A Novel Hybrid U-Net with Custom Triplet Flatten Loss Function for Liver Lesion Detection

Suraj Patil* and Dnyaneshwar K. Kirange

***Abstract*—Liver cancer ranks sixth among all cancers diagnosed globally. Due to the heterogeneous shape and size of the liver, the manual segmentation of the liver and lesions is a challenging task and time-consuming process. Most of the previous studies in this regard use traditional techniques of image processing to segment the liver and then use handcrafted features to detect lesions and tumors in the liver. The entire process is semi-automatic and results in a loss of information that affects the performance of prediction. Also, deep learning methods employed for liver lesion detection suffer from the misclassification of lesions due to an imbalance of pixel intensities and high processing computational costs. As a result, a new variant U-Net model is designed with a combination of ResNet-18 and ResNet-34 that automatically utilizes 3D contextual information of tumor tissue and detects lesions in the liver. In addition to these, a custom flattened triplet cross entropy function is designed that overcomes the problem of misclassification of lesions due to class imbalance. The novel methodology was evaluated using the benchmark LiTS17 dataset, and the best results were achieved with an accuracy, sensitivity, and specificity of 99.95%, 99.70%, and 99.85%, respectively. We were able to get a considerable reduction in error rate as well as excellent accuracy. The biomedical sector will be transformed as a result of this research.**

***Index Terms*—CT-scan images, deep learning, liver tumor segmentation, flatten triplet cross-entropy loss**

I. Introduction

The liver is an important organ of the digestive system. It is placed in the upper right quadrant of the abdomen, above the stomach, and below the diaphragm. The primary function of the liver is to produce cholesterol and lipoprotein that carries cholesterol to the cell level, which helps in building healthy cells. It secretes a yellow fluid known as bile juice, which aids in the digestion of fats and vitamins [1]. The most common liver cancer is hepatocellular carcinoma, which tends to occur by birth defect, extensive consumption of alcohol or chronic infection such as hepatitis B and C. Liver lesions are a group of abnormal cells in your liver. The preliminary diagnosis made by the doctor for these groups of cells may be a cyst, a mass, or a tumor. Liver lesions can be cancerous or non-cancerous. Non-cancerous liver lesions do not spread and are considered benign. Cancerous liver lesions spread in terms of mutation and are considered malignant. Traditionally, radiologists manually segmented liver and lesion from 3D volumetric 2D scan, which is time consuming process and has variations of 14%−22% when performed by different radiologists [2, 3]. Therefore, automatic detection of liver and lesions from large CT volumes using deep learning technology is becoming popular and highly demanded in clinical practices.

The existing techniques of detection of lesions can be classified into generative and discriminative approaches [4]. In the generative approach, domain-specific prior knowledge is used to model the distribution of tissue that segments lesions. These approaches are based on outlier detection, probabilistic distribution, tumor growth model, and joint segmentation registration. The main disadvantage of generative approaches is that it is difficult to convert domain-specific knowledge into probabilistic models. In contrast, discriminative approaches utilize the modeling of pixel intensities with class labels. Previously discriminative approaches include handcrafted features such as context awareness, multifractal texture, asymmetry, and first-order features, which are fed to conventional machine learning classifiers for prediction .

Recently discriminative approaches include the deployment of deep learning models for the segmentation and classification of lesions. One such technique that currently dominates is the convolution neural network. Both Two-Dimensional (2D) [5] and Three-Dimensional (3D) [6–12] CNN have been extensively used for the segmentation of liver and lesions. 3D network utilizes 3D contextual information for segmentation but requires high computational resources. In contrast to these, 2D network utilizes low computational resources but they cannot integrate 3D relevant information.

In this research work, we use the fusion strategy of ResNet18 and ResNet-32 with the U-Net model to utilize 3D contextual information of tumor tissue using 2D slices of 131 patients with different volumetric CT scan covering all orthogonal views of abdominal portion to get contextual information about liver. The fusion of ResNet-18 and ResNet-34 gives tissue-specific contextual information that can be utilized to segment the liver from abdominal and then segments lesions in the liver. The proposed model uses a unique technique of parallel segmentation and classification of pixels based on 3D contextual information. Moreover, we have designed our custom-based cross entropy triplet function to overcome class imbalance between black and white pixels. The rest of the paper is organized as follows: Section II briefly describes related work in this direction, Section III describes the Material and Method used for model building, Section IV describes the composite loss function for the proposed model, Section V describes network training, Section VI describes result and discussion of proposed work and finally in Section VII conclusion is presented.

Manuscript received October 22, 2022; revised November 25, 2022; accepted January 19, 2023.

Suraj Patil is with Computer Science Department at MPSTME, NMIMS University, Shirpur, India and with SSBT's COE, KBC North Maharashtra University, Jalgaon, India.

Dnyaneshwar K. Kirange is with Computer Engineering Department at SSBT's College of Engineering and Technology, Bambhori, Jalgaon, India.

*Correspondence: psuraj007@gmail.com (S.P.)

