International Journal of Computational Intelligence and Applications
Vol. 22, No. 2 (2023) 2350006 (20 pages)

DOI: 10.1142/S1469026823500062

World Scientific
www.worldscientific.com

CT Images Segmentation Using a Deep Learning-Based Approach for Preoperative Projection of Human Organ Model Using Augmented Reality Technology

Nessrine Elloumi*

SETIT, ISBS, University of Sfax, Sfax, BP 1175 - 3000 Sfax, Tunisia
ellouminessrine@gmail.com

Aicha Ben Makhlouf† and Ayman Afli‡

Laboratory of Advanced Technology and Intelligent Systems (LATIS)
ENISO, University of Sousse, Sousse 4023, Sousse, Tunisia
†*aychaa.benmakhlouf@gmail.com*
‡*aflii.aymen@gmail.com*

Borhen Louhichi

Imam Mohammad Ibn Saud Islamic University
Riyadh 11432, Saudi Arabia

University of Sousse (ISSATSo, LMS)
4023, Sousse, Tunisia
blouhichi@imamu.edu.sa; borhen.louhichi@issatso.u-sousse.tn

Mehdi Jaidane

Sahloul University Hospital, Sousse 4011, Sousse, Tunisia
mjaidane@gmail.com

João Manuel R. S. Tavares

Instituto de Cincia e Inovaco em Engenharia
Mecnica e Engenharia Industrial
Departamento de Engenharia Mecnica, Faculdade de Engenharia
Universidade do Porto, Porto 4200-465, Portugal
tavares@fe.up.pt

Received 19 July 2022
Revised 21 September 2022
Accepted 3 January 2023
Published 5 June 2023

Over the last decades, facing the blooming growth of technological progress, interest in digital devices such as computed tomography (CT) as well as magnetic resource imaging which emerged in the 1970s has continued to grow. Such medical data can be invested in numerous visual recognition applications. In this context, these data may be segmented to generate a

*Corresponding author.

precise 3D representation of an organ that may be visualized and manipulated to aid surgeons during surgical interventions. Notably, the segmentation process is performed manually through the use of image processing software. Within this framework, multiple outstanding approaches were elaborated. However, the latter proved to be inefficient and required human intervention to opt for the segmentation area appropriately. Over the last few years, automatic methods which are based on deep learning approaches have outperformed the state-of-the-art segmentation approaches due to the use of the relying on Convolutional Neural Networks. In this paper, a segmentation of preoperative patients CT scans based on deep learning architecture was carried out to determine the target organ's shape. As a result, the segmented 2D CT images are used to generate the patient-specific biomechanical 3D model. To assess the efficiency and reliability of the proposed approach, the 3DIRCADb dataset was invested. The segmentation results were obtained through the implementation of a U-net architecture with good accuracy.

Keywords: Liver; CT scans; 3D model reconstruction; medical imaging; open surgery.

1. Introduction

With reference to World Health Organization (WHO), over 325 million people in the world are infected with hepatitis, that can trigger dangerous infections that may be severe a fatal health issue.[1] The liver is among the human body's intrinsic organs. Indeed, it ensures the body's correct cyclic functioning and performs a variety of tasks, including protein synthesis and metabolite detoxification. Owing to its significance, the liver is susceptible to a variety of pathologies, involving hepatitis B and C, which may beget severe illnesses like cancer or cirrhosis. Liver cancer corresponds to the frequently diagnosed cancer and the most lethal disease.[2] The growing death rates referring to such illnesses constituted the basic impetus driving doctors and the scientific community to focus on developing reliable and feasible ways to mitigate these growing rates. For decades, MIS techniques have been heavily used to treat these common infections. They involve the usage of a thin tube attached to an image camera that can be inserted into a small hole incision made in the body. The surgeon can take appropriate actions based on the visualized images obtained from the endoscopic camera. The major merit of MIS techniques[3] lies in reducing hemorrhaging, hence, mitigating the blood transfusions need. Furthermore, the tiny incisions limit the internal organ's exposure to any external pathogens, lowering the threat of infection and decreasing the hospitalization period. However, these operations are extremely challenging in terms of surgical competence, related to the lack of organ manipulation as well as the generally partial or inadequate visual input. In this respect, Augmented Reality technology corresponds to an innovative way to manipulate and visualize a 3D model through super-imposing interactively images generated by computer onto another real object. The likelihood to interact with virtual organs in real time, requires the projection and processing of medical data in real space. To generate a precise 3D model of a specific human organ, multiple methods have been set forward resting upon the acquired preoperative patient Computed Tomography (CT) scans.[4] The main step in the 3D model generation process of a specific human organ is the input data segmentation. Basically, segmentation techniques are performed manually through image processing tool.[5]

Numerous research works have been particularly oriented to the field of medical image detection in order to segment a specific area from a set of CT scan images. Yet, these methods proved to be defective and necessitated human intervention to select the segmentation area correctly.[6] Over the last decade, the deep learning methods grounded on Convolutional Neural Networks (CNN) have outperformed the literature methods in terms of multiple visual recognition applications. The typical use of CNN resides with a classification process, where the output is a simple set of labels to an input image. Meanwhile, in biomedical image processing, the segmentation process needs to include the localization of the organ.[7] The basic target of this paper is to reconstruct a digital model of a liver organ from a patient-specific CT scan. To reconstruct the liver model, a CNN model for liver segmentation is elaborated. Regarding preoperative data, i.e, CT scans, MRI images, the U-net model is used in order to explore the liver organ. Encouraging segmentation findings of the U-net architecture were applied and validated through satisfactory accuracy values. The segmented images will serve to generate a customized biomechanical model that represents the elastic and physical behavior of the liver organ. The generated model can handle elastic deformations thanks to its physical characteristics that are parametrized with Young's modulus coefficient and Poisson's ratio values.

