SN Computer Science (2024) 5:652
https://doi.org/10.1007/s42979-024-02991-2

ORIGINAL RESEARCH

GAN-Driven Liver Tumor Segmentation: Enhancing Accuracy in Biomedical Imaging

Ankur Biswas[1] · Santi P. Maity[2] · Rita Banik[3] · P. Bhattacharya[4] · Jhunu Debbarma[1]

Received: 19 December 2023 / Accepted: 21 May 2024

Abstract

In the biomedical imaging domain, large preprocessed samples of training annotated images are required in techniques employing neural networks for effective training, which makes the method challenging and costly. Data augmentation is widely used to widen the pool of training samples by providing augmented data, enabling the learning method to capture the essential features of regularity and flexibility. However, conventional methods of data augmentation severely limit the capabilities of most training system instances and have an unfavorable influence on reliability. In this study, we addressed the issue of insufficient pre-processed datasets for producing high-quality images of liver tumors. We developed an automatic data augmentation strategy using generative adversarial network (GAN) architectures. We aimed to improve the training efficiency of conventional systems by adding GANs. The automated data augmentation approach made it easier to provide more training samples with higher quality, giving the training process access to a larger and more varied dataset. The proposed approach has been used for the automatic region of interest selection on computed tomography images, which facilitated the training procedure in the subsequent steps. Additionally, the images generated by the generative adversarial network (GAN) are leveraged in the pre-segmentation stage. Finally, we presented an efficient liver tumor segmentation technique with a geometric active contour model, achieving improvements in computational time and Dice score of 0.908, 0.872, and 0.605 on MIDAS, 3Dircadb, and LiTS datasets respectively.

Keywords Data augmentation · GAN · CT · Active contour · Dice

Introduction

The use of digital images has seen a tremendous increase in recent times in medical diagnosis by physicians. Picture archiving and communications system (PACS) are being used widely in healthcare services as they offer several advantages such as low cost, improved quality, accurate dimensions, and flexibility in sharing patient data among healthcare professionals. It also opens up the possibility of using computers to speed up the diagnosis and reduce subjectiveness. The computer-aided diagnosis (CAD) system is already being used in many countries and is expected to be extended for the diagnosis of many kinds of diseases. It will be useful in countries such as India and China where there is a huge population and diagnosis is often delayed due to acute shortage of specialist doctors. CAD has become one of the main research subjects for medical imagery and radiology diagnostics. CAD system provides a "second opinion" for radiologists to use computer analysis to make the final decisions. As a matter of fact, it is quickly inflowing the radiology mainstream. Over the years various medical image processing and analysis tools have been used extensively in CAD and Machine Learning (ML) technique has become a new edition to the CAD system [1, 2]. The diagnosis of liver tumors using magnetic resonance imaging (MRI) or computed tomography (CT) has now been a part of the routine clinical system. The initial stage in image analysis is segmentation which is an integral component of CAD systems. Tumor segmentation is a key enabling technology for

✉ Ankur Biswas
abiswas.tit@gmail.com

1 Department of CSE, Tripura Institute of Technology, Narsingarh, Tripura, India

2 Department of IT, IIEST Shibpur, Howrah, W.B., India

3 Department of EE, ICFAI University Tripura, Kamalghat, Tripura, India

4 Department of Mathematics, National Institute of Technology, Agartala, Tripura, India

Published online: 13 June 2024

the evaluation, scheduling, and guidance of medical applications. The segmentation of the liver tumor also enables structural analyses such as tumor volume estimation, which is a crucial aspect in image-driven surgery, follow-up diagnosis, and therapy [3]. CT images are one of the most widely employed imaging tools to identify and analyze tumors. This is an effective tool to investigate the variations in the structure and the shape of a liver, and noticeable lesions. These are to be used as biomarkers for preliminary clinical diagnosis and progression in primary and secondary hepatocellular tumor disorders. The better spatial resolution, higher imaging speeds, and comparably low cost over MRI have made it a popular tool for medical diagnosis.

Liver tumor segmentation (LTS) can be done in two forms: two dimensions and three dimensions (also called volumetric segmentation). The volumetric approach accomplishes the task of tumor segmentation in three dimensions (3D). It is established that the volumetric measurement provides a more accurate representation of the tumor response than the tumor size which is usually obtained through two-dimensional segmentation [4]. The volumetric segmentation facilitates medical specialists to determine the growth of tumor cells over time and the success of treating cancer. The data gathered can be used to analyze the patient's reaction to treatment and, if appropriate, settle into the therapy. The segmentation of liver tumors is a challenging task due to the considerable diversity in their location, shape, intensity, and texture. This makes it challenging to create a generalized framework to be applicable to the images irrespective of the acquisition modalities [5]. Several deep learning (DL) approaches are being explored in the automatic segmentation of tumors to achieve high degree accuracies. DL allows for the learning and retrieval of latent and hidden patterns from data (image) in order to model and predict some outcomes in real-life problems. The DL-based systems in some cases perform even better than a human in terms of accuracy. Using DL, pattern, and statistical analysis, latent relationships between the symptoms and the disease can be easily discovered. These approaches can be really helpful in providing preventive healthcare as well as a second opinion to a patient suffering from some diseases. However, the present DL framework faces several constraints and every parameter necessitates large data of annotated samples for learning to escape the over-fitting to the training set. Developing such large datasets in the medical image domain is still a challenge since most of the Electronic Medical Records (EMR) are under the Health Insurance Portability and Accountability Act. (HIPAA) compliance, so they are still not released publicly, including for scientific purposes. The data-gathering process requires the cooperation of researchers as well as radiologists, which makes it expensive and time-consuming. Manual delineation also endures from the extensive intra-rater and inter-rater inconsistencies. In practice, training data created for one study cannot be simply transferred to the other. The absence of sufficient data in the medical imaging databases prompted researchers to look for a method for expanding existing datasets with synthetic data [6, 7]. Therefore, the capability of creating medical images artificially is desperately needed. A solution to this problem that allows for more variety and enhances the dataset for the system's training phase is synthetic data augmentation using a generative model.

Generative Adversarial Networks (GANs) are a known efficient method to train an image-creating system [8]. In several different applications, including real-time object identification [9], picture segmentation, extraction of mammogram masses [10], histopathological image synthesis [11], etc., GANs have attained extraordinary recognition and success. Image-to-image transformation, which creates pictures from one modality into another modality, is an attractive use of GAN [12]. The GAN network is made up of two networks that act antagonistically, with the first network producing pictures (either genuine or false) and the second network distinguishing between real and fake images is also a popular framework in biomedical imaging [13–15]. Development of a robust and accurate technique for the segmentation of liver tumors using DL on CT Images becomes challenging yet promising in healthcare support and is the research issue of this work. The objective is to explore data augmentation techniques for enhancing training data followed by the development of automatic segmentation using a 3D volumetric approach on CT images.

The remainder of the manuscript is structured as follows: "Literature Review: Limitations, Scope and Contributions" section makes a brief literature review of related works while "Mathematical Preliminaries" section introduces the mathematical preliminaries for the proposed method. "Methodology" section describes the proposed method. "Results and Discussion" section presents the experimental results with evaluation and discussion. "Conclusion and Future Scope" section concludes the article.

Literature Review: Limitations, Scope and Contributions

To detect the tumor, lesions or other irregularities in the liver using CT images, several semi-automatic and automated segmentations have been reported [16]. This section makes a literature review on two main research issues, namely image (data) augmentation and liver tumor segmentation on CT images.