II. Related Works

Due to the advancement of machine learning over the past few decades many algorithms are employed to segment liver and tumor from CT scan images, such as thresholding , region growing, the deformable model-based method, nature inspired based methods and machine learning-based methods. Rania and Ghoniem [13] proposed a novel approach to combining bio-inspired algorithms with deep learning. The author combined deep learning networks like SegNet and U-Net. The liver extraction is done using SegNet and lesion detection is done using U-Net deep learning model. The hyperparameter tuning of this hybrid model is done using an Ant Colony nature-inspired algorithm and the proposed system was named SegNet-U-Net-ABC. Soler and Delingette *et al.* [14] have designed a fully automatic model for segmentation of the liver and lesions by considering anatomical and pathological features for the planning of hepatic surgery. Tran and Cheng *et al*. [15] introduced a new approach called TMD-UNet. He improved his model over U-Net in three ways: a) Change the node interconnection, b) use dilated convolution instead of regular convolution, and c) include multi-scale characteristics on the model's input side with thick skip connection rather than a standard skip connection. The proposed model was an experiment with seven different datasets of different modalities and got very promising accuracies of 96%. Motz and Bornemann et *al.* [16] introduced hybrid segmentation technique of adaptive thresholding and morphological operations that performs gray value analysis of ROI for segmentation of liver and lesion. The problem with this method is that user involvement is maximum during entire segmentation process. Wong and Liu *et al.* [17] proposed semi automatic method based on 2D region growing with knowledge base constraint to segment lesions from 3D volumetric scan. The best part of this technique is that it reduces user involvement in segmentation of lesions from 3D volumetric scan. Patrick *et al.* [18] proposed an automated framework for liver tumor segmentation. After the segmentation of liver slices, each image was combined to form a white image with impulse noise and grey specks that resemble a tumor. Following gaussian smoothing, the image was converted to binary using an *iso* data threshold, with tumors showing as black dots on a white background. The findings of these results on the test dataset were promising. Anil and Dayananda [19] used a multi-level deep learning network along with a Fractal Residual Network (FRN) to segment the liver from the abdomen. The model learns the probability of each superpixel in a different region of the liver and indirectly creates different classes based on the intensity of the superpixel. Further FRN is used to segment tumors from the liver with the active contour model method with an accuracy of 86% on the TCI dataset. Huang and Yang *et al.* [20] used extreme learning machine model (ELM) encoder to classify random space feature vector of each voxel of CT scan image for segmentation of liver and lesion. Cai *et al.* [21] has proposed segmentation direct loss function to trained LSTM network to segment pancreas from CT scan and showed promising results.

Based on previous studies, we present an automatic strategy to improve CT liver lesion detection. This strategy is based on past research's benefits, flaws, and constraints of liver lesion detection [22–24]. The proposed solution solves two issues: First, segment the liver using 2D contextual information from 3D volumetric data, then segment the lesions by removing healthy tissues with hybridization of U-Net with ResNet model. Second, we proposed a composite flattened triplet cross entropy function for hybrid model training by taking a mean loss of the triplet loss, cross-entropy loss, and Jaccard distance loss. This resultant composite loss function overcomes extreme class imbalances between white and black pixels for liver-lesion detection. To our knowledge, no one has done work in this direction to delineate tissue in CT scan images into three different classes: background-0, liver-1, and lesion-2, which simplifies radiologist work in proper liver segmentation for surgical operation.

III. Materials and Methods

A. Dataset

We collected datasets from the Liver Tumor Segmentation Challenge (LiTS17), which was organised in conjunction with ISBI 2017 and MICCAI 2017 [25]. This dataset is based on a worldwide platform. Various clinical locations throughout the world offer data and segmentation. In this data, we have follow-up CT scan images from various patients (both men and women) to get an idea of the various conditions and cell types. We have 131 CT scan images in various shapes and sizes (5 to 121 mm) with contrast, where each liver contains one or multiple (1–2) lesions of varying size. Each CT slice is 512×512 in size, with pixel width and height ranging from 0.56 to 0.87mm and slice thickness ranging from 1.25 to 4mm. The dataset provides ground truth for each slice to segment the liver and lesion.

B. Proposed Methodology

In this paper, we propose a novel CNN known as the efficient hybrid U-Net model, an improved variant of the popular U-Net model. The proposed framework is a hybridization of ResNet-18 and ResNet34, deep learning architectures, to get an efficient variant of the U-Net for automatic segmentation and classification of lesions in abdominal CT scan images. The framework pipeline is shown in Fig. 1.

The data preprocessing component of the proposed framework pipeline preprocesses the 2D slice of 3D volumetric data and delineates liver organs from abdominal organs with a window (−100, 400) Hounsfield Unit (HU). This makes the liver organ more distinctive from other organs for training. The proposed model is trained with a composite flattened triplet cross-entropy loss function for segmentation of the liver and lesion. The segmentation and classification of lesions are done automatically with hybrid features coming from encoders ResNet-18 and ResNet-34. To improve the performance of segmentation and classification, a Feature Pyramid Network (FPN) is used as a decoder. The FPN creates different scales of feature maps coming from the encoder. These feature maps give spatial contextual information about the liver and lesion. Finally, the performance of the proposed model is evaluated with ground truth, and metrics are analyzed concerning different model types and loss functions.

