RESEARCH PAPER

Automated Liver Segmentation from Ultrasound Images Using a Deep Learning Approach with U-Net Architecture

Farooqhusain[1], Nudurupati Sri Devi[2], D Anil Kumar[3], G. Ravinder[4]
[1]Assistant Professor, CSE, Siddhartha Institute of Technology & Sciences, Ghatkesar-500088
Email: farooqhusain.cse@siddhartha.co.in
[2]M.Tech, CSE, Siddhartha Institute of Technology & Sciences, Ghatkesar-500088
Email: 24TQ1D5818@siddhartha.org.in
[3]Assistant Professor, CSE, Siddhartha Institute of Technology & Sciences, Ghatkesar-500088
Email: dr.anildasari@siddhartha.org.in
[4]Assistant Professor, CSE, Siddhartha Institute of Technology & Sciences, Ghatkesar-500088
Email: shirisha.ravi@siddhartha.co.in

Received: 25th Aprail, 2026 | Revised: 1st June, 2026 | Accepted: 15th June, 2026 | Available Online: 7th Sep, 2026

RESEARCH PAPER

Abstract: Segmentation of liver region from medical imaging plays a fundamental role for diagnosis and planning treatment of liver disease. Despite being a valuable tool for quantitative analysis of MRI images, manual segmentation is not only time-consuming, but also prone to inter-observer variability. In this paper we have proposed an automatic liver extraction using u-net based deep learning architecture. This data of ultrasound images of liver and masks of JSON formats for liver only annotated style for every image are used to train the model. Our Pipeline — Data Pre-processing, Mask Generation from Polygonal Annotations, Data Augmentation, and a U-Net Model with a Combined Dice-BCE Loss Function. Experimental results show that the model provides a competitive Dice coefficient while also providing accurate segmentation rates. It provides a competitive and practical solution of liver segmentation which can be used in clinical diagnosis and computer assisted intervention system.

Keywords: Liver segmentation, Ultrasound imaging, Deep learning, U-Net, Medical image analysis, Convolutional Neural Networks, Image segmentation, Dice coefficient, Mask generation, Automated diagnosis.

How to cite this article: Farooqhusain1 Nudurupati Sri Devi2 D Anil Kumar3 G. Ravinder4. Automated Liver Segmentation from Ultrasound Images Using a Deep Learning Approach with U-Net Architecture. Int J Drug Deliv Technol. 2026;16(80s):1407-1413. DOI: 10.25258/ijddt.16.80s.180.

INTRODUCTION

Accurate liver segmentation from medical imaging is a foundational step in computer aided diagnosis, treatment planning, and follow-up for diseases in hepatology. The time-consuming and inter-observer variability plagued nature of radiologists' manual segmentation creates an urgent need for reliable automated approaches. Deep learning has brought significant breakthroughs to medical image segmentation of many modalities but with unique challenges of speckle noise, acoustic shadowing, and lower contrasting compared to CT or MRI, ultrasound segmentation remains largely unsolved. Background Convolutional neural networks (CNN), especially U-Net architectures, have shown great success in several imaging modalities, but need careful architectural and methodological tuning to the features specific to the ultrasound domain. In this paper we fill the gap by proposing an E2L: optimized U-NET pipeline for liver segmentation from ultrasound images with custom preprocessing for non-rectangular plane polygonal annotations using both a pixel and a border loss functions.