Data Augmentation

Data augmentation is one approach that enhances the number of required datasets to achieve better results that reduce the impact of overfitting due to the limited training datasets. This also improves the generalization of the trained network. This plays a crucial part in training the model about the features of preferred regularity and robustness when the accessibility of the samples for the training is limited. In medical image segmentation, various arrangements of transformations are generally exploited for augmentation on CT. Flipping, rotation and scaling are commonly used techniques for data augmentation-based tumor segmentation in brain and liver images [17]. Several variants of convolutional neural network (CNN) were used to achieve online data augmentation leading to an improvement in the results [18]. Shift, rotation and elastic deformations are applied by Ronneberger et al. [19] in microscopical images for training, whilst Milletari et al. [20] implemented random deformation using a dense deformation field in MRI on prostate images with B-spline interpolation. In most of the approaches of data augmentation for tumor segmentation, the diversity of the training data in terms of size, shape, location, and appearance is not changed to a great extent. GAN is another approach that has received traction, especially in the medical imaging domain for various tasks. It is a well-known approach for training a system that creates an image and achieves significantly improved recognition and success. Without any clearly defined goal function, GANs are capable of producing highly realistic pictures, and the generator may learn with very little variability. It is recommended over other deep generative models that don't necessitate an objective probability function for optimization, including Variational Autoencoders (VAEs), but produce blurry pictures mostly because of insufficient reconstruction and noise contamination. As a result, scientists working in the field of medical imaging have begun to investigate GANs for picture super-resolution in the study of retinal images, unsupervised outlier detection systems to aid in marker identification, as well as CAD-based image synthesizing [21–23]. As GAN networks enable the application of 'conditions' on both the class label and the image, conditional GANs are increasingly used to produce the necessary pictures. In recent years, a variety of GAN frameworks have been proposed to generate images of high-quality 1024X1024 resolution. This inspired us to utilize GANs to generate augmented CT Region of Interest (ROI) images for liver tumor segmentation.

Liver Tumor Segmentation on CT images

Several research contributions intended to establish CAD applications to reduce the physicians' workload are available on liver tumor segmentation in CT images. A robust multi-threshold liver segmentation framework based on the "Slope Difference Distribution" (SDD) of image histogram is also available [24]. A 3D dual path multiscale CNN (TDPCNN) model for liver and liver tumor segmentation shows a DSC value of 0.689 [25]. To improve the accuracy, the approach based on conditional random fields (CRF) is used to remove all erroneous points from the segmentation results. An implementation of two cascaded deep encoder-decoder CNN (EDCNN) obtained a high DSC of 95.22%. The model is trained to accomplish segmentation in cascade for both the liver and lesions in CT images with a limited number of images [26]. A hybrid implementation of fuzzy c-means with random walker algorithms integrated with cuckoo optimization is reported that shows results in a dice similarity coefficient (DSC) of 0.75 in 3Dircadb and 0.81 in MIDAS datasets. [27] A novel arrangement of the support vector machine (SVM), watershed, and scattered data approximation algorithms are employed on noisy CT and low-contrast MRI. It was applied to heterogeneous or hyper-/hypo-intense abnormalities in the liver and achieved a DSC of 0.83 [28]. An efficient semiautomatic method was proposed for LTS in CT volumes based on the improved fuzzy *C*-means (FCM) and graph cuts. In this work, the volume of interest on tumors in the 3Dircadb dataset is partitioned using confidence confidence-connected region growing algorithm to decrease computational cost. The obtained DSC value is 0.83 [29].

ML models using CNN on the LiTS challenge dataset achieved a DSC of 0.72. The model is developed by combining two models that worked at the voxel- and object level, which resulted in an 85% decrease in false positives when compared to the output from a neural network [30]. Another ML approach using RA-UNet extracted the liver volume of interest and successfully segmented the tumors. The model utilized the basic architecture of 3D U-Net that extorts the relevant data by merging high-level and low-level feature maps [31]. An implementation of the DL model in PET/MRI to automatically detect liver and liver tumors shows a DSC value of 0.88 and 0.53 respectively [32]. Lu et al. [33] combined 3D CNNs with graph-cut methods for an effective liver region location classification on CT images. The precise segmentation was achieved by the graph-cut approach after the probability map was obtained by 3D CNNs. Li et al. [34] proposed liver and tumor segmentation using the hybrid-dense-connected-UNet model in CT images. End-to-end training was done in the model which led to an improved result.

Scope and Contributions of the Work

A literature review on liver tumor segmentation reveals that data augmentation by both the classical approach and GAN improves the training process to a great extent. Although,

GAN-based systems have shown a significant gain in enhancing semi-supervised segmentation [35], however, the final segmentation results obtained from GAN suffer from resolution problems and are difficult to interpret as far as liver tumors are concerned [36]. Nonetheless, their use in the segmentation of 3D medical images in several different modalities highlights several promises and potential [37, 38]. The methods suggest the use of GAN-based synthetically generated images for automatic ROI extraction followed by their use in the pre-segmentation stage. From the literature, it is also seen that the Active Contour model (ACM) is found to be an effective tool for obtaining a volumetric segmentation in multi-modality medical images [39]. The present work integrates a GAN system with ACM to obtain the final volumetric segmentation to offer an improved interpretation in 3D. The integrated system offers an improvement in results than the individual ACM implementation and visualization. In brief, this work proposes an automatic technique of GAN-based data augmentation for the selection of ROI to facilitate the geodesic active contour-based LTS. We suggest and evaluate a technique for augmentation on LT on CT images where segmentation is done on the available datasets more competently. GAN-based data augmentation enhances the learning and speeds up the pre-segmentation classification step with a random forest algorithm. ACM using level set is applied then on 3D demarcations of the tumors in the liver. Simulation results exhibit that the proposed methodology offers improved segmentation over the traditional segmentation approach with or without synthetic data by (0.22–1.22) %. The method shows an improvement in dice coefficient and computational time on different datasets, especially on 3Dircadb and MIDAS. The following is a summary of the contributions made by this work:

- A framework of automatic ROI extraction with the help of GAN-based augmentation that leads to an improvement in accuracy on segmentation by (0.22–1.22) % through data augmentation.
- An improvement in visualization of the tumor in 3D that would lead to better image analysis and consequent diagnosis over existing 2D methods.

Mathematical Preliminaries

This section makes a brief introduction to the related mathematical tools and techniques.

Generative Adversarial Networks

A GAN network consists of two channels, a generator model 'G' and a discriminator model 'D', which are arranged in a position against one another and further configured to compete. The 'G' channel learns to generate the data distribution 'P_r' by making artificial pictures that seem similar to genuine photos and are challenging to distinguish in the 'D' channel. Alternatively, the 'D' network is learned to distinguish between genuine images and those created by 'G'. As seen below, the system simulates a min–max game between two players for training purposes. It is represented as,

$$\underset{G}{min}\cdot\underset{D}{max}E_{x\,P_r}\left[logD(x)\right]+E_{z\,P_z}\log[(1-D(z))] \tag{1}$$

where 'z' stands for the noise vector from the data distribution 'P_z' and x corresponds to the sample of real photos from the target distribution of data 'P_r'.

The objective of the generative model is to create new pictures that are strained from the learned distribution by performing absolute learning from the distribution of data 'P_r' using image samples such as $x^{(1)}, x^{(2)} \ldots x^{(n)}$. The basic GAN architecture involves two networks: 'G', which stands for generator, and 'D', which stands for the discriminator. These networks are learned to counterpart each other in a manner similar to a min–max game involving two participants:

$$\emptyset_{(D)}\left(\theta^D,\theta^G\right)=-\frac{1}{2}E_{x\,P_r}logD(x)-\frac{1}{2}E_z\log\left(1-D(G(z))\right) \tag{2}$$

$$\emptyset_{(G)}\left(\theta^D,\theta^G\right)=-\emptyset^D \tag{3}$$

where $\emptyset_{(D)}$ and $\emptyset_{(G)}$ stand for the cost functions of the discriminator and the generator respectively, whereas $\theta^D and \theta^G$ indicates the variables that should be tuned for the various networks. In the discriminator network, D: $x \rightarrow [0, 1]$ illustrates a mapping between a picture and the likelihood that the image is real, where 'y' represents noise in the generator network and '$x\,P_r$' is a sample from the desired distribution. During training, the cost functions for $\emptyset_{(G)}$ and $\emptyset_{(D)}$ are minimized. The equilibrium is attained when $D(G(z))=½$, at which point the discriminator is incapable of determining the distinction between the genuine and fake produced images. Because the discriminator in the architecture receives input from instances of the generator in the actual world and subsequently learns about its conclusion—whether it is real or fake—training is conducted in a semi-supervised manner. Real-time translation of images begins when the generator converts arbitrary tumor samples into objective images and is capable of producing pictures that superficially look like photos with tumors. The proposed GAN framework to synthesize tumor ROI for facilitating the subsequent pre-segmentation classification is shown in Fig. 1.