2. State of the Art

The increasing death rates due to such diseases constituted the basic impetus driving doctors as well as the scientific community to focus on developing powerful solutions to decrease these growing rates. The technological development accomplished in the 80s generated the appearance of novel surgical techniques like Minimally Invasive Surgery (MIS), which relies on using an endoscope, which stand for a thin tube attached to an image camera that is put into the body through a minor incision or a natural hole. The surgeon can treat the patient and take appropriate action through observing the real-time video acquired by the endoscope camera. Thus, less damage to the body is incurred given that small incisions are used instead of the larger cuts which are often required in traditional surgery. Minimally Invasive Surgery lessens the infection risk and provides a faster recovery time. This accounts for the wide usage of the MIS procedure in surgical operations. However, the MIS technique has various deficiencies that can handicap the surgical process. The MIS basically needs precipitous learning curves, referring to the evolved technological traits. Furthermore, the invested instruments during the surgical operation have a limited range of motion, which results in limited visual output. Inspired by its implementation in the video game industry, Augmented Reality (AR) technology has been represented as the central focus of many researchers in these recent years. In fact, numerous studies examined the use of the augmented reality technology throughout surgical operations and assessed its implementation in the medical field. Augmented reality grants a clear and large visual output and offers general surgical knowledge, which further enhances the efficiency and accuracy of the surgical operation.

2.1. *Image pre-processing*

To present an optimal analysis of human abnormalities and diseases, numerous pretreatment approaches rest on preoperative patient images of the abdomen from magnetic resonance imaging (MRI) and CT scans. These images are intrinsic in order to detect the type and level of infection to plan radiotherapy or surgery, for example. Multiple filtering methods like geometric filtering, adaptive mean filtering, maximum *a posteriori* filtering, order-statistic filtering, adaptive weighted median filtering and nonlinear diffusion were applied to mitigate the noise effects triggered by the CT scanners.[8,9] The mean filtering substitutes each pixel with the mean value of its neighbors. It creates the allusion to image smoothing and blurring.[10] The adaptive mean filtering approach invests local picture statistics like mean, variance, and correlation in order to identify and retain edges and features.[11] Moreover, this filter mitigates the effect of noise by replacing it with the local mean value. Hence, it modifies the local image features and helps in terms of decreasing the noise from various regions of the input image.[12] Geometric Filtering minimizes noise while retaining key details. It uses a nonlinear, iterative algorithm to adjust the value of neighboring pixels grounded on their relative value.[13] Another approach relies on the usage of neuronal networks to process the input images, such as in Ref. 14. For color image processing based on the RGB color model, Lee *et al.* elaborated a CNN multilayer neural network structure. In this approach, each primary color is given to a separate CNN layer for parallel processing. Self-organizing maps (SOM)[15] are also widely used in this field, since they can extract the primary color content of the input images.[16]

2.2. *Liver segmentation*

With reference to recent works, volumetric images like CT and MRI scans were well employed in the field of medical analysis of liver diseases. Thus, the data produced by scanners, and the time needed to assess the resulting countless slices, have been growing swiftly, which makes the analysis of physicians and radiologists a more sophisticated task and a time-consuming process. Recent progress permeating image processing methods gave rise to the emergence of new solutions implementing deep learning models that target an accurate segmentation of liver. Multiple research works were centered around applying specific algorithms for upscaling to polish and brush up the noise filtering and image features.[17] Lu *et al.*[18] introduced an outstanding method for detecting the liver through eliminating parts that do not include liver organ from the CT images. However, the removal of the part is carried out manually, which presents certain limitations. CT images were divided automatically using CNN to detect regions of the liver area. Authors in Ref. 19 performed a method based on deep learning approach to select the brain in terms of multiple patch locations that are contrast-enhanced, namely, computed tomography scans of the brain, abdomen and pelvis which were classified into 10 categories. Concerning image classification, GoogLeNet and AlexNet architectures were trained and assessed to

classify natural images rather than medical ones. The method based on clustering rests upon the concept that, the similarity and distance between samples belonging to the same class are high and short, respectively, in n-dimensional feature space. Methods based on Clustering are automated and capable of doing multiple segmentations.[20] Fuzzy c-means was invested by Zhao *et al.* to carry out initial segmentation. This segmentation was later smoothed by morphological image processing techniques. Comparing with K-means, Fuzzy c-means proves to be efficient in terms of thoroughly segmenting the liver as well as its ability to enhance segmentation. Foruzan *et al.*[21] computed the starting border of the active contour using K-means algorithm. Hierarchical agglomeration clustering and self-organized maps are involved by other classification approaches for MRI images.[22] Nasrullah *et al.*[23] recorded a technique that invests a 2D CNN with a sliding window approach to locate lung nodules, after separating them through the use of the lung's 64 by 64 pixel sliding window with a 24-pixel overlay. The 64 by 64 window is fed into the 2D CNN classifier that applies binary classification to specify whether the window is positive or negative. De Vos *et al.*[24] revealed a technique for automatically localizing one or more anatomical features in 3D medical images by using a convolutional neural network to identify their presence in coronal, sagittal, and axial CT slices. The presence of interesting anatomical features in coronal, axial and sagittal slices extracted from 3D medical scans is detected using a convolutional neural network in each slice. Eventually, the neural network outputs are merged to generate 3D bounding boxes of the organ of interest. Christ *et al.*[25] introduced an automated segmentation technique through the use of cascaded fully convolutional neural networks (FCNs) to segment liver and lesions in CT and MRI abdominal images. This approach rests upon two FCNs: the Region of Interest (ROI) is obtained by the first FCNs, which separate the liver from the abdomen organ. The findings were then fed into the second FCN to identify the lesions. Using the 3DIRCADb dataset, the method accomplished a mean Dice of 93.1% for liver segmentation and 56% for tumor segmentation.