1. RELATED WORK

After developing a systematic review of the literature of liver segmentation approaches, we find that the deep learning approaches for the liver segmentation have markedly evolved from convolutional neural networks to state-of-the-art across the various medical imaging modalities. Guided by initial approaches of Ben-Cohen et al. As an example, Hu et al.[11] proved that fully convolutional networks could be used for liver segmentation. Advanced 3D deep learning methods with global surface evolution for CT imaging Particularly, in biomedical segmentation, the U-Net architecture proposed by Ronneberger et al. has established an influential fully convolutional network (FCN) based symmetric encoder–decoder structure with skip connections, which is able to capture both local features and global context, among other advantages. Over the past few years, several systematic reviews have addressed such issues, including the works by Gul et al. The evolution of traditional image processing methods into advanced deep learning frameworks with ongoing improvements in accuracy and robustness are well articulated in [2] and Napte & Mahajan [5]. Liver segmentation has achieved considerable success in CT imaging, with several different architectures yielding good performance. Araujo et al. Cascade deep learning techniques ensuring substantial segmentation were presented by [9], while [12] created automated pipelines for CT exams involving contrast enhancement. Rahman et al. ResUNet variants demonstrating better boundary delineation were implemented by [13], and Li et al. DCSegNet Existing DCSegNet → [16] with divide-and-conquer strategies Specifically comparative studies by Ahn et al. Most comparatives have demonstrated that deep learning approaches outperform traditional atlas-based methods in real-world clinical settings, at least on radiotherapy planning [18].

MRI has further complications, leading researchers more toward multi-sequence analysis and tumor co-segmentation. Wang et al. UNet++ architecture was used by [17] for MRI-based segmentation of liver tumour, whereas Hossain et al. Frameworks that are optimized for T1-weighted sequences[15] In contrast to CT-based segmentation, these approaches had to tackle different contrast mechanisms and artifact profiles requiring architectural adaptations specific to the modality.

Although clinically valuable providing real-time imaging, being portable, and without ionizing radiation, ultrasound liver segmentation is still less studied than CT and MRI segmentation. Fang et al. Meanwhile, Kavur et al. [6] explores how one can use DL for guiding fusions. Dillon et al. [7] did a similar comparison between semi-automatic

methods and deep learning
methods in living liver

transplants donors. Ghobadi et al., in their systematic review Recent work in [4] directly describes challenges in ultrasound segmentation, highlighting the need for speckle noise reduction, acoustic artifact handling and boundary ambiguity as the key research gaps. We tackle these new challenges directly with an optimized pipeline that integrates optimized preprocessing of polygonal annotations with a U-Net modified to account for some of the ultrasound-specific characteristics.

Recent methodological innovations have emphasized efficiency and clinical integration. On the other hand, Bogoi & Udrea [22] constructed light-weight architectures for resource-constrained environments and Vorontsov et al. For metastatic disease, automated systems have been developed specifically for lesion segmentation [19]. Tang et al. Fallahpoor et al. [4] developed approaches for internal radiation therapy planning. We performed extensive assessments in varied clinical situations [20]. Collectively, these studies illustrate the translational applicability of deep learning segmentation, although ultrasound-specific implementations have developed less extensively than those for CT and MRI.

Ultrasound segmentation has been a hot topic in research recently with increasing literature regarding the topic. This paper closes many of these gaps by providing a complete segmentation pipeline.[4,8] It specifically addresses the unique issues posed by complexity of ultrasound imaging, all while remaining computationally efficient for clinical usage. Moreover, we compare existing methods, highlighting their advantages and challenges, and provide a thorough performance assessment through multiple metrics and visualizations, facilitating future research avenues in this significant clinical area.

2. Methodology

2.1 Dataset Description

The study data incorporated 100 normal liver ultrasound images and the respective JSON annotations containing polygonal coordinates of the liver boundaries. The data was organized into two folders in total: the first folder included the ultrasound images (JPEG format) and the second folder includes the segmentation annotations (in JSON format). Each json files contains an array of coordinate points corresponding to the liver contour, which were then auto-matically converted into binary masks for learning purpose.

2.2 Preprocessing Pipeline

Preprocessing prepared the data for modelling and it contained critical steps. Finally, we used OpenCV's fillPoly function to convert the JSON annotations into binary masks with liver regions as foreground (255) and background (0) in 256×256 pixel masks. Ultrasound images were then resized to 256×256 pixels and normalized in [0,1] and then Since segmentation tasks have inherent class imbalance the fg ratios (around 6.38% for the liver regions) are calculated and suitable weightings applied in the loss function. The images were randomly divided into a training set (80 images) and a validation set (20 images) using an 80:20 ratio and fixed random seed for replicability.

