculum learning strategy to develop an overcomplete U-Net model for liver tumor segmentation, addressing challenges posed by variations in tumor characteristics, and enabling accurate learning of tumor artifacts for improved diagnosis and therapy planning in liver cancer treatment.

Hansch*et al*. [10]proposed a deep learning-based approach utilizingthe late hepatocellular phase of DCE-MR in a multi-model training strategy andan anisotropic 3D U-Net architecture was proposed for segmentation of liver tumor.

Rahman *et al*. [11] proposed a more capable method for liver and tumor segmentation in CT image volumes by combining ResNet and UNet models into a hybrid ResUNet model, addressing the gap in liver segmentation and performing region of interest assessment using a publicly available 3D dataset IRCADB01.

Zhang *et al*. [12] proposed a Decoupled Pyramid Correlation Network (DPC-Net), incorporating attention mechanisms to exploit high- and low-level features in FCN. The network utilized SemCor modules to enhance spatial details and SpaCor modules to emphasize global semantic information.

Sahli*et al*. [13] proposed a new analytical detection method for liver tumor segmentation in CT images was proposed, utilizing an encoder-decoder structure and automatic algorithms based on architectures of Seg-Net and U-Net to improve the localization and breakdown of metastatic lesions in medical imaging.

Zheng *et al*. [14] proposed, a 4D DL model based on convolutional long short-term memory (C-LSTM) and 3D convolutional for hepato-cellular carcinoma (HCC) lesion segmentation, utilizing dynamic contrast-enhanced (DCE) MRI images through 4D information to extract spatial and temporal domain features.

Amin *et al*. [15] offered, a three-part model comprising synthetic image generation, segmentation and localization was employed, utilizing an optimized generative adversarial network for synthetic image generation and an improved localization model incorporating YOLOv3 detector with deep features from pretrained ResNet-50 models removed.

SJIF Impact Factor (2023): 8.574| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 8 | Issue: 7 | July 2023 - Peer Reviewed Journal

Table 1: Comparative analysis of the existing methods on liver tumour segmentation

Authors	Methods Used	Advantages	Disadvantages
Amin *et al.* [6]	VGG16	-improves image quality and the accuracy.	-potentially limiting its applicability
Li *et al.* [7]	UNet++	-improves the accuracy and robustness of the Network.	-Its performance may vary depending on the specific dataset and task.
Balasubramanian *et al.* [8]	APEST Net	-provides accurate liver tumor segmentation and classification.	-Influenced by the quality and variability of the input CT images.
Tummala and Barmanda[9]	U-Net	-capturing high-level structures effectively.	- performance and generalizability of the proposed approach on the data used.
Hansch*et al.*[10]	3D U-Net	-improves the accuracy and efficiency	- influenced by factors such as dataset variability, imaging artifacts.
Rahman *et al*.[11]	ResUNet	-improved efficiency and accuracy from CT image volumes.	-influenced by factors such as variations in tumor characteristics, image quality.
Zhang *et al.* [12]	DPC-Net	-effectively leverages both low- and high-level information.	-may increase computational requirements.
Sahli*et al.* [13]	U-Net	-aiding in diagnostic andtreatment effectiveness.	- applicability to different datasets.
Zheng *et al.*[14]	C-LSTM	-It combines spatial and temporalinformation.	-increase computational fortraining and inference.
Amin *et al.* [15]	U-Net	-improved localization, allowing for more accurate.	-limit the model's performance in handling specific variations.