2.3. *Augmented reality technology*

Augmented Reality (AR) corresponds to one of the rising digital media technologies in view of its interactive context. The rise of the AR systems started in the 1990s, when Kasapakis *et al.*[26] established the concept of the reality–virtuality continuum to describe the conceivable combination of real and virtual objects. It characterizes the coexistence of real and virtual objects in the same virtual environment, as shown in Fig. 1.

Over the last decade, technological evolution in the medical sector has been conducive to the appearance of novel techniques like Minimally-invasive surgery, which assists surgeons in terms of manipulating organs in a straightforward manner investing accurate tools. This technique helps advanced technologies like augmented reality in operations rooms. Numerous researchers explored the usage of this

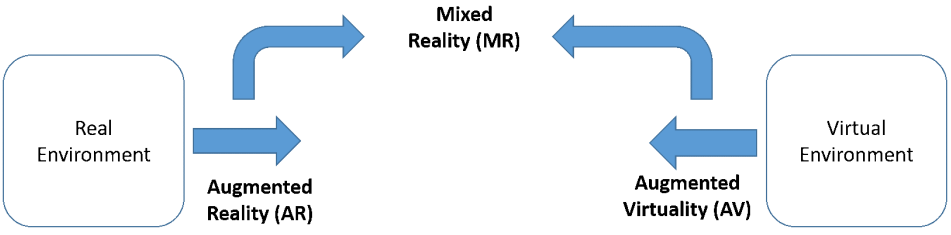

Fig. 1. Representation of the Reality–Virtuality continuum.

technology during surgical procedures, as well as its application in the medical area in general. AR provides a broader and global visual feedback and grants crucial surgical information, thus promoting the efficiency and accuracy of the surgical intervention. Indeed, one of the first image-guided surgery implementations was undertaken by Müller *et al.*[4] who evaluated the efficiency of a mobile computer-assisted Percutaneous Nephrolithotomy (PCNL). The iPad-based system applies a tracking server for image processing and a tablet to record the captured images, that are transferred back and forth with the server via Wi-Fi. The system confirmed its potential in terms of ensuring percutaneous interventions through its intuitive navigation process, which settled multiple shortcomings faced when using ultrasound and fluoroscopy techniques. In a performed experiment, their system proved its efficiency through accelerating the procedure by 20% compared to fluoroscopy. Although this approach yielded outstanding results, a scarcity of guidance information was detected and the absence of a real depth view begot certain obstacles in achieving the target areas in different cases, which justifies the high radiation exposure compared to the ultrasound technique. Additionally, this framework was checked on a prepared gelatin mixture, that is not distinguishable from kidney tissue. In the case of liver surgery, one of the thorny tasks that can be found during the Minimally invasive surgery is the partial surface view acquired by the laparoscopic camera, which confirms access to internal structures like tumors and vessels. To surmount this problem, Plantefve *et al.*[27] proposed an interesting approach based on non-rigid registration of preoperative and intraoperative data. The collected preoperative data served to erect a biomechanical model that characterizes the behavior of the organ. The registration process applies a three-dimensional point cloud generated from stereo-endoscopic camera-view. This pipeline comprises a semi-automatic liver segmentation and an automatic generation of the biomechanical model. To assess the feasibility and effectiveness of their approach, the researchers used two sets of data belonging to a healthy liver. The registration findings displayed a better exactitude through getting a mean error below 3 mm in almost all the cases. In spite of the goodness of fit as well as the robustness of this approach, using a high number of parameters might complicate the registration process and trigger latency in the overall pipeline. Besides, the relevant depth perception of the internal vascular network could enhance the visualization of the superimposed images. In the case of a

hyper-elastic organ, the liver is basically subjected to a large elastic deformation owing to pneumoperitoneum, i.e., CO_2 gas insufflation that constitutes a true challenge for an initial alignment.[28] Sahi *et al.*[17] addressed this issue by elaborating a method for the simultaneous estimation of the elastic transformation of the liver as well as the transformation of the endoscopic camera. Their approach rests upon an automatic detection of target area boundaries through segmenting the intraoperative 3D point cloud, which helps to generate the pneumoperitoneum model. The experiment on real *in-vivo* data of human hepatic surgery and synthetic data corroborated the relevance of this approach by obtaining lower mean errors compared to the existing method with a lower computational time, which makes their method substantially reliable and feasible during surgical operations. Regarding the validation process, a revision of the patient-specific parameters could be considered to enhance further the estimation process.

3. Proposed Approach

This work aimed to implement a novel approach to generate the digital model of human liver shape given as input a preoperative CT image, in order to model a 3D geometric model. The generated model will be projected in the real word using Meta 2 AR Headset. The developed approach in Fig. 2 starts through collecting the dataset and preparing the volumetric images to be fed into the CNN. The training process of the network results in a solution to segment a set of images of the liver organ that can be used as input for the 3D reconstruction step of the liver using InVesalius 3 software,[29] a medical 3D reconstruction tool for CT/MRI images. The reconstructed 3D model is exported in OBJ format to be manually post-processed in Blinder 3D tool in order to smooth any surface imperfections, if needed. For this reason, the final model can be used to perform a real-time projection within the Unity game engine using the Meta 2 AR headset. Nowadays, Deep learning techniques are the most frequently used techniques at the level of medical imaging approaches. CNNs have been invested to automatically divide CT scans to localize a search space for radiologists.