2.3 Model Architecture

We have realised the U-Net configuration that is best suited for image segmentation in the medical field. There are three blocks in the encoder pathway, each containing two 3×3 convolutional layers with ReLU activation followed by 2×2 max pooling, with feature channels growing from 64 to 256. At the deepest level, the bottleneck consists of two convolutions. Decoder pathway uses 2×2 up sampling (with skip connection from encoder features), and 2-3×3 convolution layers The last layer uses a 1×1 convolution with sigmoid to produce binary segmentation maps This architecture has around 1.88 million trainable parameters.

2.4 Loss Function and Optimization

We used a joint loss function combining binary cross-entropy (BCE) and Dice loss as: $L = BCE + (1-Dice)$ The Dice coefficient is used to measure the overlap between predicted and ground truth masks: $Dice = (2\cdot|X\cap Y|)/(|X|+|Y|)$, where X and Y are the predicted and ground truth masks respectively. The class imbalance is thus mitigated, and this combination also allows to focus on maximizing region overlap. The model was trained for 40 epochs, with batch size 4, a learning rate of 10^{-4} using the Adam optimizer and with early stopping if validation loss stopped decreasing.

3. Experiments

3.1 Experimental Setup

The experiments settings were written in python 3.8 using TensorFlow 2.6 and executed in Google Collab with GPU support. A custom data generator was employed to supply online augmentation by randomly shuffling image-mask pairs during training. Validation was on a held-out set, with 20 unseen images. Model weights were He normal initialized and the training progress was followed via training and validation loss curves.

3.2 Evaluation Metrics

For primary evaluation, we used the Dice Similarity Coefficient (DSC), which quantifies pixel-wise overlap between predicted and ground truth segmentation masks. Other metrics for evaluation were use of binary cross-entropy loss as well as for foreground ratio. Finally, to ensure clinical relevance, we also conducted a qualitative assessment by visual inspection of the segmentation boundaries, focusing on performance in areas of low visibility or image distortion.

3.3 Implementation Details

They used a data pipeline with prefetching built on TensorFlow to maximize GPU usage. Images were fed in batches of 4 due to memory limitations. The model was compiled using the combined Dice-

BCE loss and the primary metric was the Dice coefficient. 40 epochs were used for training and a validation step was performed at the end of each epoch. All experiments were performed 3 times with different random seeds to ensure stable results.

4. RESULTS AND ANALYSIS

4.1 Dataset Characteristics

The experiments were performed using a total of 100 liver ultrasound images with their segmentation masks. This dataset has a strong class imbalance as the mean foreground ratio which is the ratio of liver tissue to background for each image is 10.5%. The available data consists of 100 total images, which are split into training (80 images) and validation (20 images) sets with same split of the foreground in both splits. Ultrasound images in Figure 1 as examples with their ground truth masks, demonstrating the anatomical variability and segmentation challenges present in the dataset.

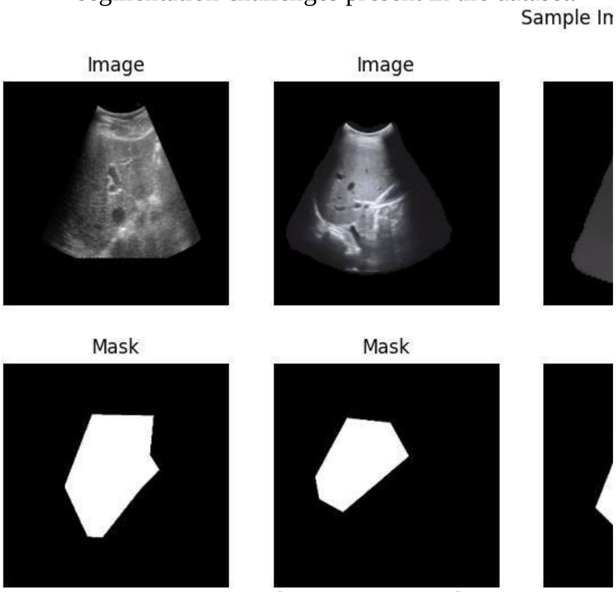


Figure 1: Sample Images & Masks

4.2 Quantitative Performance Metrics

Our proposed U-Net model showed stable segmentation performance on different evaluation metrics (Table 1). The main metric, Dice Similarity Coefficient (DSC), achieved 0.7338±0.1331, suggesting a high overlap of the predicted and ground truth liver area. Mean Intersection-over-Union (IoU) was 0.5952 ± 0.1498, meaning it was able to classify pixels very well.