Learning Wasserstein GAN (WGAN)

The divergence criterion that GANs reduce is discontinuous in terms of the generator parameter complexity that is frequently

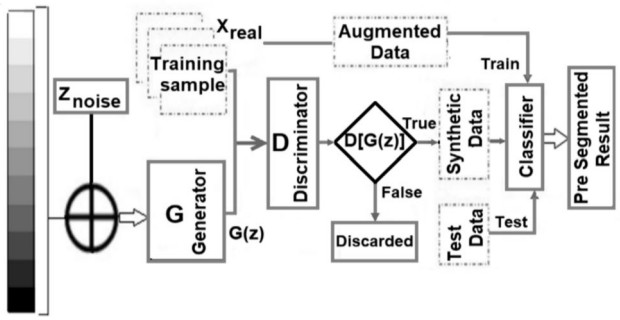

Fig. 1 GAN framework for ROI image synthesis and pre-segmentation tumor classification

present during training. In order to change the probability distributions *'q'* into *'p'* the author of [40] proposed the distance W(q, p) utilizing Earth-Mover (EM) (also known as Wasserstein-1) to determine the distance between the two distributions. This method minimizes mode collapse while achieving sustained learning. Even though the two distributions in this scenario are not overlapping and are located on lower-dimensional regions, $W(q, p)$ is still thought of as continuous and provides a meaningful and even representation of the distance between them. Kantorovich-Rubinstein duality is used to generate the value function of WGAN [41] to achieve,

$$\min_{G} \max_{D \in L} E_{x\,P_r}[D(x)] - E_{z\,P_g}[D(z)] \tag{4}$$

where L represents the 1-Lipschitz function set and 'P_g' being the distribution of the model. Minimizing the value function of the generator reduces $W(P_r, P_g)$ under an optimum discriminator (referred to as critic which is not learned for classifications. By creating the critic function, whose gradients about the input are executed superiorly apart from its GAN matching component, the WGAN value function facilitates the efficiency of the generator. The findings confirmed that, unlike GANs, the WGAN value function embodied sample quality correlation. To enforce the Lipschitz constraint on the critic, weights are bound to persist in narrow windows like [− 0.01, 0.01]. This compacts the space of the parameter [− c, c]. A subset of the k-Lipschitz functions for some k that depends on 'c' and the critic architecture are the functions that satisfy the lower and upper bounds necessary to uphold the Lipschitz continuity constraint. Utilizing EM distance, WGAN achieves steady learning (or Wasserstein-1 metrics),

$$W(P_g, P_r) = \inf_{P \in \prod (Pg, Pr)} E_{x,z\,P_g} \|x - z\| \tag{5}$$

where $\prod (Pg, Pr)$ joint distributions 'p' and marginal are 'Pg' and 'Pr' correspondingly and ideally expressing the mass transfer from one distribution to another.

Active Contour Model

Active contour is a mathematical platform or model that utilizes the energy and forces present in the image to separate the target object. It establishes an isolated borderline or curvature enclosing the ROI for segmentation. The two types of ACM level set implementation are edge-based ACM and region-based ACM. The developing curve is watched in the second category until it reaches the edge of the desired item in the picture, which is carried out using knowledge of the slope or spatial relation. Although these models are sensitive to noise, but achieve enhanced results for the segmentation of the objects with the presence of strong edges. Conversely, the execution of region-based techniques depends on the region's statistical information in order to mature the curve intending to delineate the items present in the image and possessing the competence to delineate the image edges which are slightly weak. In ACM-based LTS, an initiatory confined curve is enhanced and abated utilizing the intensity or texture information of the image, till the tumor boundary is conformed. But complications arise while disbanding the curve if the structure comprises of collective in detached regions. The ACM formulation of the image, $g(x)$ after the pre-segmentation phase, is a parametric contour 'Γ' reflecting the tumor region's border that evolves over time *'t'* as given by

$$\frac{\partial \Gamma}{\partial t} = [g(\Gamma) + \alpha K_\Gamma] N_\Gamma \tag{6}$$

where the mean curvature of C denoted by K_Γ with N_Γ being the unit outward normal of Γ while 'α' is a scalar parameter. The initial seed points are necessary for the contour's initiation for the image-obsessed energies to further drive the contours in the direction of the tumor's borders. Two forces that are fascinating with images are internal forces, $f_{internal}$ that are sparked by the curve for retaining the contour smooth during twisting or bending, and external forces, $f_{external}$ which are planned from the ROI's available data to evolve the contour toward the boundary of the tumor or other specific characteristics inside ROI. Active contour execution is made easier and faster by level-set formulation. The contour evolution derivation method is provided by,

$$\Gamma(t, s) = f_{internal} + f_{external} \tag{7}$$

where Γ(t, s) is the contour at any point in time t, characterized by x, and f is the normal force that acts on the curve. The functional curve is provided by,

$$\int_\Gamma E_{internal}(\Gamma(s)) + E_{external}(\Gamma(s)) ds \tag{8}$$

The given problem's optimal approach is resolved using the Euler Lagrange technique as,

$$\Gamma_{optimal} = \underset{\Gamma \in F}{argmin} E(\Gamma(s)) \tag{9}$$

to determine the value of 'Γ' that justifies the usage of the energy function, E to determine the contour of the least value. The Euler–Lagrange hypothesis is intended to reduce,

$$\int_{\Gamma} E(s, \Gamma, \Gamma\prime, \Gamma\prime\prime) ds. \tag{10}$$

Therefore, it must be resolved,

$$\partial E/\partial \Gamma - \partial/\partial s\, \partial E/\partial \Gamma\prime + \partial^2/\partial s^2 \partial E/\partial \Gamma\prime\prime = 0 \tag{11}$$

Equation (10), which must be solved, was transformed into a time differential system where the contour function $\Gamma(s)$ is now a time-dependent function $\Gamma(s, t)$,

$$\frac{\partial \Gamma}{\partial t}(s,t) = \frac{\partial E}{\partial \Gamma}(s,t) - \partial/\partial s \cdot \partial E/\partial \Gamma\prime(s,t) + \partial^2/\partial s^2 \cdot \partial E\, \partial \Gamma\prime\prime(s,t) \tag{12}$$

Consequently, if we insert the following equation into the energy function 'E', it must be evaluated,

$$\begin{aligned} \frac{\partial \Gamma}{\partial t}(s,t) = {} & \alpha\left\{\partial^2 \Gamma/\partial s^2(s,t)\right\} - \beta\left\{\partial^4 \Gamma/\partial s^4(s,t)\right\} \\ & + \delta \nabla\left(\left\|\nabla\left(G_n x I\right)\right\|^2\right)(\Gamma(s,t)) \end{aligned} \tag{13}$$

Dual forces are taken into account viz., RC and CF. While RC forces the contour inwards and outwardly, CF maintains a smoother contour boundary and stops contour leakage. For consistency and effectiveness in computing, 'C' is inherently expressed as a function's 'ϕ' zero level set specified on g(x), and the evolution Eq. (7) is represented as a 'ϕ' progression in the small area all over the 0-level set. The seed points in the tumor are initialised. The contour grows in areas for positive values $g(x)$ i.e., when $P(x \epsilon T) > P(x \epsilon \Omega \setminus T)$, while contracts in areas for all negative values of g(x) throughout evolution. The scalar 'α' controls how smooth the contour 'C' is. The convergence criteria of contour evolution could be provided manually, which provides a 3D depiction on-screen and allows the progression to be stopped and restarted whenever necessary. The contour evolution is also terminated using a specified value (max_iteration) as a convergence criterion. The extracted tumor is then acquired.