For the segmentation process, a deep learning technique grounded on the U-net architecture, which outputs the segmented images of the input liver slices from a set of volumetric images, was implemented.

The training approach relies upon several steps, as shown in Fig. 3. First, data preprocessing operations are implemented to the input set of CT scans images, which includes resizing them given as parameter the desired images size and implementing the windowing technique with Hounsfield Units measure to recalculate their pixels' intensity using the resized images. This step goes along with a separation between the CT images and their corresponding mask. Afterward, the training process is executed using the CNN model based on those sets of images to get the segmented images containing the liver.

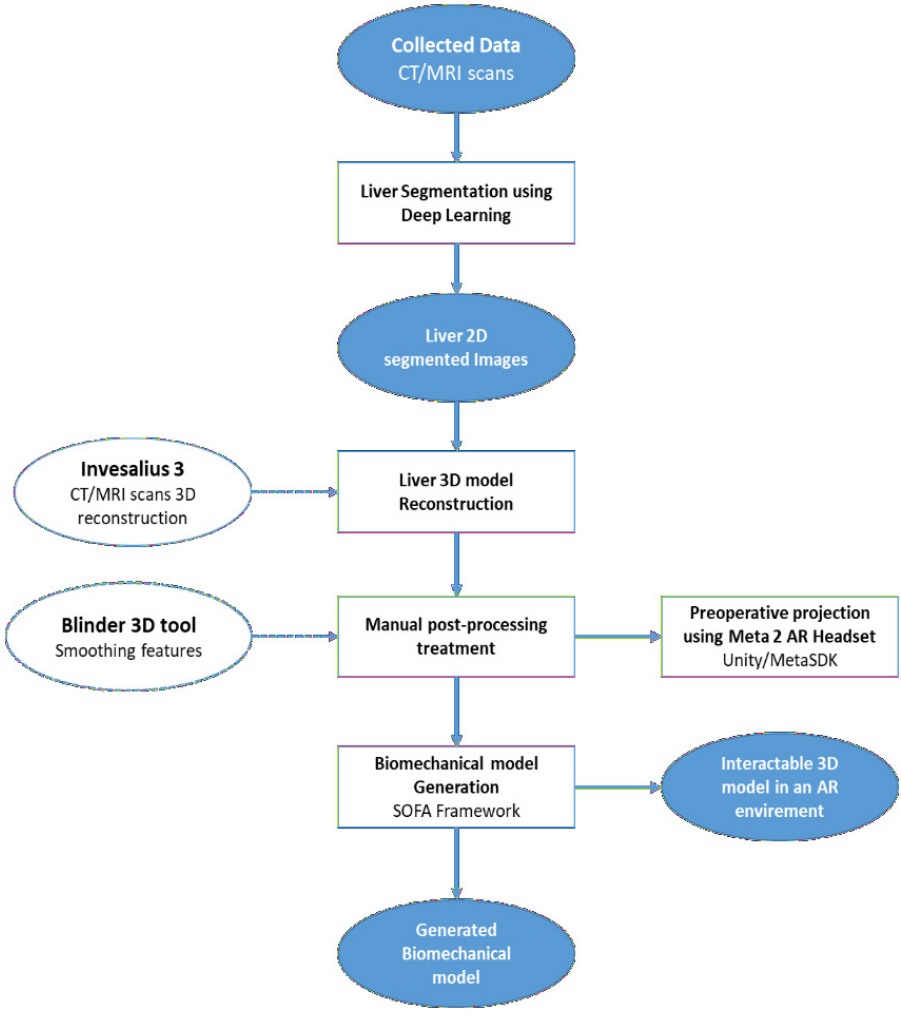

Fig. 2. Overview of the proposed approach.

The algorithm follows these steps:

(1) Calculating the mean for data centering, assuming as a parameter each resized image.
(2) Applying data normalization for each image.
(3) Calculating the weights of the best prediction using U-net architecture.
(4) Creating the mask for the training dataset (corresponding to the liver)
(5) Loading the training data.
(6) Loading the generated liver mask
(7) Matching the predicted images from the training process with the generated mask to obtain the segmented images.

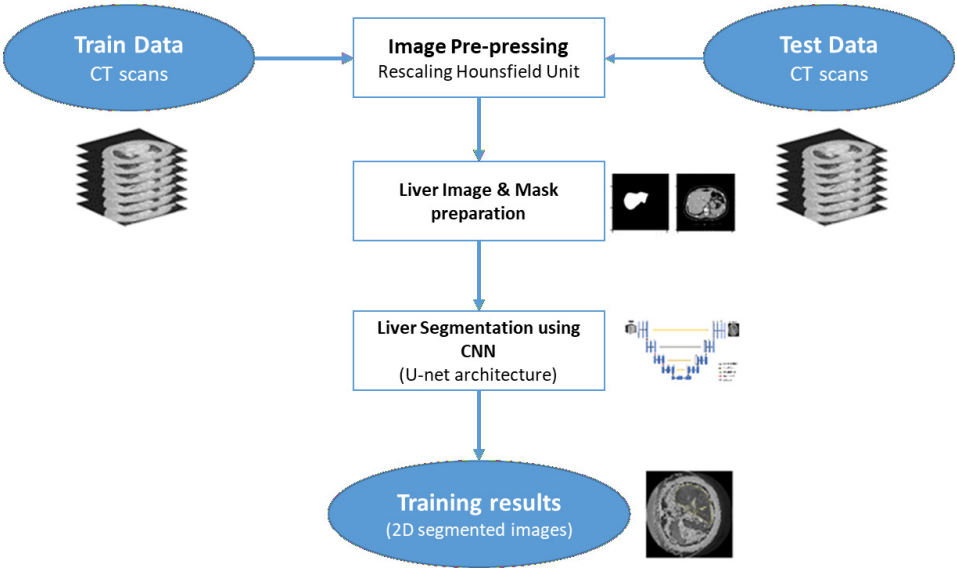

Fig. 3. Training flow diagram.