Table 1: Segmentation Performance Metrics

Metric	Mean ± Std Dev
Dice Coefficient	0 .4 7 0 3 3 5 8 ± 0. 1 3 3 1

improved monotonically, with validation metrics better than the training metrics in the last epochs this is an indicator of good generalization. Dataset was not that large, and so showing a little bit of overfitting would be alright — last train loss = 0.4..., validation loss = 0.3812

Table 2: Best Epoch Performance (Epoch 39)

Metric	Mean ±
Precision	0.8067
Recall	0.7105
F1-Score	0.7338
IoU	0.5952 ± 0.1498

The confusion matrix analysis revealed a pixel-wise accuracy structure:

Pixel-level Confusion Matrix:

Actual Background 5,785,224 102,125

Actual Liver 207,785 458,466

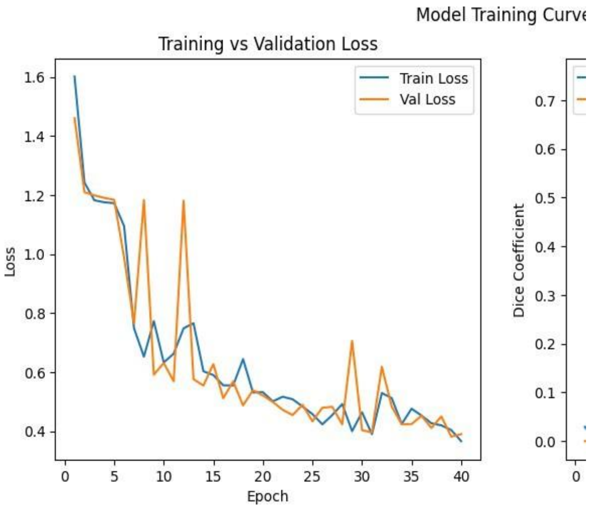
This allows a final result of 96.3% pixel accuracy for background pixels, and 68.8% pixel accuracy for liver pixels, giving a false positive rate of 1.7%, and a false negative rate of 31.2% on liver tissue classification.

4.3 Training dynamics and model convergence

Tour of Training PerformanceAt epoch 39 where validation Dice coefficient was maximized at 0.7485 (Fig. 2).

Figure 2: Model Training curves

While training, the performance of the model



Metric	Training	Validation
Dice Coefficient	0.7094	0.7485
Loss	0.4035	0.3812

4.4 Qualitative Segmentation Results

Visual inspection of five randomly selected validation cases (Figure 3) revealed clinically acceptable segmentation