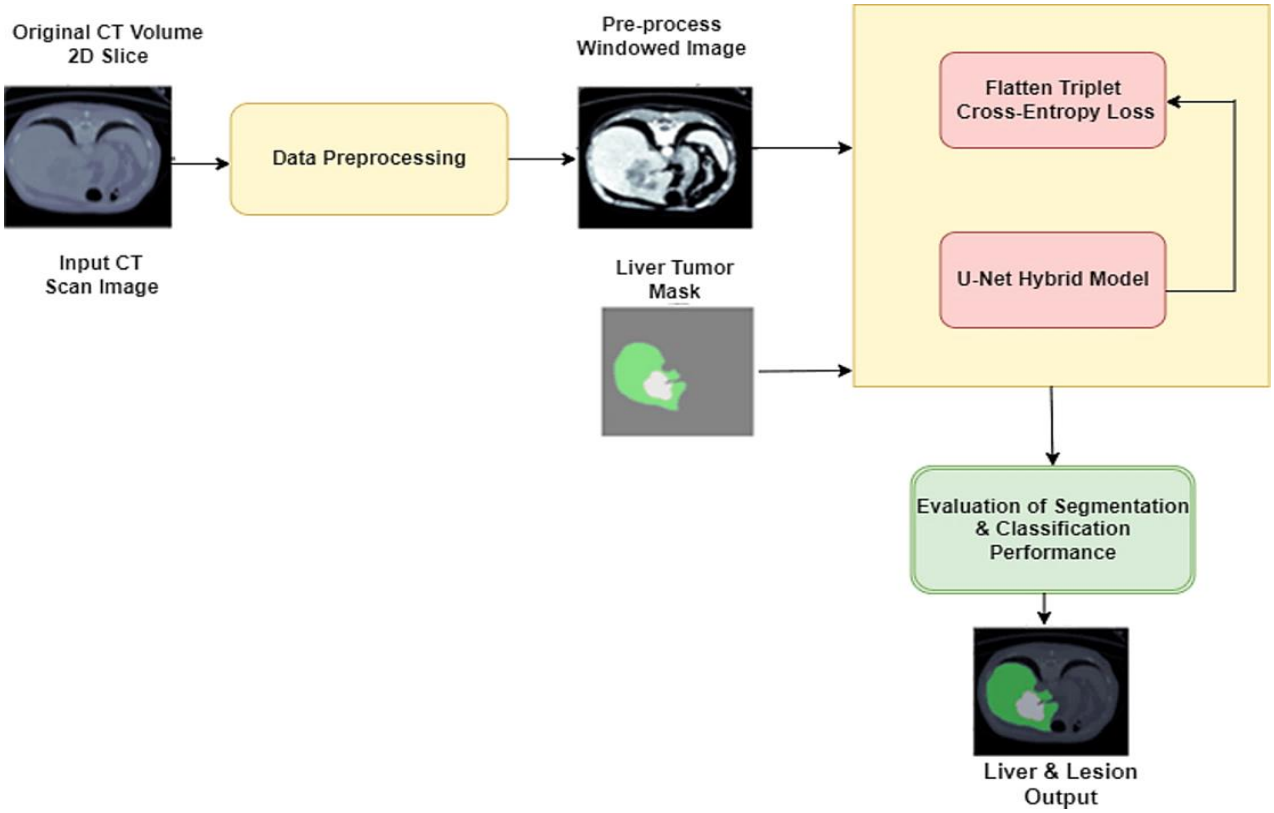

Fig. 1. Proposed liver lesion detection pipeline.

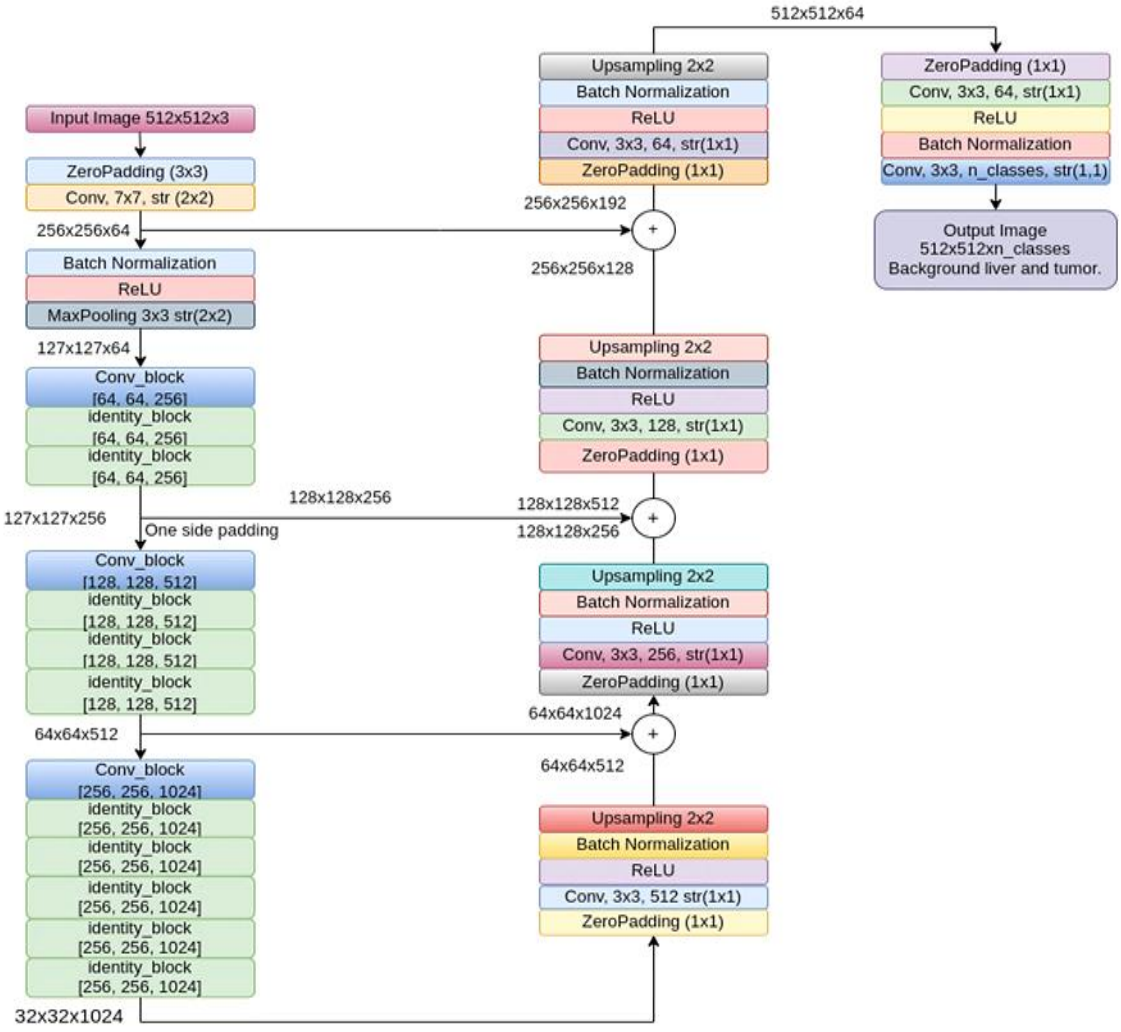

Fig. 2. Proposed hybrid U-net model with combinatorial ResNet-18 and ResNet-34.