Methodology

In the proposed methodology, the dataset is initially pre-processed using Contrast Limited Adaptive Histogram Equalization (CLAHE) [42] and normalizing local contrast with the goal of improving local contrast in the pictures and enhancing tumor local contrast. The comparison of original and processed images is shown in Fig. 2. The method eases the synthetic data augmentation and further training process for ROI extraction which in turn alleviates the pre-segmentation classification in the successive step.

Figure 3 depicts the overall schematic interactive framework with Data augmentation and ACM-based Segmentation. The methodology in the proposed scheme for the extraction of tumors comprises three main stages:

Tumor ROI Extraction

To improve the computation time, the ROI region around the tumor should first be extracted from the image. This minimizes the total number of pixels to be taken into account, which in turn speeds up processing. The main problem in automatic segmentation is the selection of ROI for the tumor, which varies in size, location and appearance. Training the system for ROI selection requires high-volume labelled training datasets. We enhanced the data in two ways to increase the training set and enhance the segmentation outcome:

1. Creation of new synthetic ROI pictures that are learned from existing datasets using generative models;
2. Classical augmentation employing various image editing techniques on the original pictures.

Classical Augmentation

Augmentation is usually employed to expand the training data so as to reduce the overfitting issue. Each tumor ROI was first rotated at arbitrary degrees, and then each rotated ROI was repeatedly flipped across all orientations. Each ROI was downscaled to a uniform pixel size of $(64 \times 64 \times 64)$ by means of bicubic interpolation. GANs are also utilized for ROI Synthesis. Figure 4 shows the result of data augmentation and ROI image synthesis.

Pre-segmentation Stage

In this stage, pre-segmented image, $g(x)$ is obtained from the ROI image within the range of intensities [− 1, 1], given by,

$$g(x) = f(x \epsilon P_{tumor}) - f(x \epsilon \Omega ¥ P_{tumor}) \tag{14}$$

where 'P_{tumor}' is the tumor (foreground) and 'Ω' is the image domain. To detect the foreground/ background probabilities of the pre-segmentation result, the Random Forest (RF) classification is used. For obtaining the result the system is trained on Liver Tumor Segmentation Challenge (LiTS) 2017 [43] datasets containing 200 3D abdominal CT scans data and augmented data generated and synthesized for ROI extraction. For each pixel in the tumor, a feature

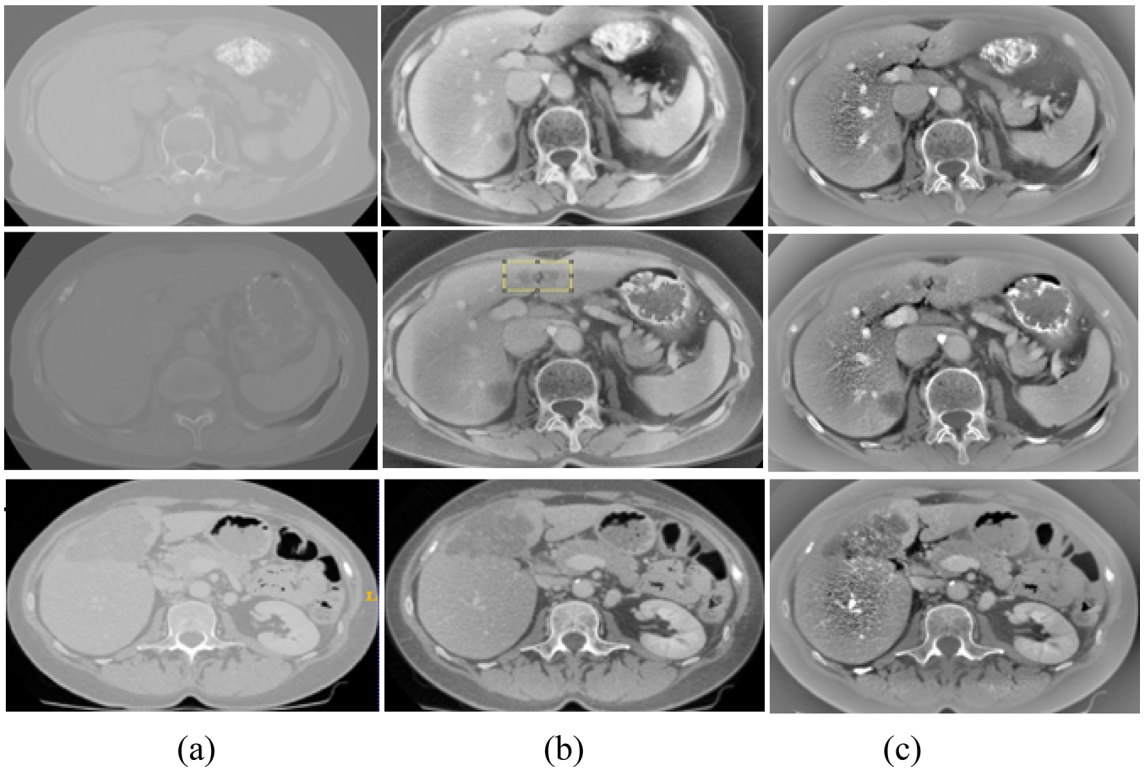

Fig. 2 Result of pre-processing **a** original image, **b** CLAHE, **c** normalised local contrast enhanced image

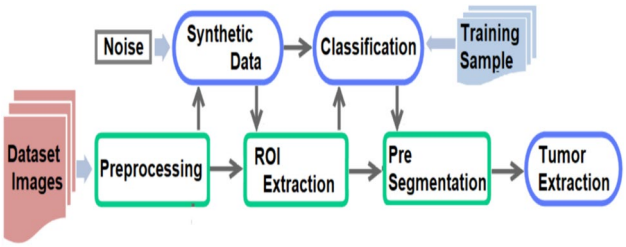

Fig. 3 The overall framework for liver tumor segmentation

vector consisting of the intensity of the pixel in the image and intensities of the neighboring pixels is created allowing the inclusion of texture data in the classification algorithm. The positions of the pixel are also included in the feature vector, enabling the RF classifier to use spatial features. The classifier is then applied to each pixel in the picture domain after being trained on this feature information, yielding the probabilities $P^n(x)$ for each pixel x and each n class. The foreground and background classes of pixels are used to train the classifier. Using the probability for the foreground and background pixels, $f(x \epsilon P_{tumor})$ and $f(x \epsilon \Omega \mid P_{tumor})$, respectively, are derived. Further trained model is tested and validated on test data from LiTS challenge, 3Dircadb [44] and MIDAS [45] dataset provided with ground-truth of liver and liver tumor segmentation as shown in Figs. 5 and 6. Addition of augmented data in the training process improves the final segmentation result.

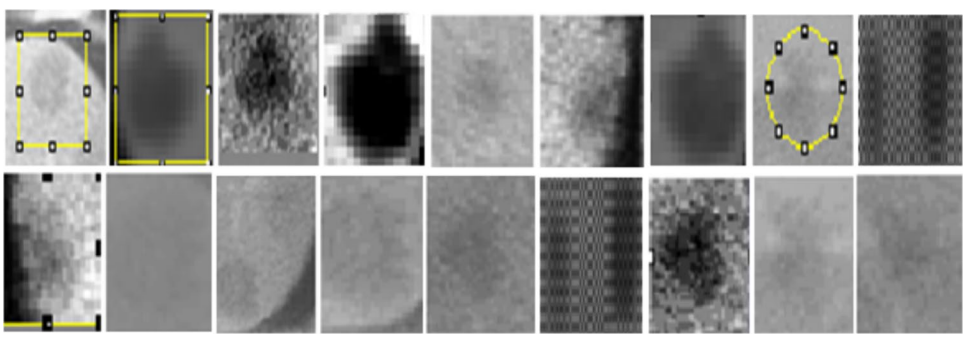

Fig. 4 Synthetic images using data augmentation and GAN from the original dataset

Active Contour Evolution

The tumors cannot be created in three dimensions using ROI photos or images from the pre-segmentation stage. Active contour evolution is therefore crucial for improving tumor visibility. Algorithm 1 illustrates the active contour implementation in the proposed algorithm in a generic version.