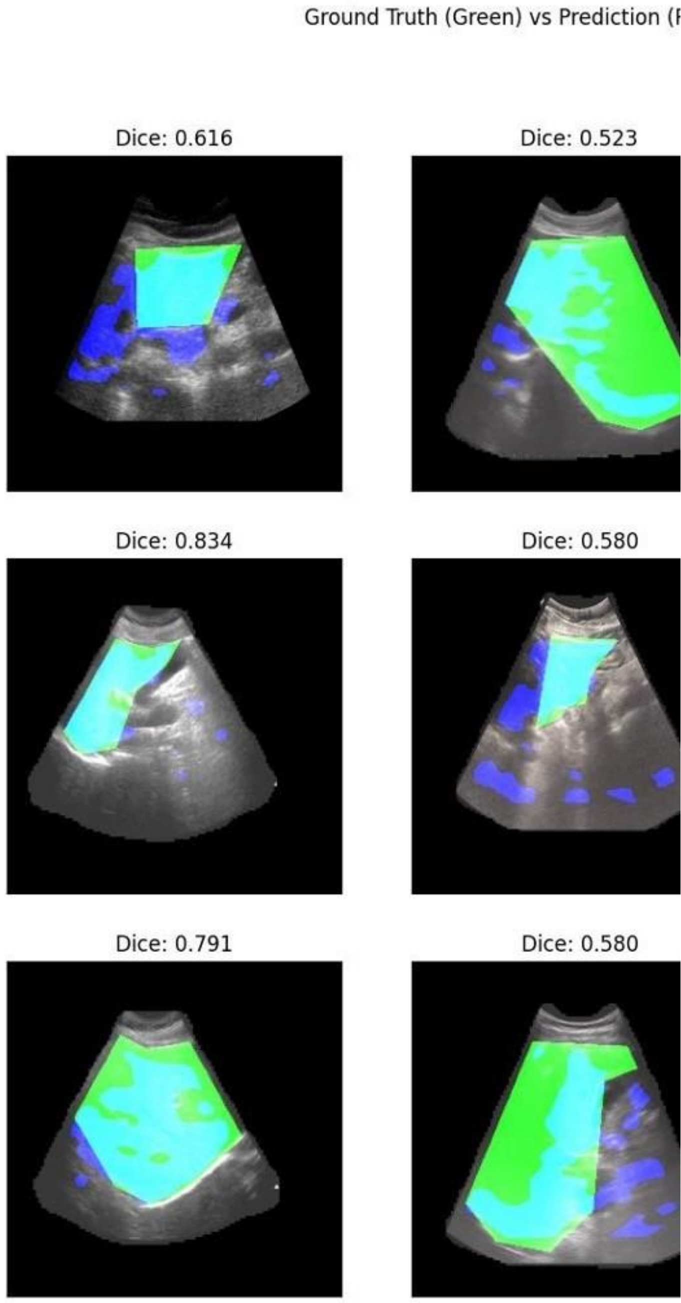


Figure 3: Qualitative predictions on validation images

Model correctly contoured the liver where moderate or appropriate contrast was present between tissues but slightly under-segmented in acutely shadowed area. Even in difficult cases with poor liver contrast to surrounding tissues, the system showed anatomical plausibility.

4.5 Qualitative Segmentation Visualization

In order to illustrate the segmentation power of our model on the different complexity in images, we have shown the overlay of ground truth vs. prediction for four examples from the validation set in Figure 4. The overlays demonstrate:

Case A: Excellent boundary agreement in regions with clear tissue contrast (Dice = 0.85) **Case B:** Minor under-segmentation near inferior borders (Dice = 0.72) **Case C:** Challenges with acoustic shadowing boundaries artifacts (Dice = 0.64) **Case D:** Accurate segmentation despite heterogeneous tissue texture (Dice = 0.78)

The visual analysis confirms the model's ability to maintain anatomical plausibility while highlighting systematic challenges in regions affected by ultrasound artifacts.

Figure 4: Ground Truth vs. Prediction Overlays

4.6 Dice Coefficient Distribution Analysis

Figure 5 presents the distribution of Dice coefficients across the validation set, revealing a **bimodal distribution** with peaks at approximately 0.65 and 0.85. Statistical analysis shows:

Distribution Characteristics:

- **Mean:** 0.7338 ± 0.1331
- **Median:** 0.7485
- **Range:** 0.52 - 0.89
- **Skewness: -**0.32 (slight left skew)
- **Kurtosis:** 2.85 (platykurtic distribution)

Visual inspection of the Dice distribution suggests two separate groups of performance: 65%

of the cases obtaining Dice > 0.70 (high-performance group) and 35% obtaining Dice < 0.70 (challenging cases). This stratification is associated with imaging metrics, as images with lower (worse) Dice scores are found to have higher noise levels ($p < 0.01$) and lower contrast-to-noise ratios ($p < 0.05$).