3. PROPOSED DEEP TRANSFER ATTENTION NETWORK MODEL

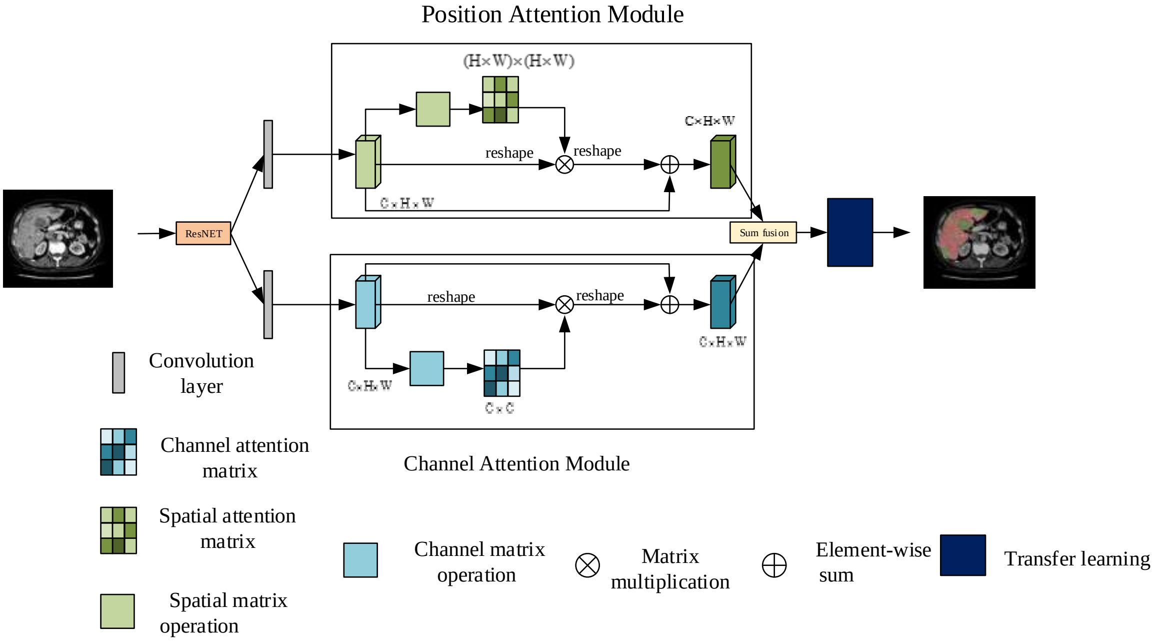


Figure 1: Structure of the Proposed Deep Transfer Attention Network Model

We proposed a novel framework called Deep Transfer Attention Network (DTAN) for automated liver tumor segmentation in fig 1. The DTAN model integrates transfer learning and attention mechanisms, leveraging a pre-trained CNN for feature extraction and an attention mechanism to emphasize relevant regions, enabling accurate segmentation by capturing fine-grained spatial dependencies and intricate features.

3.1. Dataset Preparation

A comprehensive dataset of liver MRI images, including both tumor and non-tumor cases.Preprocess the dataset by standardizing the image intensities, resizing the images to a consistent resolution, and dividing them into training and testing sets.

3.2. Pre-training the Deep Transfer Attention Network (DTAN)

Utilize a pre-trained DTAN, ResNet, which has been trained on a large-scale dataset (e.g., ImageNet).Remove the fully connected layers from the DTAN, keeping the convolutional layers intact.Freeze the weights of the pre-trained DTAN to prevent them from being modified during training.Pre-training DTAN refers to the process of training the network on a large dataset or a related task before fine-tuning it for a specific target task. This pre-training step helps initialize the DTAN's weights and learn general features that can be useful for various computer vision tasks.The main advantage of pre-training a DTAN is that it enables transfer learning, where knowledge gained from one task or dataset is transferred to another task. By leveraging pre-trained weights, the DTAN can benefit from general visual representations learned from a large dataset, even when the target dataset is limited. This can lead to faster convergence, improved generalization, and better performance, especially when the target dataset is small or lacks diversity.Overall, pre-training a DTAN provides a powerful strategy to initialize and leverage convolutional neural networks for various computer visiontasks, reducing the need for training large models from scratch and enabling effective transfer of learned knowledge.

ResNet

A pretrained Residual network with the dilated strategy is employed as the backbone. Residual Network (ResNet) is a deep learning architecture that has been widely used in various computer vision tasks, including image segmentation. Liver tumor segmentation is one such task where ResNet has been applied successfully. The goal of liver tumor segmentation is to identify and delineate the boundaries of tumors within liver images. This task is crucial in diagnosis of medical image, treatment planning, and monitoring the progress of liver diseases. Deep learning techniques, such as ResNet, have shown promising results in automating this process. By utilizing residual connections, ResNet is known for its ability to train very deep neural networks. These connections allow the network to absorb residual mappings, permitting the network to capture fine-grained details effectively. This is especially important for tasks like liver tumor segmentation, where accurate boundary delineation is essential. Liver tumor segmentation is a complex task, and the ResNet architecture is just one of many possible approaches. Researchers continually explore and develop new methods, combining various architectural components and techniques, to further improve the efficiencyand accuracyof segmentation of liver tumor.