The U-net architecture, as it has been, effectively invested in the medical analysis community. CNN employs several convolutional layers in order to filter the inputs for meaningful information. These convolutional layers include learned parameters, allowing these filters to automatically modify the extraction of the most useful information for the available task at hand. Pooling layers are also used for limited translation and rotation invariance. Pooling in turn minimizes memory consumption, enabling additional convolutional layers to be used.

U-net was applied and devised in medical-image segmentation.[30] As illustrated in Fig. 4, U-net involves a symmetric path that applies convolution to rescale the obtained information to the input data. U-net returns a mask that slices the input image. In the case of medical images, each slice is used as 2D input information at each iteration until the full volume is cycled. Its architecture relies on an encoder network followed by a decoder network. The encoder network reduces the spatial dimensions in each layer while increasing the channels, whereas the decoder rises the spatial dimensions and reduces the channels. In Fig. 4, each blue box corresponds to a multichannel feature map. The number of channels is indicated at the top of the box. White boxes denote copied feature maps. The different operations are depicted in the arrows. The encoder on the left side rests on applying two 3×3 convolutions repeatedly. Each convolution is followed by ReLU and batch normalization. The spatial dimensions are then decreased using a 2×2 max pooling procedure. At each down sampling stage, the number of feature channels is doubled while cutting the spatial dimensions in halves. At the decoder side, each step consists of an up sampling of the feature map followed by a 2×2 transpose convolution, which cuts the number of feature channels in halves. Concatenation with the matching feature map

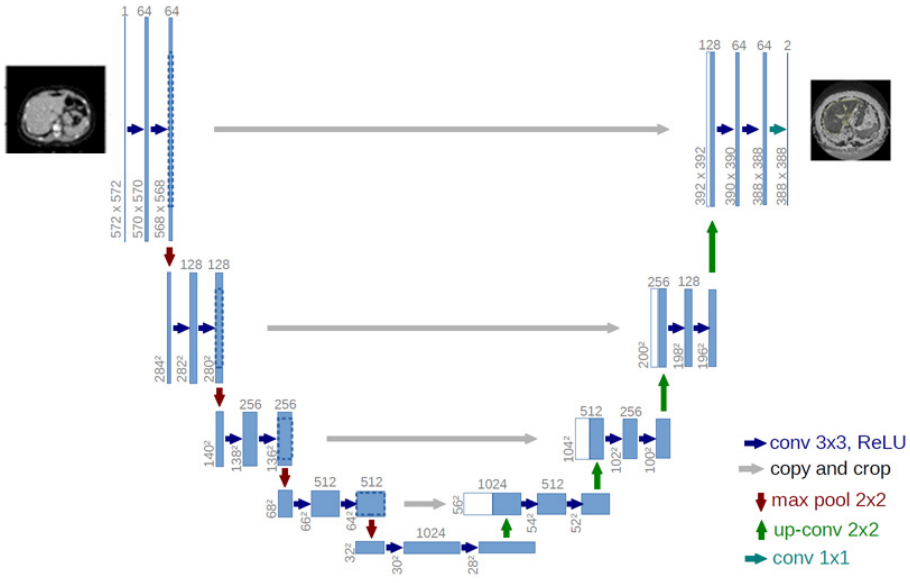

Fig. 4. U-Net architecture.

from the contracting path is also present, along with a 3×3 convolutional neural network (each followed by a ReLU activation function). A 1×1 convolution is employed in the final layer to translate the channels to the appropriate number of classes.

3.1. *Dataset preparing*

In order to assess the goodness of fit of the introduced approach, the 3DIRCADb dataset[31] was applied. This dataset contains livers of a wide diversity and complexity of lesions. It involves 10 men and 10 women computed tomography scans acquired from different European hospitals using various types of CT scanners. In this research work, the model is trained using volumetric CT scans of 13 patients, where each patient has a set of 116 slices, and each slice has a dimension of 512×512. The Nibabel package is used in order to convert the format of the input CT scans from the DICOM to NIfT format.

3.2. *Data pre-processing*

Owing to memory constraints during the training process, the dataset is downscaled. Slices that did not have any segmentation were collected to minimize substantial loss of liver information. This ensures that after rescaling, data segmentation is as dense as feasible. After importing and preparing the dataset, pre-processing operations were implemented to each slice. To exclude unnecessary organs, a CT windowing technique was invested. This technique computes the pixel intensity estimated by

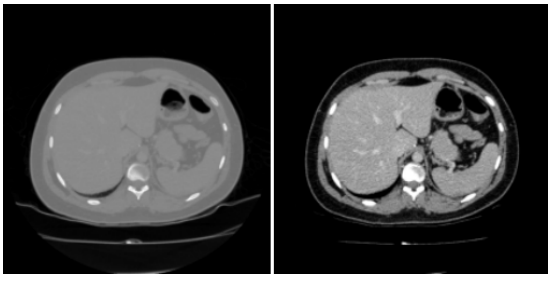

Fig. 5. Original raw input CT slice (on the left), and after the enhancement process (on the right).

the Hounsfield Units (HU)1, which stands for a function integrating the radio density, u, of a given substance.

$$\text{Hounsfield Units (HU)} = 1000\frac{\mu - \mu_{\text{water}}}{\mu_{\text{water}} - \mu_{\text{air}}}, \tag{1}$$

where μ, μ_{water} and μ_{air} express the linear attenuation coefficients of the subjects water, air and tissue, respectively. The Hounsfield unit values were windowed in the range [35, 350]. The result of the conducted pre-processing implemented on a raw medical slice is displayed in Fig. 5. In the training before feeding the model with the input images, the images were resized from 512×512 to 256×256 in order to speed up the process, given that smaller images usually yield better and faster training results.