C. Hybrid U-NET Architecture

U-Net architecture is a kind of encoder and decoder architecture that is more popular for semantic segmentation in the medical domain and has shown some promising results. But, it has the disadvantage of extracting multi-scaled receptive field features at high-level feature maps, which diminishes lesions with visible scale variations. A ResNet can be used for the encoder/down-sampling section of the U-Net (the left half of the U) as shown in Fig. 2. The stride function in ResNet just converts high-resolution pixels into low-resolution pixels and increases the channel after every step in downsampling. We have used a ResNet-18/ResNet-34 design with the U-net as hybridization and found it to be a very efficient method that trains faster and consumes less memory. Fig. 2 shows the internal architecture of the proposed hybrid U-Net with ResNet-18 and ResNet-34. In this architecture, the U-Net hybridization is implemented using a bounding box that may be identified by some numerical characteristics concerning the image boundary. On a per-pixel basis, semantic image segmentation, also known as pixel-level classification, allows us to detect the receptive field of the liver and lesion within the same input image class. With this procedure, we can detect the tumor in the liver. Our proposed odel consists of semantic segmentation. It has a down and up-sampling path. In the contracting path, we used ResNet-18 and ResNet-34. Every down-sampling channel increases, while the pixels decrease. A 3×3 convolution layer is applied repeatedly, and the Relu function is used after each convolution to convert the linear function to a non-linear function. It is completed after each down-sampling 2×2 max-pooling operation with stride 2. After up-sampling the feature map, a 3×3 max pooling ("up-convolution") halves the feature map, and a ReLU is done at each stage of the expanding path. Batch normalization is used to normalize the gradients coming from the previous layer, which results in a lower number of epochs during training and loss. To improve the performance of segmentation, a feature pyramid network (FPN) is used as a decoder. At various layers of the proposed hybrid U-Net, FPN creates different scales of feature maps, which give different characteristics of liver tumor types and irregular shapes. The use of FPN in the proposed network significantly improves the performance of segmentation.

IV. Composite Loss Function

The loss function is a function that measures the variance between the algorithm's actual output and the target outcome. The idea behind designing the composite loss function is to overcome the problem of class imbalance in terms of pixel distribution between three classes. For training the proposed deep U-net hybrid model we refer to pixel class distribution between three classes (background, liver, lesions) and not the dataset itself. The pixel distribution between the three classes is shown in Fig. 3. The dataset only incorporates lesion information and not the severity of the lesion stage. Hence an attempt is made for only lesion detection. Here we applied the custom loss function, which takes the mean loss of all training sets that passes sequentially through different loss functions such as triplet loss, flattened cross-entropy, and Jaccard loss. The comparison of the custom loss function with Dice loss for different hybrid deep learning models is done to analyze its performance.

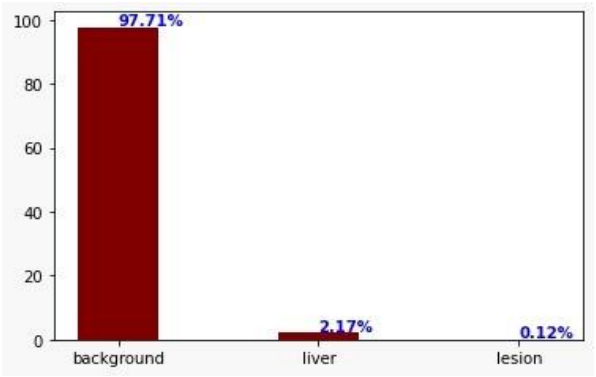

Fig. 3. Pixel distribution of three classes.

A. Triplet Loss

It is an optimization approach that computes distance as triplet between baseline input images *A* (anchor-A) to *P* (Positive image-*P*) and *N* (Negative image-*N*) images to determine feature vector for class prediction $y\_(k = 0,1,2)$ as background (b): 0, liver (l): 1, tumor/lesion (t): 2 as shown by embedding function *F* in Eq. (1). as a triplet loss *L (A, P, N).* It takes each pixel of the baseline input window image and computes distances between the positive input pixel and negative input pixel of the foreground and background and computes training loss. The distance between the baseline input image and the positive input image is minimized and the distance between the baseline input image and the negative input image is maximized by some factor α as shown in Eq. (2). The resultant loss *L* for a single triplet is shown in Eq. (3).

$$L(A,P,N) = \sum_{i=0}^{N} F(A_i^{y_k}, P_i^{y_k}, N^{yk}) \quad (1)$$

$$L\ (A,\ P,\ N) = \|F(A)\text{-}F(P)\|^2 + \alpha \leq = \|F(A)\text{-}F(N)\|^2 \quad (2)$$

$$L\ (A,\ P,\ N) = = \text{Max}(\|F(A)\text{-}F(P)\|^2 - \|F(A)\text{-}F(N)\|^2 + \alpha, 0) \quad (3)$$

The overall triplet loss function of the network for all CT images is given in Eq. (4).

$$TL = \sum_1^N L(A^i, P^i, N^i) \quad (4)$$

B. Flatten Triplet Cross Entropy Loss

Here, we try to pass the predicted output of a network with a triplet loss function to a softmax cross-entropy function. Since we have flattened the output of a 2-dimension array into a single vector and then computed the loss, we termed our resultant loss function a flattened triplet cross entropy function (FTCL), as shown in Eq. (5). with *N* has the total number of all pixels, y_k is the labels of that pixel (0 for background, 1 for the liver, and 2 for the tumor), $p\ (y_k)$ is the predicted probability of pixel that belongs to the liver, lesion,and background in CT scan images. When classifying foreground and background pixels, the FTCL can penalize false positives and false negatives.