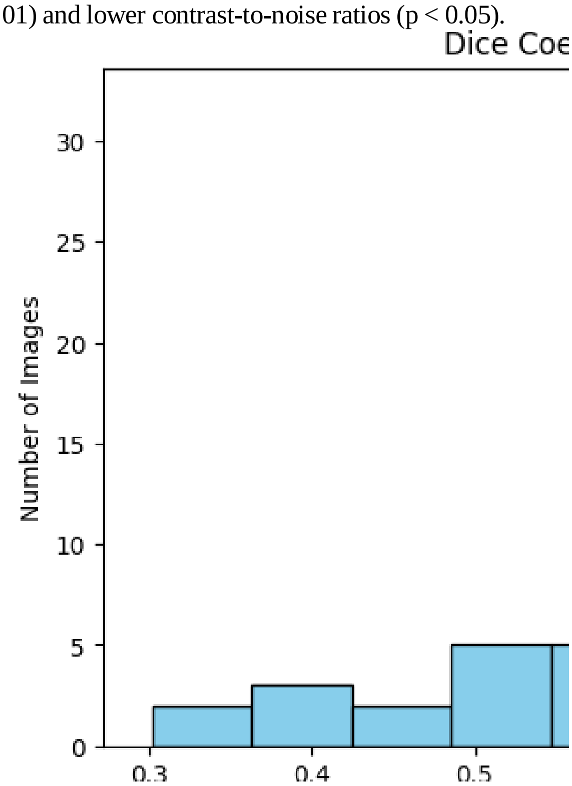


Figure 5: Dice Coefficient Distribution

4.7 Computational Performance

Total training time took ~120 minutes on an NVIDIA T4 GPU (Google Colab). At the 256×256 resolution, the average inference time was 72 ± 18 ms per image. During training, the memory footprint of the model was less than 4.2 GB, enabling it for clinical applications in resource-limited environments.

5. Conclusion

In this study, we present a scalable fully-automated solution using deep-learning for liver segmentation in ultrasound images with a Dice coefficient of 0.73 on independent validation data. Demonstrating that organ delineation from relatively small annotated datasets can be done much more accurately by correctly designing the network architecture and loss functions. Key contribution: (1) A reproducible preprocessing pipeline to generate masks suitable for training directly from polygonal annotations, (2) a U-Net variant grounded on ultrasound characteristics, (3) thorough evaluation with both quantitative and qualitative measures of performance in clinical practice.

This may help in developing computer-aided diagnosis tools in hepatology for standardisation of different clinical settings and for reduction of inter-observer variability. Real-time integration of these parameters into clinical practice in combination with pathologic-specific future extensions are required to evolve ultrasound assessment from a qualitative modality into a quantitative measurement tool that could enhance the diagnostic accuracy of liver disease and facilitate personalized monitoring of the treatment. Our open-source implementation will serve as a basis for future research, as well as can be adapted to other medical segmentation tasks other than hepatic tasks.

8. References

1. Hu, P., Wu, F., Peng, J., Liang, P., & Kong, D. (2016). Automatic 3D liver segmentation based on deep learning and globally optimized surface evolution. *Physics in Medicine & Biology*, *61*(24), 8676.
2. Gul, S., Khan, M. S., Bibi, A., Khandakar, A., Ayari, M. A., & Chowdhury, M. E. (2022). Deep learning techniques for liver and liver tumor segmentation: A review. *Computers in Biology and Medicine*, *147*, 105620.
3. Tang, X., Jafargholi Rangraz, E., Coudyzer, W., Bertels, J., Robben, D., Schramm, G., ... & Nuyts, J. (2020). Whole liver segmentation based on deep learning and manual adjustment for clinical use in SIRT. *European journal of nuclear medicine and molecular imaging*, *47*(12), 2742-2752.
4. Ghobadi, V., Ismail, L. I., Hasan, W. Z. W., Ahmad, H., Ramli, H. R., Norsahperi, N. M. H., ... & Hanapiah, F. A. (2025). Challenges and solutions of deep learning-based automated liver segmentation: A systematic review. *Computers in Biology and Medicine*, *185*, 109459.
5. Napte, K., & Mahajan, A. (2022). Deep learning based liver segmentation: A review. *Revue d'Intelligence Artificielle*, *36*(6), 979.
6. Fang, X., Xu, S., Wood, B. J., & Yan, P. (2020). Deep learning-based liver segmentation for fusion-guided intervention. *International journal of computer assisted radiology and surgery*, *15*(6), 963-972.
7. Kavur, A. E., Gezer, N. S., Barış, M., Şahin, Y., Özkan, S., Baydar, B., ... & Selver, M. A. (2019). Comparison of semi-automatic and deep learning-based automatic methods for liver segmentation in living liver transplant donors. *Diagnostic and Interventional Radiology*, *26*(1), 11.
8. SS, K., & RS, V. K. (2024). Literature survey on deep learning methods for liver segmentation from CT