The obtained enhanced image has higher contrast and the image becomes more detail. From Fig. 5, one can confirm that the liver region is highlighted and becomes easier to identify in the enhanced image than in the original one.

3.3. *3D model reconstruction*

The elaborated approach aims to generate a customized 3D model of a human organ and its projection in real world using AR headset technologies. From perspective, a generation of a visual mesh seems to be highly needed. In order to generate the visual mesh for the organ, the extracted data from the patients preoperative CT scan images have to be processed and reconstructed. Following the liver segmentation process, the output images were invested to produce a 3D model that can be projected in the preoperative visualization step, which was performed within InVesalius 3,[29] as in Fig. 6, which is an open-source tool dedicated to reconstruct 3D models from CT and MRI images, which provides manual and semi-automatic segmentation tools and grants the likelihood to import CT images, analyze and export files to OBJ, and other common formats.

3.4. *Biomechanical model generation*

Generating the biomechanical model was based on a set of physical-based parameters. After obtaining an input mesh from the segmented images, the SOFA[5]

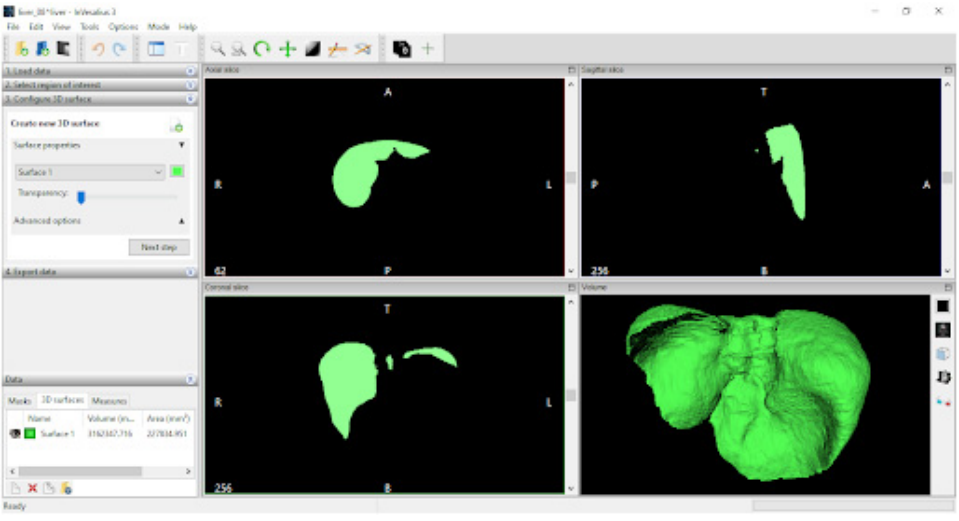

Fig. 6. Example of a liver 3D model generation within InVesalius 3.

(a) (b) (c)

Fig. 7. SOFA main components: biomechanical model (a), Collision model (b), Visual model (c).

framework was applied to represent and define the physical properties of the liver mesh. SOFA is a biomechanical simulation engine that provides a set of modeling and deformation methods. This framework is built upon three major components: visual, mechanical, and collision models — Fig. 7. Hence, SOFA provides a multi-modal method to manage the physical features and properties of the organs during the simulation process. The biomechanical model comprises a collection of physical-based parameters that correspond to the organs deformation properties. To represent the organ, several models, such as the finite element model or mass–spring systems, can be invested. The collision model serves basically to identify collisions between different models in the scene. Indeed, at each computing step and calculate their effect on each other. The third model is the visual model, which serves to display the deformable object. The three models are then the main components that define the deformable behavior of the liver model. We imported the reconstructed 3D model from InVesalius as a visual representation of the liver and determined the collision and biomechanical model's physical properties, as portrayed in Figs. 8 and 9.

The biomechanical model is the most significant component, serving to specify the physical properties of the object under analysis. To simulate the physical behavior of

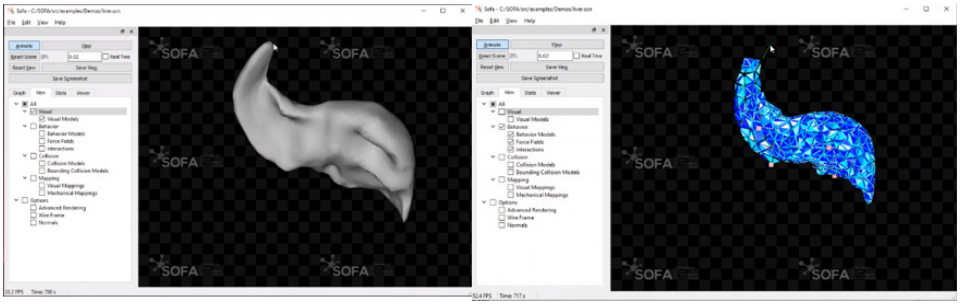

Fig. 8. Liver visual model (left) and mechanical model (right) (during deformation) within SOFA.