$$FTCL = -\left(\frac{1}{N}\right)\sum_{k=1}^{N} y_k \,.\, log\big(p(y_k)\big) + (1 - y_k).\, log\left(1 - \big(p(y_k)\big)\right) \quad (5)$$

Due to pixel class imbalance, the flattened triplet cross entropy function alone is inadequate for precise segmentation of the liver and lesion. To bypass this issue of pixel class

imbalance, FTCL along with Jaccard distance is used. The Jaccard distance measures how dissimilar two sets of data are [24]. The Jaccard loss function is defined as:

$$L_j = 1 - \frac{|Y_+ \cap \hat{Y}_+|}{|Y_+ \cup \hat{Y}_+|} = 1 - \frac{\sum_{k\in Y}(y_k \wedge \hat{y}_k)}{\sum_{k\in Y}(y_k \vee \hat{y}_k)}\ 1 - \frac{\sum_{f\in Y_+}(1 \wedge \hat{y}_f)}{|Y_d| + \sum_{b\in Y_-}(0 \vee \hat{y}_b)} \quad (6)$$

where Y, $\hat{Y}$ represent the ground truth of the labels and model prediction. Respectively Y_+ and Y_- represent the foreground pixel and background pixels set, and $|Y_+|$ represents the cardinality of the foreground pixel as shown in Eq. (6).. A similar notation is also applied for model prediction. The y_k and $\hat{y}_k \in \{0,1\}$ are indexed pixels in Y and $\hat{Y}$. Since $\hat{y}_k$ is the representation of model prediction probabilities, its value will always lie between 0 and 1 and therefore we can approximate the Jaccard loss function as a triplet cross entropy function as shown in Eq. (7), and the model will then be updated by Eq. (8), where *k* represents the predicted value for that pixel. To minimize the loss of the model, and get an updated gradient to update the weights we take the derivative of loss w.r.t network prediction. This assigns a larger gradient to the foreground pixel that automatically balances the background pixel and foreground pixel.

$$L_j = 1 - \frac{\sum_{f\in Y_+} min(1, \hat{y}_f)}{|Y_+| + \sum_{b\in Y_-} max(0, \hat{y}_b)} = 1 - \frac{\sum_{f\in Y_+} \hat{y}_f}{|Y_d| + \sum_{b\in Y_-} \hat{y}_b} \quad (7)$$

$$\frac{dL_{jac}}{d\hat{y}_k} = \begin{cases} \frac{-1}{|Y_+| + \sum_{b\in Y_-} \hat{y}_b} for\ k \in Y_+ \\ \frac{-\sum_{f\in Y_+} \hat{y}_f}{(|Y_+| + \sum_{b\in Y_-} \hat{y}_b)^2} for\ k \in Y_- \end{cases} \quad (8)$$

With the help of the Jaccard loss function, we can make sure that the model gives the same amount of weight to each of the three classes: the background class, the liver class, and the lesion class. In this, we use FTCL and Jaccard together to get the best optimal results by minimizing information loss during training. So, here is our final custom loss function:

$$FTCL = FTCL + L_j \quad (9)$$

V. Model Implementation

The entire experiment was implemented using Python on google colab pro by importing the required packages. We used fastai to implement the deep learning models. In this research, the liver tumor segmentation dataset has been used. In this dataset, a total of 133 CT scan images are present which contain the liver with a lesion. Lesions are abnormal cells present in the liver. In our research, the main objective is to predict the lesions in the segmented liver of the abdominal CT scan images.

A. Data Preprocessing

Data preprocessing is an essential technique for improving data quality. After collecting the dataset, the next step is the preprocessing of CT images. First, all the images were converted into grayscale. After that, the windowing steps are applied to the CT-scan images. For windowing the liver organ, we used the window of (−100, 400) HU, so that most of the irrelevant organs were removed from the CT scan slices. This improves the resolution of the liver and tumor among the surrounding tissue. After completion of the windowing process, custom JPG files were created to train the hybrid model. There are 67,072 slices total in 131 volumes of .nii files. Each volume has approximately has 512 slices.

For training the model, it is important to do annotation of all CT scan images. The annotation of all CT scan images is done with the help of a domain expert using mongo image analysis tool [26]. Since tumor lesions are small, the model concerning single or multiple lesions may affect prediction accuracy. Labeling the ground truth in the (LiTS17), obeyed the following rules: background, liver, and tumor lesion areas were labeled as 0, 1, and 2, respectively. Fig.4.shows different labeling methods for the training dataset.

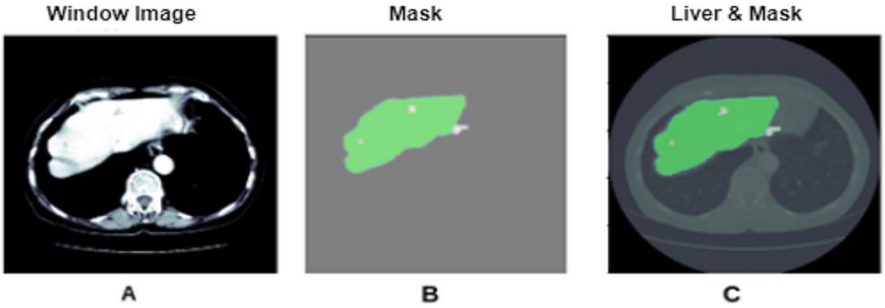

Fig. 4. (A) The windowing CT image. (B) Shows the Liver and tumor by using distinct labels bright green area denotes the Liver and white denotes the tumor. (C) Show the predicted CT image with the background, liver, and tumor.