Results and Discussion

The proposed framework has been implemented and validated on three datasets of liver CT sequences: LiTS, 3Dircadb and MIDAS. In this section, the various parameters, and architecture used in the framework along with time requirements and validations are discussed. Figure 7 shows

Input: Liver image dataset
Output: 3D tumor segmentation in the liver images

1. *Load the image*
2. *Apply classical data augmentation on the image.*
3. *Train a GAN system using the dataset to synthetically generate images.*
4. *Apply the trained GAN system to generate synthetic images for the given input.*
5. *Perform automatic ROI extraction using the dataset images and synthetically generated images obtained from the GAN system specifically targeting liver tumors.*
6. *Apply pre-segmentation techniques for liver tumor images to the extracted ROIs.*
7. *Initialize the ACM (Active Contour Model) for tumor segmentation in liver images.*
   For each image, perform the following:
   a. *Set the initial contour for the ACM model within the liver region.*
   b. *Iterate until convergence:*
   c. *Update the contour using the ACM algorithm, focusing on the tumor boundaries within the liver.*
8. *Obtain the tumor segmentation result using the final contour obtained from the ACM model.*
9. *Visualize the segmented tumor in 3D.*

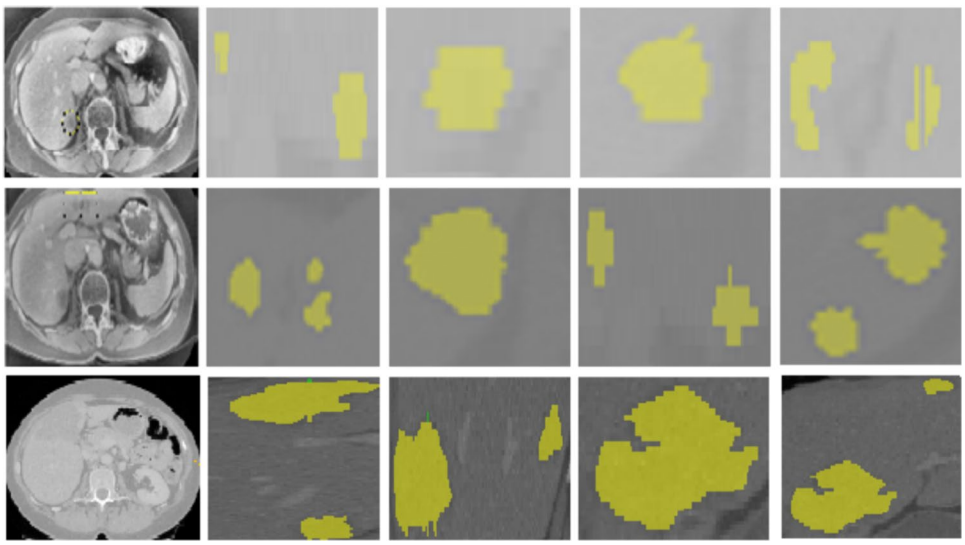

Fig. 5 The pre-segmentation of tumor from the original dataset of LiTS and 3Dircadb. **a** Processed image, **b–e** tumor ROI

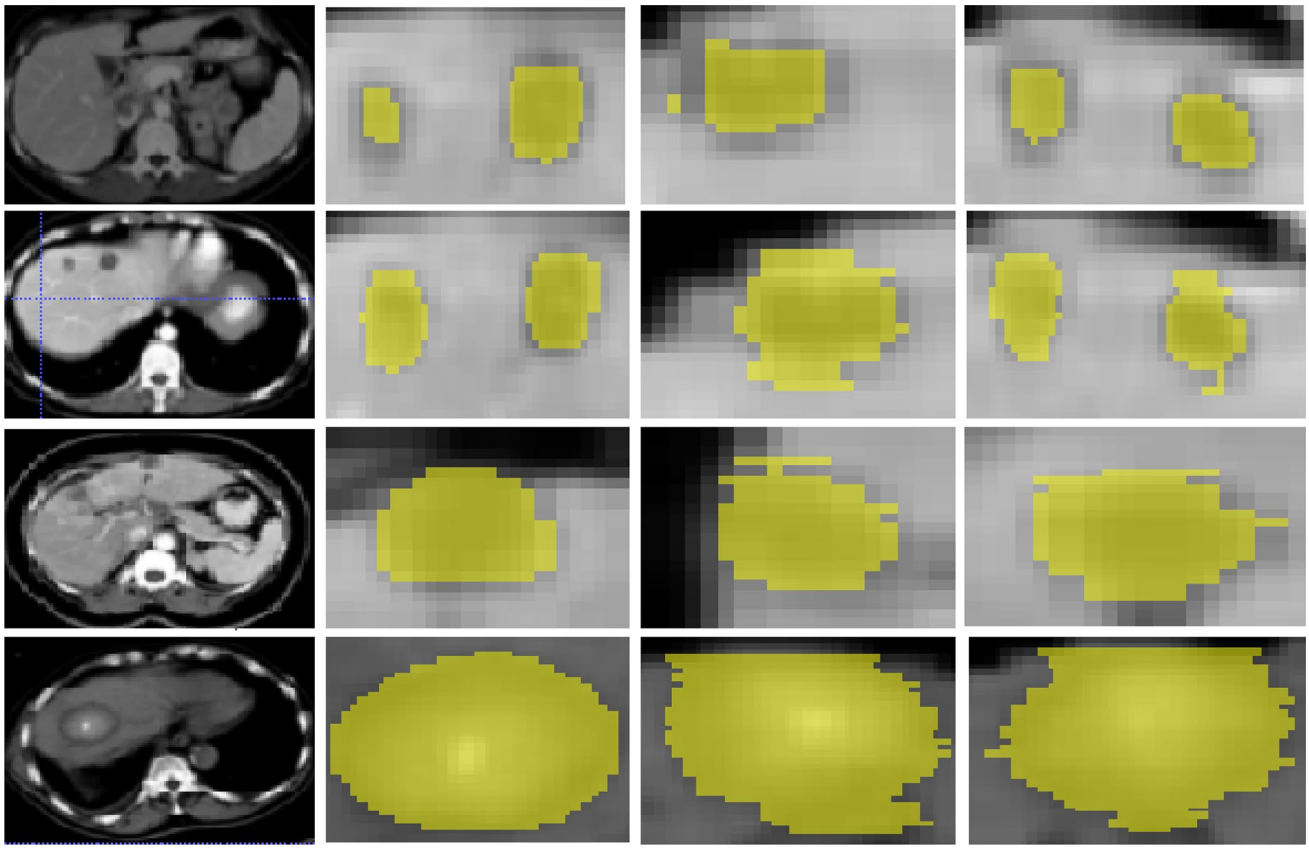

Fig. 6 The pre-segmentation of tumor from the original dataset of MIDAS dataset. **a** Original image, **b**–**d** tumor ROI

the intermediary outcomes of the suggested methodology in three datasets, where (a) shows the pre-processed image, (b)–(d) are the liver tumor image cases in three planes, and (e) is the final extracted tumor in 3D. For each liver image in each plane, the goal was to accurately delineate the boundaries of the tumor. The parameters used to obtain the tumor using active contour are tabulated in Table 1.

The GAN Training

The commonly used WGAN architecture is preferred using the neural network module of PyTorch. CuDNN is used to accelerate neural network operations: convolution, pooling, normalization, and activation functions. We used batch normalization to make the model suitable for the GAN architecture. Leaky ReLUs with alpha 0.2 were chosen in order to provide a tiny slope for the negative values, as suggested in [46, 47]. The size of feature maps in both networks is 64, the number of training epochs is 50, the learning rate for the generator is 0.0002 and discriminator is 0.0005, and the Beta1 hyperparameter for Adam optimizers is 0.5. Since maximum pooling results in a sparse gradient, average pooling is used to simplify the GAN learning. In the last layer, a 'Tanh' activation is used to generate the output image. Binary Cross Entropy loss is used to train both the generator and the discriminator. Losses for both networks are noted during training for monitoring and visualization. The modifications to the architecture made it easier to practice steadily and perform better. The discriminator's architecture was kept similar to the U-Net architecture [48, 49]. The U-Net architecture, which consists of an encoder-decoder structure with skip links, is frequently used for semantic segmentation tasks. The patches of $64 \times 64 \times 64$ are extracted from 3D CT images to train the model. The suggested network utilized Adam's optimizer with a batch size of 16.