Residual Block

The development of ResNet, also known as residual networks. These ResNets are constructed using Residual Blocks, which have played a pivotal role in overcoming this challenge.
The mathematical equation of Residual block is as shown in equation (1),

$$H(y) = F(y) + y \tag{1}$$

Rearranging the Equation (1), we get

$$F(y) = H(y) - y \tag{2}$$

In the Residual block, the output of the considerable layers denoted as H(x), is obtained based on the input vectors x within a residual block. To learn the key objective of the residual block is the true output $H(y)$ as shown in Eq 1 and fig. 1 $F(y)+y$ by utilizing a feedforward neural network that incorporates identity shortcut connections. These shortcut connections are responsible for maintaining computational complexity and occur due to the input. To express this mathematically we can refer to Eq2, where $F(y)$ represents the residual function

The function $F(y)$ is often represented by matrix multiplication interlaced with activation functions and normalization operations.

Convolutional Layers (CL)

In the Deep Transfer Attention Network (DTAN), the last two ResNet blocks undergo modifications to extend the final feature map dimension and increase the dimension reduction. The CL within the DTAN module plays a crucial role in automated tumor detection, also segmentation within liver medical images. It applies learnable filters to input image patches, extracting relevant features such as edges, textures, and tumor-specific patterns. The CL learns these features through training and helps highlight tumor regions, supporting diagnosis, treatment planning, and patient monitoring for liver cancer.

SJIF Impact Factor (2023): 8.574| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 8 | Issue: 7 | July 2023 - Peer Reviewed Journal

Calculating the spatial dimensionality of the convolution layer's output, we tend to use the following formula as,

$$\frac{(R-V)+2P}{S+1} \quad (3)$$

where V denoted as the input volume size, (i.e., height × breadth × depth),
R denoted as the receptive field size
P is the amount of zero padding set and
S denoted as the stride

Position and Channel Attention Module (PAM & CAM)

The features are provided into two, parallel attention modeland channel attention model. To capture spatial relationships, features are processed by parallel spatial and channel attention modules, generating a spatial attention matrix, performing matrix multiplication with the original features, and obtaining final representations through element-wise summation.

Position Attention Model (PAM)

To capture rich contextual relationships, position attention module is introduce, that enhances local features by coding a broader series of contextual information and adaptively aggregating spatial contexts. Assumed a limitedfeature $\boldsymbol{L} \in \Re^{C \times H \times W}$ we first fed it into a CL to generate $\boldsymbol{M}$ as well as $\boldsymbol{K}$ are two new feature maps, where $\{R, L\} \in \Re^{C \times H \times W}$. At that time we reshape them to $\Re^{C \times N}$, here $N = W \times H$ represents the totalnumber of pixels. After that a multiplication of matrixis performed among the transpose of $\boldsymbol{K}$ as well as $\boldsymbol{M}$.

$p \in \Re^{N \times N}$:

$$p_{ij} = \frac{\exp(M_j \cdot K_i)}{\sum_{j=1}^{N} \exp(M_j \cdot K_i)} \quad (4)$$

where p_{ij} measures the j^{th} position's impact on i^{th} position. The most similar feature representations of thoseposition contributes to greater correlation among them.

while, we fed feature $\boldsymbol{L}$ into a CL to create a $\mathbf{D} \in \Re^{C \times H \times W}$ as new feature map and reshape it to $\Re^{C \times N}$. And multiplication of matrix between $\mathbf{D}$ and the transpose of S is performed and the result is reshape to $\Re^{C \times H \times W}$.Then, a scale parameter α is multipliedand an element-wise sum operation is performed with the features $\boldsymbol{L}$ andthe final output $\mathbf{E} \in \Re^{C \times H \times W}$ is obtained:

$$F_i = \alpha \sum_{j=1}^{N} (s_{ij} D_j) + A_i \quad (5)$$

Channel Attention Model (CAM)

From the PAM,the CAM is altered and directly the map of channel attention $K \in \Re^{\mathbf{C} \times \mathbf{C}}$ from the original features $\mathbf{L} \in \Re^{C \times H \times W}$ is calculated. Specifically, we reshape $\mathbf{L}$ to $\Re^{C \times N}$, and then a matrix multiplication between $\mathbf{L}$ and the transpose of $\mathbf{L}$ is performed.