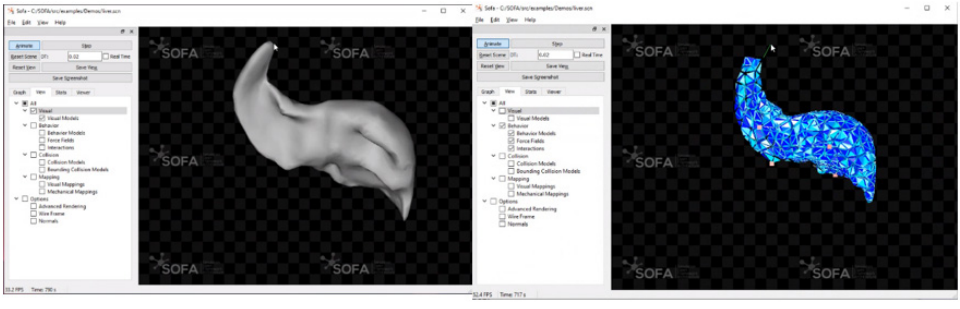

Fig. 9. Liver visual model (left) and mechanical model (right) (in rest shape position) within SOFA.

the liver, the Finite Element Method was used with a Poisson ratio of 0.3 and Young modulus of 30 KPa. This technique relies on a linear representation of the stress–strain relationship and allows significant displacements or rotations in the model, which provides an excellent compromise between behavior realism and computing efficiency.

4. Experiments

In the liver segmentation stage, the elaborated U-net architecture was implemented using Google Colab notebook in Python, the virtual GPU is running on Googles. In order to identify the parameters and layers of the neural network, TensorFlow and Keras libraries were invested. The model was trained for 20 epochs, taking about 1.5 h per patient, approximately. The weights were updated using batch-wise training, with a batch size of 10 after each epoch. Dice-coefficient and loss function were applied along with Adam optimizer with a learning rate of 0.0001 in order to update the weights. Each layer was followed by the Rectified Linear Unit (ReLU) and nonlinear activation function. The data were split 60% for training the network, 20% for the testing phase and 20% for the validation phase. Three metrics, which are commonly used in the field of the automatic medical image segmentation, were used to assess the precision of the proposed segmentation, :the Intersection over Union

(IoU), the Accuracy 3, and the Dice Similarity Coefficient (DSC). The Accuracy specifies semantic segmentation, and reports the percentage age of pixels in the images that were appropriately classified 2. The pixel accuracy is defined in terms of the proportion of True Positive, (TP) and True Negative (TN) according to the total number of cases,

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \tag{2}$$

On the other hand, DSC determines the similarity and the ratio between two sets of data (A and B).

$$\text{Dice Similarity Coefficient (DSC)} = 2\frac{B \cap A}{|B| + |A|}, \tag{3}$$

where A and B stand for two sets of input data. $|A|$ and $|B|$ refer to the number of elements in the A and B datasets, respectively, and $\cap$ represents the intersection of two datasets. The Intersection over Union IoU 4 is an evaluation metric serving to specify the accuracy of an object on a particular dataset, to evaluate the performance of classification methods. IoU is computed through dividing the area of overlap by the area of union between the expected segmentation and the ground truth:

$$\text{IoU}(A, B) = \frac{|A \cap B|}{|A \cup B|}, \tag{4}$$

where A and B express the prediction and ground truth bounding boxes. The intersection $|A \cap B|$ comprises the pixels found in the prediction mask and the ground truth mask. The union $|A \cup B|$ comprises all pixels found in the prediction or ground truth mask. To validate the proposed method, the IRCAD dataset was used. After training the model for 20 epochs (116 steps per epoch) on 13 patients with 1508 images, the effectiveness of the suggested segmentation method was tested on 6 patients with a total of 812 slices (116 slices per patient). To assess the performance of the proposed segmentation method, the Dice coefficient, the accuracy, and the IoU metrics are outlined in Table 1. The model performance is shown to be promising with a minimum value of 0.95 in the three-evaluation metrics. With the conducted experiments, the proposed method was tested on a set of patients separately for a more reliable and objective assessment, as in Table 2. Table 2 highlights the obtained results from six patients data. The goodness of fit of the model was proven through the outstanding findings indicated by the accuracy, IoU, and Dice values.

Table 1. Metrics obtained by the proposed method in the IRCAD dataset.

Dice	Accuracy	IoU
0.9787	0.9976	0.9586

Table 2. Segmentation metrics obtained by the proposed solution per patient.

Patient	Dice	Accuracy	IoU
Patient 1	0.9793	0.9977	0.9596
Patient 2	0.9726	0.9969	0.9470
Patient 3	0.9791	0.9976	0.9593
Patient 4	0.9717	0.9967	0.9969
Patient 5	0.9784	0.9975	0.9578
Patient 6	0.9725	0.9974	0.9482

(a) (b) (c)

Fig. 10. Results of the segmentation approach: Raw, i.e., input CT slice from the IRCAD dataset, liver region mask (b) and, i.e., segmented image.