Computational Time

The projected methodology was accomplished in the Windows GPU platform with an Intel Core i7-4790 processor running at 3.6 GHz and 16 GB of RAM with the help of open-source platforms ImageJ (https://imagej.net) and ITK-Snap (www.itksnap.org). Python TensorFlow was used to develop WGAN on a Windows GPU platform. For each $64 \times 64 \times 64$ picture that was shrunk from the entire data, the overall training time was around 2 h. After the training procedure is done, segmentation takes a total of 10–20 s, plus an additional 60 s for selecting the ROI. Thus, for a

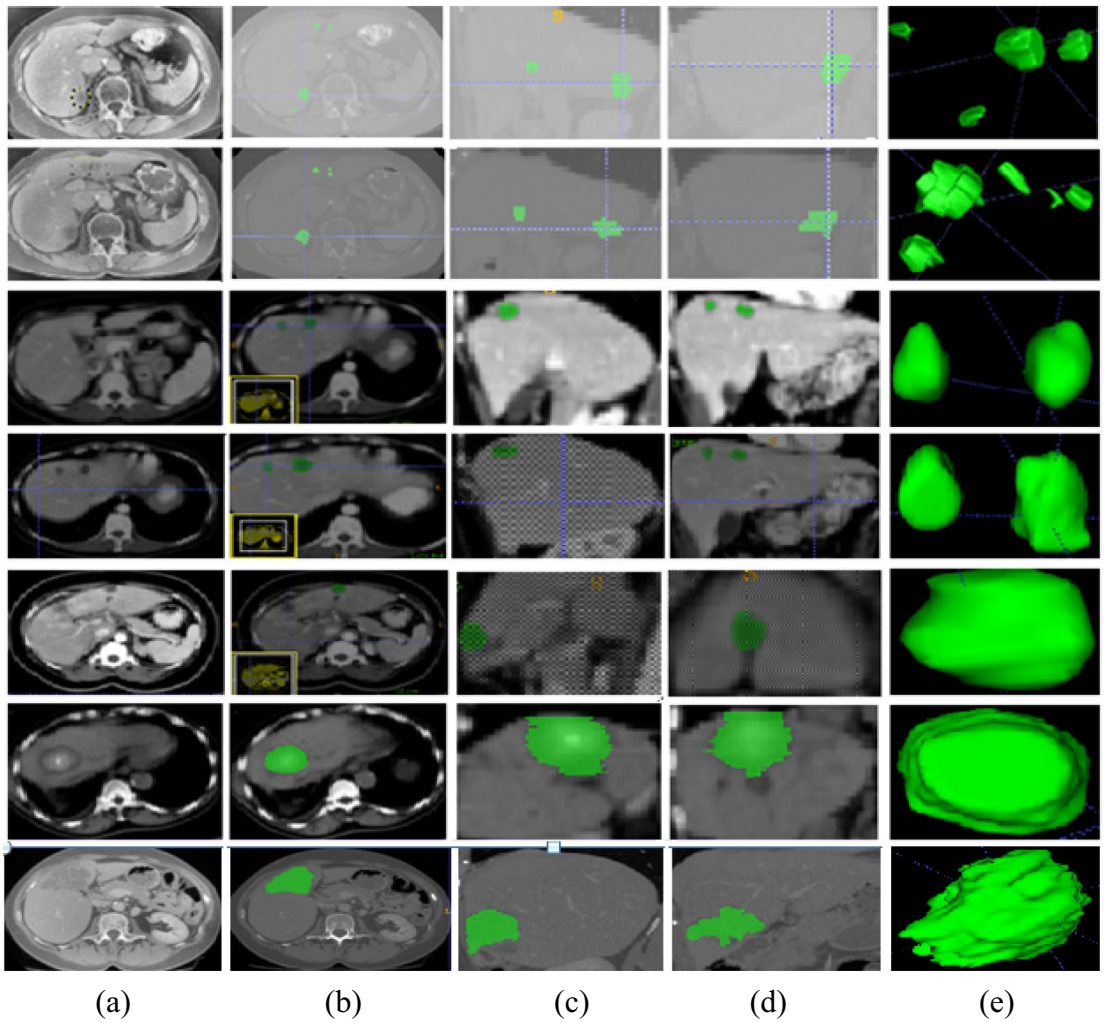

Fig. 7 Findings of the suggested method (**a** CLAHE preprocessed images, **b–d** tumor demarcation employing active contour in three orthogonal viewpoints via the image, in the axial, coronal, and sagittal planes, **e** 3D view of tumor segmentation in its final form)

specific tumor on a CT picture, an overall time of fewer than 90 s is required, which is much less time than the 30–40 min required for manual delineation.

Table 1 The variables utilized to extract the tumor from the ROI image in 3 dimensions are shown in Fig. 5

Segmentation type	Parameters	Value
Pre-segmented tumor (random forest)	Tree numbers	200
Tumor (using active contour)	Advection	2.20
	Propagation	1.00
	Curvature	1.00
	Grayscale tolerance	30.00
	Region competition function	1
	Smoothing function	0.2

Evaluation

To assess the grade of the segmentation, two performance evaluation criteria are used, DSC and JSC [50]. DSC displays the similarity between two supplied picture samples that yields numeric numbers between 0 and 1. A DSC score of 1 indicates that the segmentation result (image) is perfectly aligned with the ground truth. It is numerically represented as,

$$dice(I_1, I_2) = 2X\left|intersection(I_1, I_2)\right| / \left(\left|I_1\right| + \left|I_2\right|\right)$$

JSC, often referred to be Intersection over Union, is a number that is used to assess how similar and dissimilar images are. When the segmentation result (image) has a JSC score of 1, it means that it exactly fits the ground truth. It is numerically represented as,

$$jaccard(I_1, I_2) = \left|intersection(I_1, I_2)\right| / \left|union(I_1, I_2)\right|$$

where I_1, I_2 are two images.

Other metrics include False positive (FP), False negative (FN), Hausdorff Distance (HD), Standard Surface Distance (SSD), and Maximum Surface Distance (MSD). When a pixel is incorrectly identified as belonging to a target item or condition when it really does not, this is known as FP. A FP might occur, for instance, if a pixel was identified as belonging to a tumor when it was actually healthy tissue. On the other hand, a FN occurs when a pixel that is part of a target item or condition is mistakenly identified as not part of it. For example, it would constitute a FN if a damaged pixel was labeled as normal. The greatest distance between two groups of pixels is measured by the HD. It determines the separation between the farthest and closest pixels in two sets. The average distance between comparable pixels on two surfaces is determined using SSD. The average difference between pixels in a segmented tumor and pixels in a real tumor, for instance, is measured in medical imaging. And the maximum distance between any two matching pixels on two surfaces is determined using the MSD calculation. It determines the biggest discrepancy between segmented object pixels and reference or genuine object pixels. Table 2 compiles the quantitative analysis of segmentation outcomes attained utilizing the recommended plan.

On the MIDAS, LiTS, and 3Dircadb datasets, the proposed approach achieved Dice scores of 0.908, 0.872, and 0.605, respectively. The JSC score of 0.831, 0.773 and 0.434 on three datasets are also competitive and the False positive (FP) and False negative (FN) values are within limits. It shows that the proposed technique produces acceptable liver tumor segmentation results. There were several submissions about tumor segmentation on 3Dircadb dataset. We reached a DSC of 0.872, JSC of 0.773, False positive of 0.063, False negative of 0.185, Hausdorff Distance (HD) of 14.734, Standard Surface Distance (SSD) of 1.473, Maximum Surface Distance (MSD) of 8.062. In comparison with other methods, the proposed method outperformed all methods as listed in Table 3.