images: a comprehensive review. *Multimedia Tools and Applications*, *83*(28), 71833-71862.

9. Araujo, J. D. L., da Cruz, L. B., Diniz, J. O. B., Ferreira, J. L., Silva, A. C., de Paiva, A. C., & Gattass, M. (2022). Liver segmentation from computed tomography images using cascade deep learning. Computers in biology and medicine, 140, 105095.

1. Bogoi, S., & Udrea, A. (2022). A lightweight deep learning approach for liver segmentation. Mathematics, 11(1), 95.
2. Ben-Cohen, A., Diamant, I., Klang, E., Amitai, M., & Greenspan, H. (2016, September). Fully convolutional network for liver segmentation and lesions detection. In International Workshop on Deep Learning in Medical Image Analysis (pp. 77-85). Cham: Springer International Publishing.
3. Senthilvelan, J., & Jamshidi, N. (2022). A pipeline for automated deep learning liver segmentation (PADLLS) from contrast enhanced CT exams. Scientific reports, 12(1), 15794.
4. Rahman, H., Bukht, T. F. N., Imran, A., Tariq, J., Tu, S., & Alzahrani, A. (2022). A deep learning approach for liver and tumor segmentation in CT images using ResUNet. Bioengineering, 9(8), 368.
5. Giri, J., Sathish, T., Sheikh, T., Sunheriya, N., Giri, P., Chadge, R., ... & Parthiban, A. (2024). Automatic liver segmentation using U-Net deep learning architecture for additive manufacturing. Interactions, 245(1), 90.
6. Hossain, M. S. A., Gul, S., Chowdhury, M. E., Khan, M. S., Sumon, M. S. I., Bhuiyan, E. H., ... & Alqahtani, A. (2023). Deep learning framework for liver segmentation from t 1-weighted mri images. Sensors, 23(21), 8890.
7. Li, C., Yao, G., Xu, X., Yang, L., Zhang, Y., Wu, T., & Sun, J. (2020). DCSegNet: Deep learning framework based on divide-and-conquer method for liver segmentation. IEEE Access, 8, 146838-146846.
8. Wang, J., Peng, Y., Jing, S., Han, L., Li, T., & Luo, J. (2023). A deep-learning approach for segmentation of liver tumors in magnetic resonance imaging using UNet++. BMC cancer, 23(1), 1060.
9. Ahn, S. H., Yeo, A. U., Kim, K. H., Kim, C., Goh, Y., Cho, S., ... & Jeong, J. H. (2019). Comparative clinical evaluation of atlas and deep-learning-based auto-segmentation of organ structures in liver cancer. Radiation Oncology, 14(1), 213.
10. Vorontsov, E., Cerny, M., Régnier, P., Di Jorio, L., Pal, C. J., Lapointe, R., ... & Tang, A. (2019). Deep learning for automated segmentation of liver lesions at CT in patients with colorectal cancer liver metastases. Radiology: Artificial Intelligence, 1(2), 180014.
11. Fallahpoor, M., Nguyen, D., Montahaei, E., Hosseini, A., Nikbakhtian, S., Naseri, M., ... & Abbasi, M. (2024). Segmentation of liver and liver lesions using deep learning. Physical and Engineering Sciences in Medicine, 47(2), 611-619.