B. Training Strategy of the U-Net Hybrid Model

TABLE I: Network Training Hyperparameter

Training parameters	Values
Loss Function	Flatten triplet cross-entropy loss
Batch size	16
Learning rate	0.1
Epoch	120
Image size	128×128

For the training model, we need to create a data loaders object in Fastai. With the data loader, we can pass the arguments that specify the feature variables (images) and the target variable (a category for each image). Each input to model training consists of an input image, a liver mask, and a tumor mask. For training, we resize 512×512 images to 128×128 images for storage reasons. The proposed hybrid model is trained with flattened triplet cross entropy with hyperparameter as shown in Table I. As we are using an encoder of a pre-trained network, the training time with 15GB tesla VRAM Nvidia GPU is 1.05 minutes for each epoch and

the simulation runs for 120 epochs with a total training time of two hours with variable learning rate which decays for every epoch , while the detection time is very fast and takes 0.416 sec for each subject with same GPU configuration.

VI. Result and Discussion

Tumors are among the leading causes of mortality globally. Liver tumors are among the most common types of tumors that people get. Recently many deep learning algorithms, have been developed for the diagnosis of tumors and lesions in MRI and CT scan images. Using the existing traditional deep learning techniques to detect liver lesions, the results are unsatisfactory due to the heterogenous shape and size of the liver. Moreover, the drawbacks of the existing algorithms are that they consume more time to detect a liver lesion and do not get more accurate results. To overcome the drawbacks of existing techniques, we propose a U-Net hybrid model with a flattened triplet cross-entropy function and it is used to trained the model. Using this approach, we can utilize the 2D contextual information of the liver and lesion to generate 3D spatial information for liver lesion detection. The evaluation of the model performance using various metrics such as accuracy, sensitivity, and specificity is recorded. Table II shows results recorded for liver lesion detection for all samples of an entire dataset that covers all lesion sizes and samples with a subset of lesions with a size >10 mm that require immediate medical attention by a radiological expert for prompt care and necessary action.The various models that are evaluated in the form of a confusion matrix are U-Net+ResNet32 model with Dice Loss; U-Net+Hybrid model with Dice Loss; U-Net+ResNet32 model with Custom Loss (CL); and U-Net+Hybrid model with Custom Loss (FTCL). The comparison of the proposed hybrid model with other models is shown in Fig. 5.

TABLE II: Comparison of Proposed Hybrid Model with Custom Loss Function with Existing UNet+ResNet with Dice LossFunction

Performance of model with all lesion sizes									
Model	**Accuracy**			**Sensitivity**			**Specificity**		
	Background	**Liver**	**Lesion**	**Background**	**Liver**	**Lesion**	**Background**	**Liver**	**Lesion**
UNet+Resnet32+DL	90.01%	92.21%	91.21%	91.23%	91.21%	91.19%	91.14%	90.09%	92.29%
UNet+hybrid+DL	92.34%	93.33%	92.34%	93.36%	93.45%	92.25%	92.09%	92.96%	93.56%
UNet+Resnet32+FTCL	94.45%	94.45%	94.45%	94.56%	95.65%	94.00%	94.56%	94.45%	94.67%
UNet+hybrid+FTCL	99.84%	99.85%	99.95%	99.70%	99.85%	99.93%	99.70%	99.21%	99.50%
Performance of model with lesion size > 10 mm									
UNet+hybrid+FTCL	99.90%	99.95%	99.97%	99.75%	99.85%	99.95%	99.80%	99.50%	99.75%

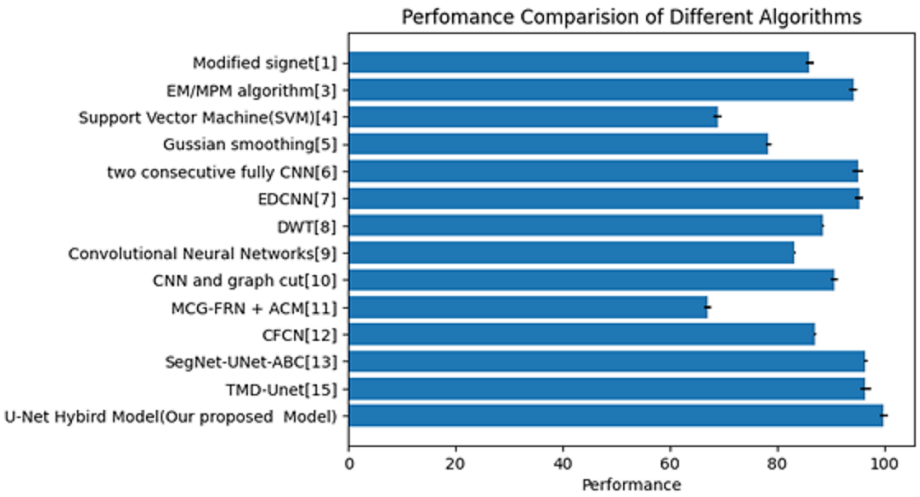

Fig. 5. Performance comparison of different algorithms.

A. Result of the UNet+ResNet32 Model with Dice Loss

In this model, we calculate the result in the form of a confusion matrix. We achieve an accuracy of 90.01%, 92.21%, and 91.21% in terms of background, liver, and tumor. The sensitivity evaluation achieved a result of 91.23%, 91.21%, and 91.19% in terms of background, liver, and lesion. The specificity evaluation achieves a result of 91.14%, 90.09%, and 92.29% in terms of background, liver, and tumor. The model does not achieve a good accuracy score, so to improve the accuracy score, we proposed a U-Net hybrid model.