Conclusion and Future Scope

In this study, we have developed a hybrid architecture that effectively and efficiently extracts liver tumors from CT volumes. Our novel approach uses generative adversarial networks (GANs) to extract three-dimensional (3D) structures pixel-by-pixel, increasing accuracy and reducing time complexity. The proposed method enables the professional

Table 2 Evaluation of segmentation results

Dataset	DSC	JSC	FP	FN	HD	SSD	MSD
MIDAS	0.908	0.831	0.140	0.039	2.492	0.274	1.414
3Dircadb	0.872	0.773	0.063	0.185	14.734	1.473	8.062
LiTs	0.605	0.434	0.521	0.178	37.424	0.506	1.732

Table 3 Comparison of the proposed method's quantitative segmentation results to those of other innovative algorithms on the 3Dircadb and MRI/PET

Previous methods	Models	Segmentation structure	DSC
Moghbel et al. [26]	EDCNN	Liver tumors	0.750
Foruzan and Chen [27]	Fuzzy c-means	Liver tumors	0.820
Wu et al. [28]	SVM	Liver tumors	0.830
Chlebus et al. [29]	FCM with graph cuts	Liver tumors	0.840
Jin et al. [30]	CNN	Liver tumors	0.830
Zhuofu Deng et al. [31]	3D U-Net	Liver tumors	0.850
Fallahpoor et al. [32]	DL model	Liver tumors on MRI	0.530
Su et al. [51]	DV-Net	Liver tumors	0.754
Shi et al. [52]	Extended Res-UNet	Liver tumors	0.792
Chen et al. [53]	FRA-UNet	Liver tumors	0.689
Li et al. [54]	Improved U-Net Model	Liver tumors on LiTs	0.690
Ozcan et al. [55]	AIM-Unet	Liver tumors	0.655
Hettihewa et al. [56]	MANet	Liver tumors	0.6400 ± 0.279
You et al. [57]	Contour-induced parallel graph reasoning	Liver tumors	0.741
Proposed	GAN and ACM	Liver tumors	0.872

use of three-dimensional region growth, which can be useful for managing medical treatment. The requirement for human labor is reduced by choosing seed points once and applying them to all future slices. In contrast to the traditional slice-by-slice method, which is time-consuming and ineffective, the initial setup for active contour generation involves minimal effort. On the MIDAS, LiTS, and 3Dircadb datasets, the proposed method obtained Dice similarity coefficients (DSC) of 0.908, 0.872, and 0.605, respectively. These results demonstrate the effectiveness of the proposed paradigm. The recommended approach also shows potential for use with other medical image modalities. It can help surgeons examine the tumor for prospective medical choices and therapy planning, leading to better patient care.

The given architecture has the potential to be easily extended to additional cutting-edge imaging modalities in addition to its relevance to liver tumor segmentation in CT volumes. Images of the brain, lungs, breast, and other anatomical areas for malignancy delineation are included, as are PET, CT, and 4D ultrasound studies. Additionally, the proposed strategy may be used to identify bone fractures in cost-effective X-ray images.

Funding No funding was received for conducting this study.

Data Availability The datasets generated during and/or analyzed during the current study are available in the LiTS – Liver Tumor Segmentation Challenge (LiTS17) organized in conjunction with ISBI 2017 and MICCAI 2017, Ircad Dataset (www.ircad.fr/research/3dircadb) and MIDAS Dataset (www.insight-journal.org/midas/collection/view/38).

Declarations

Conflict of interest The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Rana M, Bhushan M. Machine learning and deep learning approach for medical image analysis: diagnosis to detection. Multimed Tools Appl. 2023;82:26731–69.
2. Gupta M, Mishra A. A systematic review of deep learning based image segmentation to detect polyp. Artif Intell Rev. 2024;57:7. https://doi.org/10.1007/s10462-023-10621-1.
3. Wesdorp NJ, Zeeuw JM, Postma SCJ, et al. Deep learning models for automatic tumor segmentation and total tumor volume assessment in patients with colorectal liver metastases. Eur Radiol Exp. 2023;7:75. https://doi.org/10.1186/s41747-023-00383-4.
4. Ray S, Hagge R, Gillen M, Cerejo M, Shakeri S, Beckett L, Greasby T, Badawi RD. Comparison of two-dimensional and three-dimensional iterative watershed segmentation methods in hepatic tumor volumetrics. Med Phys. 2008;35(12):5869–81.
5. Hu Q, et al. Label-free liver tumor segmentation. In: 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, BC, Canada. 2023. p. 7422–32. https://doi.org/10.1109/CVPR52729.2023.00717.
6. Bilic P, Christ P, Li HB, Vorontsov E, Ben-Cohen A. The liver tumor segmentation benchmark (LiTS). Med Image Anal. 2023;84:102680. https://doi.org/10.1016/j.media.2022.102680.
7. Biswas A, Bhattacharya P, Maity SP, Banik R. Data augmentation for improved brain tumor segmentation. IETE J Res. 2021;69:2772–82. https://doi.org/10.1080/03772063.2021.1905562.
8. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, et al. Generative adversarial nets. In: Advances in neural information processing systems. 2014. p. 2672–80.
9. Zhou C, Kong S, Sun J. Review of generative adversarial networks in object detection. In: Liang Q, Wang W, Mu J, Liu X, Na Z, editors. Artificial Intelligence in China. AIC 2022. Lecture notes in electrical engineering, vol. 871. Singapore: Springer; 2023. https://doi.org/10.1007/978-981-99-1256-8_20.
10. Ma J, He Y, Li F, et al. Segment anything in medical images. Nat Commun. 2024;15:654. https://doi.org/10.1038/s41467-024-44824-z.
11. Gour M, Rajpoot R, Jain S. Histopathological image synthesis with generative adversarial networks for nuclei segmentation. SN Comput Sci. 2024;5:123. https://doi.org/10.1007/s42979-023-02440-6.
12. Armanious K, Jiang C, Fischer M, Küstner T, Hepp T, Nikolaou K, Gatidis S, Yang B. MedGAN: medical image translation using GANs. Comput Med Imaging Graph. 2020;79:101684. https://doi.org/10.1016/j.compmedimag.2019.101684.
13. Jeong JJ, Tariq A, Adejumo T, et al. Systematic review of generative adversarial networks (GANs) for medical image classification and segmentation. J Digit Imaging. 2022;35:137–52. https://doi.org/10.1007/s10278-021-00556-w.
14. Saad MM, O'Reilly R, Rehmani MH. A survey on training challenges in generative adversarial networks for biomedical image analysis. Artif Intell Rev. 2024;57:19. https://doi.org/10.1007/s10462-023-10624-y.
15. Li G, Wang J, Tan Y, et al. Semi-supervised medical image segmentation based on GAN with the pyramid attention mechanism and transfer learning. Multimed Tools Appl. 2024;83:17811–32. https://doi.org/10.1007/s11042-023-16213-z.
16. Gul S, Khan MS, Bibi A, Khandakar A, Ayari MA, Chowdhury MEH. Deep learning techniques for liver and liver tumor segmentation: a review. Comput Biol Med. 2022;147:105620. https://doi.org/10.1016/j.compbiomed.2022.105620.
17. Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. J Big Data. 2019;6:60. https://doi.org/10.1186/s40537-019-0197-0.
18. Kwiatkowski D, Dziubich T, et al. Comparison of selected neural network models used for automatic liver tumor segmentation. In: Abelló A, et al., editors. New trends in database and information systems. ADBIS 2023. Communications in computer and information science, vol. 850. Cham: Springer; 2023. https://doi.org/10.1007/978-3-031-42941-5_44.
19. Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. In: MICCAI. 2015. p. 234–41.
20. Milletari F, Navab N, Ahmadi SA. V-net: fully convolutional neural networks for volumetric medical image segmentation. In: 2016 Fourth International Conference on 3D Vision (3DV). 2016. p. 565–71.
21. Iqbal A, Sharif M, Yasmin M, Raza M, Aftab S. Generative adversarial networks and its applications in the biomedical image segmentation: a comprehensive survey. Int J Multimed Inf Retr. 2022;11(3):333–68. https://doi.org/10.1007/s13735-022-00240-x.
22. Ahmad W, Ali H, Shah Z, et al. A new generative adversarial network for medical images super resolution. Sci Rep. 2022;12:9533. https://doi.org/10.1038/s41598-022-13658-4.