$K \in \Re^{C \times C}$:

$$Q_{ij} = \frac{\exp(M_i \cdot M_j)}{\sum_{j=1}^{C} \exp(M_i \cdot M_j)} \quad (6)$$

where k_{ji} measures the i^{th} channel's impact on the j^{th} channel. Moreover, a multiplication of matrix among the transpose of K and $\mathbf{L}$ is performed and then reshape their result to $\Re^{\mathbf{C} \times \mathbf{H} \times \mathbf{W}}$ finally the result by a scale parameter β and implement an element-wise sum operation is performedwith $\mathbf{L}$ to obtain the final output $\mathbf{E} \in \Re^{C \times H \times W}$:

SJIF Impact Factor (2023): 8.574| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 8 | Issue: 7 | July 2023 - Peer Reviewed Journal

$$F_i = \beta \sum_{j=1}^{K} (x_{ji} M_i) + M_j \quad (7)$$

Attention Mechanism (AM)

The AM in the DTAN model selectively weights and emphasizes important regions or features in the input data, enhancing the model's ability to capture fine-grained spatial dependencies and emphasize tumor regions. It dynamically assigns weights based on relevance, guiding the model's focus during liver tumor segmentation. The attention mechanism consists of an Encoder module that transforms the input data into a semantic vector and a Decoder module that generates the output data based on the transformed vector. The attention mechanism's equation defines the weight assignment process.

$$u_j = \tanh(W_j h_j + b_j) \quad (8)$$

$$\alpha_j = \frac{\exp(u_j^T u_n)}{\sum_j \exp(u_j^T u_n)} \quad (9)$$

$$h_j = \sum_j \alpha_j h_j \quad (10)$$

where: α_j is the attention score for $j-th$ word sentence; u_i is the result of a full connection operation of the hidden layer vector; h_j ;W_j and b_j are the height, weight matrix and bias term of attention calculation; u_n is a randomly improved context vector, which is efficient.

Sum Fusion

To leverage contextual information, we combine features from two attention modules by transforming their outputs with a CL and performing an element-wise sum fusion. A final CL generates the prediction map, while avoiding memory-intensive cascading operations. Our straightforward attention modules seamlessly integrate into the FCN pipeline, enhancing feature representations without a significant increase in parameters.

Transfer Learning

Transfer learning in DTAN leverages a pre-trained CNN, such as ImageNet, to improve feature extraction, overcome limited data challenges, and enhance generalization in liver tumor segmentation. The pre-trained model captures generic visual features applicable to various tasks. Transfer learning addresses data limitations and aids faster convergence, preventing overfitting. By extracting discriminative features, DTAN achieves accurate liver tumor segmentation, while adaptability and reduced complexity are ensured by removing fully connected layers from the pre-trained model.

4. EXPERIMENTAL RESULTS

The DTAN model is evaluated on a labelled training dataset and its segmentation performance is optimized. The trained model is then tested on an independent dataset using metrics such as Hausdorff distance, Dice coefficient, specificity and sensitivity.

3.6.1. Hausdorff distance

$$H(A,B) = \max(h(A,B), h(B,A))$$

The Hausdorff distance quantifies dissimilarity between sets/shapes, reflecting their shape and location. It is used in image processing and computer vision for tasks like shape matching and evaluation, aiding segmentation accuracy and image alignment.

3.6.2. Dice coefficient

$$(2*|A \cap B|)/(|A|+|B|)$$

The Dice coefficient measures the overlap between predicted and ground truth masks in image segmentation, indicating segmentation accuracy. A higher coefficient signifies better performance, commonly used in medical image analysis for tasks like tumor segmentation.

4.1 Comparative Analysis

Table 2 shows the performance of proposed DTAN model with existing methods such as DPC-Net [12], and C-LSTM [14]. The proposed DTAN model achieved the better segmentation performance than other methods for liver tumor segmentation. Our method achieves dice value of 82.11% and Hausdroff Distance value of 13.78%.