B. Result of the U-Net Hybrid Model with Dice Loss

In this model, we calculate the result in the form of a confusion matrix. We achieve an accuracy of 92.34%, 93.33%, and 92.34% in terms of background, liver, and lesion. The sensitivity evaluation achieved a result of 93.36%, 93.45%, and 92.25% in terms of background, liver, and lesion.The specificity evaluation achieves a result of 92.09%, 92.96%, and 93.56%. The proposed U-Net hybrid model also does not achieve a good accuracy score due to the increase in information loss during training. To get around this problem, we came up with a custom loss function called Flatten triplet cross-entropy loss.

C. Result of the U-Net+ResNet32 Model with Custom Loss (FTCL)

We use the previous model, which is UNet+ResNet. In this model, we use a custom loss function. We calculate the result in the form of a confusion matrix. The accuracy we achieve is 94.45%, 94.45%, and 94.45% in terms of background, liver, and lesion. The sensitivity evaluation achieved a result of 94.56%, 95.65%, and 94.00%. The specificity evaluation achieved a result of 94.56%, 94.45%, and 94.67%. The model accuracy score is satisfactory, but the model gets trained more on black pixels than white pixels due to class imbalance, so we need to improve the accuracy score as well as sensitivity and specificity. So, we combine the features of ResNet18+RestNet32 with U-Net to get hybrid features with a custom loss function.

D. Result of the UNet+Hybrid Model with Custom Loss (FTCL)

In this model, we calculate the result in the form of a matrix. We achieve an accuracy of 99.84%, 99.85%, and 99.95% in terms of background, liver, and lesion. The sensitivity evaluation achieved a result of 99.70%, 99.85%, and 99.93% in terms of background, liver, and lesion. The specificity evaluation achieved a result of 99.70%, 99.21%, and 99.50%. When we combine our proposed hybrid model with a custom loss function, we find that accuracy increases significantly. We achieved a very good accuracy score and improved the performance of liver lesion detection. The proposed hybrid model was tested with the images, that were not included during the training process. The results of the training are normally higher than those achieved by testing. The results of the testing are shown in Fig. 6.

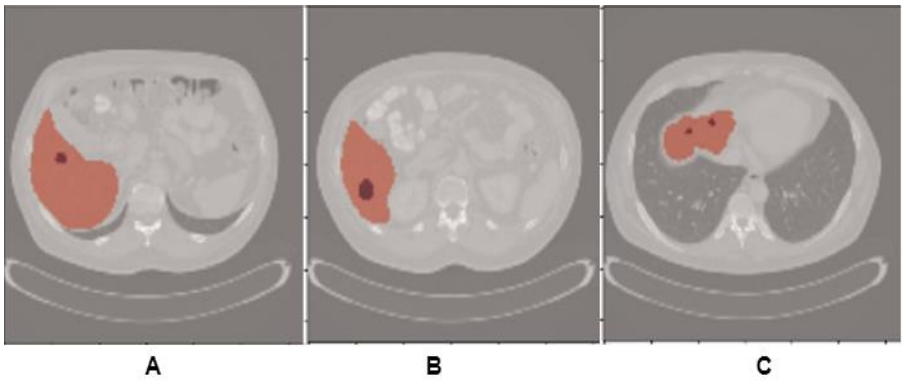

Fig. 6. Samples of the results of testing images (A, B, C): Identify the liver and then detect the lesion position highlighted with a dark red dot. (A): Liver with a lesion, (B): Liver with lesion size >10mm, (C): Liver with multiple lesions less than 10 mm.

VII. Conclusion

In recent years, deep learning has made remarkable breakthroughs in cancer diagnosis and treatment. Some research has also been conducted in the detection of liver tumors, but it has limitations. So, in this paper, we employed four deep learning models: U-Net and ResNet with dice loss; the U-Net hybrid model with dice loss; U-Net and ResNet model with custom loss; and the U-Net hybrid model with custom loss. To improve the accuracy score, we proposed a custom loss function, which is known as flattened triplet cross-entropy loss, and we proposed the U-Net hybrid model with a flattened triplet cross-entropy loss to determine the abnormal growth of cells. In this article, we compare our proposed hybrid model to our other models and previous related work models in terms of several parameters such as sensitivity, specificity, and accuracy. After that, we compared it to earlier studies and saw a significant increase in accuracy. After computing, the metrics, our accuracy, sensitivity, and specificity results are 99.95%, 99.70%, and 99.85% respectively. As can be seen from these results, the U-Net hybrid model with flattened triplet cross-entropy loss is an excellent fit for our objectives. Our method appears to be more effective than the previous method. This research will have a significant impact on the biomedical industry.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Suraj Patil had conducted the research, designed the model, trained and tested the model by analyzing its performance, and written the paper.

Dr. Dnyaneshwar Kirange had guided towards the research work. Both approved the final version of the paper.

Acknowledgement

The author extends thanks to Dr. Nilesh Dhamne, DNB Medical Oncology consultant, Kolhapur Cancer Center for valuable guidance and suggestions to carry out this research.

References

[1] Almotairi, G. Kareem, M. Aouf, B. Almutairi, and M. Salem, "Liver tumor segmentation in CT scans using modified SegNet" *Sensors*, vol. 2, Mar 2020. doi: https://doi.org/10.3390/s20051516

[2] S. P. Kumar, Z. Mohammed, H. Wesam, T. Tarazi, and T. Basant, "AI-driven novel approach for liver cancer screening and prediction using cascaded fully convolutional neural network," *Journal of Healthcare Engineering (Hindawi)*, vol. 2022, Feb 2022. doi: https://doi.org/10.1155/2022/4277436

[3] X. Zhang, J. Tian, D. Xiang, and X. Li, "Interactive liver tumor segmentation from CT scans using support vector classification using watershed," in *Proc. Annual International Conference of IEEE Engineering in Medical and Biological Society*, 2011, pp. 6005–6008.

[4] C. Y. Lin, M. Wu, J. A. Bloom, I. J. Cox, and M. Miller, "Rotation, scale, and translation resilient public watermarking for images," *IEEE Trans. Image Process.*, vol. 10, no. 5, pp. 767–782, May 2020.