23. Feng X, Lin J, Feng C, Lu G. GAN inversion-based semi-supervised learning for medical image segmentation. Biomed Signal Process Control. 2024;88(Part B):105536. https://doi.org/10.1016/j.bspc.2023.105536.
24. Siri SK, Kumar SP, Latte MV. Threshold-based new segmentation model to separate the liver from CT scan images. IETE J Res. 2020;68:4468–75. https://doi.org/10.1080/03772063.2020.1795938.
25. Lu M, Yaoyu T, Sihang B. Liver tumor segmentation based on 3D convolutional neural network with dual scale. J Appl Clin Med Phys. 2019;21:44–157. https://doi.org/10.1002/acm2.127841.
26. Budak U, Guo Y, Tanyildizic E, Sengur A. Cascaded deep convolutional encoder-decoder neural networks for efficient liver tumor segmentation. Med Hypotheses. 2020;134:109431.
27. Moghbel M, Mashohor S, Mahmud R, Iqbal Bin Saripan R. Automatic liver tumor segmentation on computed tomography for patient treatment planning and monitoring. EXCLI J. 2016;15:406–23.
28. Foruzan AH, Chen YW. Improved segmentation of low-contrast lesions using sigmoid edge model. Int J Comput Assist Radiol Surg. 2016;11(7):1267–83.
29. Wu W, Wu S, Zhou, Z, Zhang R, Zhang, Y. 3D Liver tumor segmentation in CT images using improved fuzzy C-means and graph cuts. BioMed Res Int. 2017;Article ID 5207685.
30. Chlebus G, Schenk A, Moltz JH, Ginneken BV, et al. Automatic liver tumor segmentation in CTwith fully convolutional neural networks and object-based postprocessing. Sci Rep. 2018;8(1):15497.
31. Jin Q, Meng Z, Sun C, Wei L, Su R. RA-UNet: a hybrid deep attention-aware network to extract liver and tumor in CT scans. 2018. http://arxiv.org/abs/1811.01328.
32. Fallahpoor M, Nguyen D, Montahaei E, et al. Segmentation of liver and liver lesions using deep learning. Phys Eng Sci Med. 2024. https://doi.org/10.1007/s13246-024-01390-4.
33. Lu F, Wu F, Hu P, Peng Z, Kong D. Automatic 3D liver location and segmentation via convolutional neural network and graph cut. Int J Comput Assist Radiol Surg. 2016;12:171–82. https://doi.org/10.1007/s11548-016-1467-3.
34. Li X, Chen H, Qi X, et al. H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes. IEEE Trans Med Imaging. 2018;37(12):2663–74. https://doi.org/10.1109/TMI.2018.2845918.
35. Melek Y, Durmus E. Artificial intelligence in radiation oncology. Artif Intell Med Imaging. 2021;2(2):13–31.
36. Sbai O, Couprie, C, Aubry M. Unsupervised image decomposition in vector layers. 2019. arXiv:1812.05484v2 [cs.CV].
37. Khaled A, Ghaleb TA. MRI-GAN: generative adversarial network for brain segmentation. In: Sheng B, Bi L, Kim J, Magnenat-Thalmann N, Thalmann D, editors. Advances in computer graphics. CGI 2023. Lecture notes in computer science, vol. 14495. Cham: Springer; 2024. https://doi.org/10.1007/978-3-031-50069-5_21.
38. Li Z, Wang Y, Yu, J. Brain tumor segmentation using an adversarial network. In: International MICCAI Brainlesion Workshop. Springer; 2017. p. 123–32.
39. Hemalatha RJ, Thamizhvani TR, et al. Active contour based segmentation techniques for medical image analysis. In: Medical and biological image analysis. 2018. Robert Koprowski, IntechOpen. https://doi.org/10.5772/intechopen.74576.
40. Villani C. Optimal transport: old and new, vol. 338. Springer Science & Business Media; 2008.
41. Arjovsky M, Chintala S, Bottou L. Wasserstein GAN. 2017. arXiv preprint arXiv:1701.07875.
42. Zuiderveld K. Contrast limited adaptive histogram equalization. In: Graphics gems IV. Academic Press Professional, Inc. 1994. p. 474–85.
43. Christ P, Ettlinger F, Grun F, Lipkova J. Kaissis G. Lits—liver tumor segmentation challenge. http://www.lits-challenge.com.
44. Ircad Dataset. Available from: www.ircad.fr/research/3dircadb.
45. Midas Dataset. Available from: www.insight-journal.org/midas/collection/view/38.
46. Salimans T, Goodfellow I, Zaremba W, Cheung V, Radford A, Chen X, Chen X. Improved techniques for training gans. In: Lee DD, Sugiyama M, Luxburg UV, Guyon I, Garnett R, editors. Advances in neural information processing systems, vol. 29. Curran Associates, Inc.; 2016. p. 2234–42.
47. Xun S, Li D, Zhu H, Chen M, Wang J, et al. Generative adversarial networks in medical image segmentation: a review. Comput Biol Med. 2022;140:105063. https://doi.org/10.1016/j.compbiomed.2021.105063.
48. Wang J, Peng Y, Jing S, et al. A deep-learning approach for segmentation of liver tumors in magnetic resonance imaging using UNet++. BMC Cancer. 2023;23:1060.
49. Zang L, Liang W, Ke H, et al. Research on liver cancer segmentation method based on PCNN image processing and SE-ResUnet. Sci Rep. 2023;13:12779. https://doi.org/10.1038/s41598-023-39240-0.
50. Nasrulloh AV, Willcocks CG, Jackson PTG, et al. Multi-scale segmentation and surface fitting for measuring 3-D macular holes. IEEE Trans Med Imaging. 2018;37(2):580–9.
51. Su J, Liu Z, Zhang J, Sheng VS, et al. DV-Net: accurate liver vessel segmentation via dense connection model with D-BCE loss function. Knowl Based Syst. 2021;232:107471. https://doi.org/10.1016/j.knosys.2021.107471.
52. Shi J, Kamata S-I. Extended Res-UNet with hierarchical inner-modules for liver tumor segmentation from CT volumes. In: 2022 4th International Conference on Robotics and Computer Vision (ICRCV), Wuhan, China. 2022. p. 169–74. https://doi.org/10.1109/ICRCV55858.2022.9953223.
53. Chen Y, Zheng C, Hu F, Zhou T, et al. Efficient two-step liver and tumour segmentation on abdominal CT via deep learning and a conditional random field. Comput Biol Med. 2022;150:106076. https://doi.org/10.1016/j.compbiomed.2022.106076.
54. Li H, Liang B. Liver tumor computed tomography image segmentation based on an improved U-Net model. Appl Sci. 2023;13(20):11283. https://doi.org/10.3390/app132011283.
55. Özcan F, Uçan ON, Karaçam S, Tunçman D. Fully automatic liver and tumor segmentation from CT image using an AIM-Unet. Bioengineering (Basel) 6. 2023;10(2):215. https://doi.org/10.3390/bioengineering10020215.
56. Hettihewa K, Kobchaisawat T, Tanpowpong N, et al. MANet: a multi-attention network for automatic liver tumor segmentation in computed tomography (CT) imaging. Sci Rep. 2023;13:20098. https://doi.org/10.1038/s41598-023-46580-4.
57. You Y, Bai Z, Zhang Y, Li Z. Contour-induced parallel graph reasoning for liver tumor segmentation. Biomed Signal Process Control. 2024;92:106111. https://doi.org/10.1016/j.bspc.2024.106111.