SJIF Impact Factor (2023): 8.574| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 8 | Issue: 7 | July 2023 - Peer Reviewed Journal

Table: 2 Dice and Hausdroff Distance for proposed and existing models

Methods	Dice (%)	Hausdroff Distance (%)
DPC-Net [12]	76.4	5.339
C-LSTM [14]	80.9	12.76
Proposed DTAN model	82.11	13.78

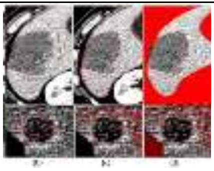
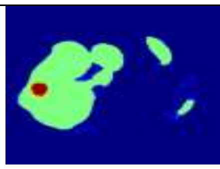
Input image	ResNet	Position attention Model	Channel attention Model	Output Image
				

Figure: 2 Segmentation Results

5. CONCLUSION

In this paper weintroduces the Deep Transfer Attention Network (DTAN), which combines transfer learning and attention mechanisms to achieve state-of-the-art liver tumor segmentation performance. By capturing intricate features and spatial relationships, DTAN provides accurate results, holding promise for aiding radiologists and clinicians in liver tumor diagnosis, treatment planning, and patient management, enhancing precision and efficiency in clinical settings.

REFERENCES

1. *Rela M, Suryakari NR, Reddy PR (2020) Liver tumorsegmentation and classification: A systematic review. 2020 IEEE-HYDCON, pp.1-6.*
2. *Mohite MVB, Belagali MP. Methods for Liver TumorSegmentation and Detection: A Review.*
3. *Jagtap J, Habeeb A, Jha A, Aggarwal S, Gupta K (2021) Bibliometric Review on Liver and Tumour Segmentation using Deep Learning.*
4. *Kesaratti M, Dixit S, Khodanpur BI. A Survey on Different Methods for Liver Segmentation.*
5. *Napte K, Mahajan A (2022) Deep Learning based Liver Segmentation: A Review. Revue d'IntelligenceArtificielle, 36(6):979.*
6. *Amin J, Anjum MA, Sharif M, Kadry S, Crespo RG (2023) Visual Geometry Group based on U-Shaped Model for Liver/Liver Tumor Segmentation. IEEE Latin America Transactions, 21(4):557-564.*
7. *Li J, Liu K, Hu Y, Zhang H, Heidari AA, Chen H, Zhang W, Algarni AD, Elmannai H (2023) Eres-UNet++: Liver CT image segmentation based on high-efficiency channel attention and Res-UNet++. Computers in Biology and Medicine, p.106501.*
8. *Balasubramanian PK, Lai WC, Seng GH, Selvaraj J (2023) APESTNet with Mask R-CNN for Liver Tumor Segmentation and Classification. Cancers, 15(2):330.*
9. *Tummala BM, Barpanda SS (2023) Curriculum learning based overcomplete U-Net for liver tumor segmentation from computed tomography images. Bulletin of Electrical Engineering and Informatics, 12(3):1620-1629.*
10. *Hänsch A, Chlebus G, Meine H, Thielke F, Kock F, Paulus T, Abolmaali N, Schenk A (2022) Improving automatic liver tumor segmentation in late-phase MRI using multi-model training and 3D convolutional neural networks. Scientific Reports, 12(1):1-10.*
11. *Rahman H, Bukht TFN, Imran A, Tariq J, Tu S, Alzahrani A (2022) A Deep Learning Approach for Liver and TumorSegmentation in CT Images UsingResUNet. Bioengineering, 9(8):368.*
12. *Zhang Y, Yang J, Liu Y, Tian J, Wang S, Zhong C, Shi Z, Zhang Y, He Z (2022) Decoupled pyramid correlation network for liver tumor segmentation from CT images. Medical Physics, 49(11):7207-7221.*
13. *Sahli H, Ben Slama A, Labidi S (2022) U-Net: A valuable encoder-decoder architecture for liver tumors segmentation in CT images. Journal of X-Ray Science and Technology, 30(1):45-56.*
14. *Zheng R, Wang Q, Lv S, Li C, Wang C, Chen W, Wang H (2022) Automatic liver tumor segmentation on dynamic contrast enhanced mri using 4D information: deep learning model based on 3D convolution and convolutional lstm. IEEE Transactions onMedical Imaging, 41(10):2965-2976.*
15. *Amin J, Anjum MA, Sharif M, Kadry S, Nadeem A, Ahmad SF (2022) Liver tumor localization based on YOLOv3 and 3D-semantic segmentation using deep neural networks. Diagnostics, 12(4):823.*