[5] H. Nader, M Abdel-massieh, M. Hadhoud, and K. M. Amin, "Fully automatic liver tumor segmentation from abdominal CT scans," in *Proc. IEEE International Conference on Computer Engineering and System*, Dec. 2010.

[6] N. Gruber, "A joint deep learning approach for automated liver and tumor segmentation," *Deep AI*, 2019.
[7] U. Budak, Y. Guo, E. Tanida, and A. Sengur, "Cascaded deep convolutional encoder-decoder neural networks for efficient liver tumor segmentation," *Medical Hypotheses*, vol.1 34., Jan 2020. doi: https://doi.org/10.1016/j.mehy.2019.109431
[8] L. Ali, K. Khelil, S. K. Wajid, Z. U. Hussain, and A. Howard, "Machine learning based computer-aided diagnosis of liver tumors," in *Proc. IEEE 16th International Conference on Cognitive Informatics & Cognitive Computing*, 2017, pp. 139–145.
[9] R. Vivanti, A. Ephrat, L. Joskowicz, O. Karaaslan, N. Lev-Cohain, and J. Sosna, "Automatic liver tumor segmentation in follow-up CT studies using convolutional neural networks," in *Proc. Patch-Based Methods in Medical Image Processing Workshop*, vol. 2, October 2015.
[10] F. Lu, F. Wu, P. Hu, Z. Peng, and D. Kong, "Automatic 3D liver location and segmentation via convolutional neural network and graph cut," *International Journal of Computer Assisted Radiology and Surgery*, vol. 12, pp. 171–182, 2017.
[11] Z. Bai, H. Jiang, S. Li, and Y. D. Yao, "Liver tumor segmentation based on multi-scale candidate generation and fractal residual network," *IEEE Access Journal*, vol. 7. pp. 82122–82133, 2019.
[12] C. Patrick *et al*., "Automatic liver and tumor segmentation of CT and MRI volumes using cascaded fully convolutional neural networks," arXiv preprint arXiv:1702.05970, Feb. 2017.
[13] M. Rania and Ghoniem, "A novel bio-inspired deep learning approach for liver cancer diagnosis," *Journal Information*, MDPI, Nov 2020. doi: https://doi.org/10.3390/info11020080
[14] L. Soler, H. Delingette, G. Malandain, J. Montagnat, N. Ayache, C. Koehl, O. Dourthe, B. Malassagne, M. Smith, D. Mutter *et al*., "Fully automatic anatomical, pathological, and functional segmentation from CT scans for hepatic surgery," *Computer Aided Surgery*, vol. 6, no. 3, pp. 131–142, 2001.
[15] S. T. Tran, C. H. Cheng, T. T. Nguyen, M. H. Le, and D. G. Liu, "Tmd-unet: Triple-unet with multi-scale input features and dense skip connection for medical image segmentation," *Healthcare Journal*, vol. 9, Jan 2021. doi: https://doi.org/10.3390/healthcare9010054
[16] J. H. Moltz, L. Bornemann, V. Dicken, and H. Peitgen, "Segmentation of liver metastases in CT scans by adaptive thresholding and morphological processing," in *Proc. MICCAI Workshop*, vol. 41, no. 43, 2008, p. 195.
[17] D. Wong, J. Liu, Y. Fengshou, Q. Tian, W. Xiong, J. Zhou, Y. Qi, T. Han, S. Venkatesh, and S.-C. Wang, "A semi-automated method for liver tumor segmentation based on 2d region growing with knowledge-based constraints," in *Proc. MICCAI Workshop*, vol. 41, no. 43, 2008, p. 159.
[18] C. Patrick *et al*., "Automatic liver and tumor segmentation of CT and MRI volumes using cascaded fully convolutional neural networks," arXiv preprint, arXiv:1702.05970, Feb. 2017.
[19] B. C. Anil and P. Dayananda, "Automatic liver tumor segmentation based on multi-level deep convolutional neural networks and fractal residual network," *IETE Journal of Research*, Feb 2021. doi: https://doi.org/10.1080/03772063.2021.1878066
[20] W. Huang, Y. Yang, Z. Lin, G.-B. Huang, J. Zhou, Y. Duan, and W. Xiong, "Random feature subspace ensemble based extreme learning machine for liver tumor detection and segmentation," in *Proc. Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, 2014, pp. 4675–4678.
[21] J. Cal *et al*.," Improving deep pancreas segmentation in CT and MRI images via recurrent neural contextual learning and direct loss function," arXiv:1707.04912v2[cs.CV], 18 July 2017.
[22] A. Hoogi, C. F. Beaulieu, G. M. Cunha, E. Heba, C. B. Sirlin, S. Napel, and D. L. Rubin, "Adaptive local window for level set segmentation of CT and MRI liver lesions," *Medical Image Analysis*, vol. 37, pp. 46–55, 2017.
[23] C.-L. Kuo, S.-C. Cheng, C.-L. Lin, K.-F. Hsiao, and S.-H. Lee, "Texture-based treatment prediction by automatic liver tumor segmentation on computed tomography," in *Proc. 2017 International Conference on Computer, Information and Telecommunication Systems (CITS)*, 2017, pp. 128–132.
[24] S. Khan and T. Yairi, "A review on the application of deep learning in system health management," *Mechanical System and Signal Processing*, vol. 107, pp. 241–265, 2018.
[25] A. Jose, C. Luana *et al*., "An automatic method for segmentation of liver lesions in Computed tomography images using deep neural network," *Expert System with Application*, vol. 180, April 2021. doi: https://doi.org/10.1016/j.eswa.2021.115064
[26] Multi-image Analysis GUI. Mango for the Desktop. [Online]. Available: http://rii.uthscsa.edu/mango/mango.html