An example of segmented images is presented in 10, along with the corresponding input raw CT slice, and liver mask. Figure 11 foregrounds the output images obtained by the proposed model, where the segmented liver region is marked in yellow. To weigh the reliability of the obtained results, a comparison between the elaborated method and related methods on the 3DIRCADb dataset was performed, as in Table 3. The obtained values for Dice Similarity Coefficient corroborated that the suggested approach outperforms the other methods reported in the literature. These results confirm that the developed implementation of the U-net architecture by adjusting its hyperparameters, such as the number of epochs and the batch size, could help achieve more accurate and precise results. Based on this study, it can be, suggested that the combination of a set of U-net architectures to tackle axial, coronal, and sagittal CT slices independently can be interesting to guarantee an even more accurate liver segmentation process. In addition, as the model robustness depends largely upon the type of the used dataset, a set of diverse datasets is required to validate the efficiency of the architecture. A custom generated set of CT scans prepared by our collaborators in the Hospital of Sahloul, in Sousse, Tunisia is planned to be used as input in our proposed segmentation approach.

The elaborated method proved to be successful in terms of segmenting the liver organ from real patient-specific CT scans obtained from the university hospital of Sahloul (Sousse, Tunisia). After training the model for 20 epochs (166 steps per epoch) on 20 patients with 2803 images in total, we evaluated its robustness by testing it on real CT scans of a single patient with a total of 95 slices (Fig. 12).

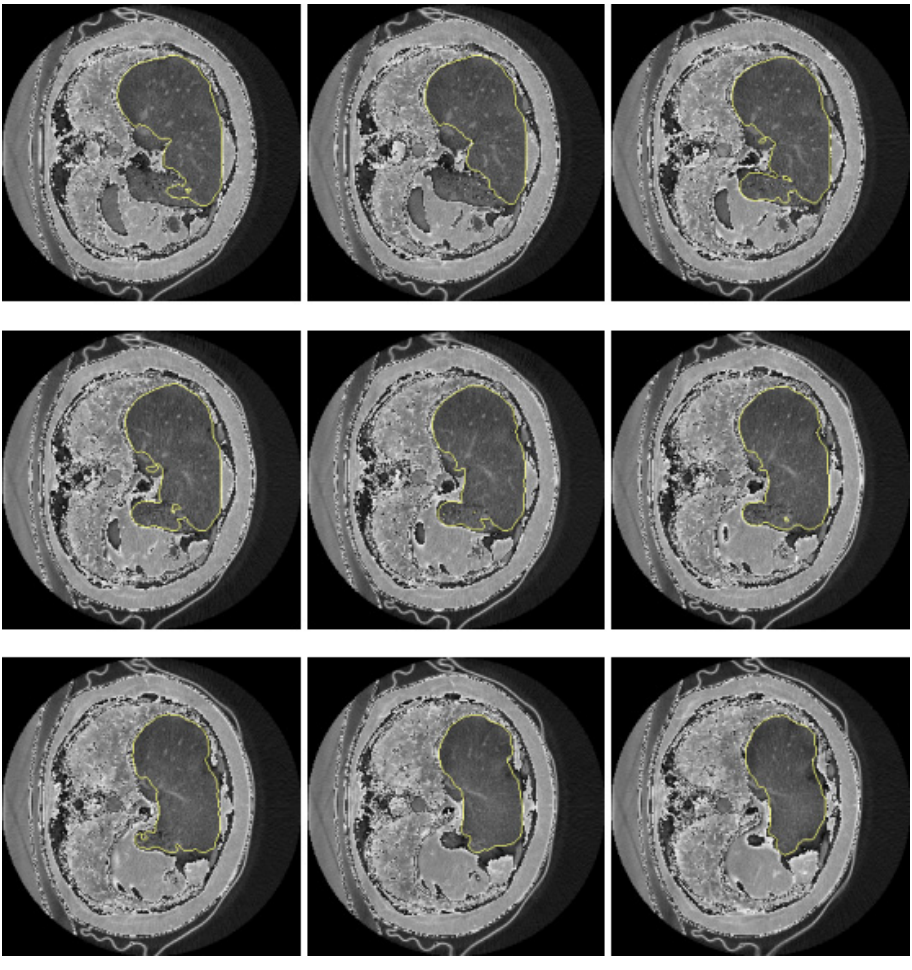

Fig. 11. Output testing images by the proposed model.

Table 3. Comparison between liver segmentation results obtained by the proposed method and those of related methods on the 3DIRCADb dataset.

Method	Dice (%)
Li *et al.*[32]	67.4
Affane *et al.*[33]	70.0
Alirr *et al.*[6]	91.8
Christ *et al.*[7] (Cascaded U-Net)	93.1
Christ *et al.*[7] (Cascaded U-Net + 3D CRF)	94.3
Ahmad *et al.*[34]	95.2
Budak *et al.*[35]	95.2
Proposed approach	**97.8**

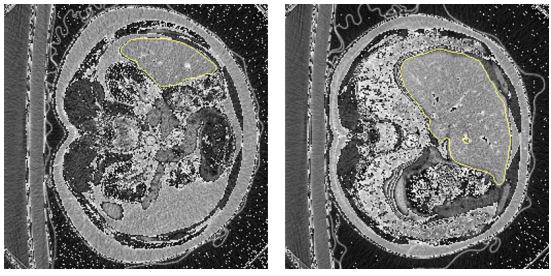

Fig. 12. Output images of a real data segmentation by the proposed model.

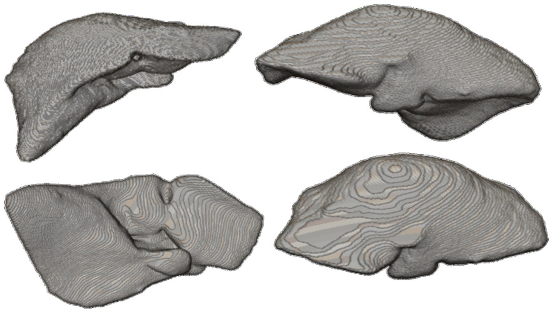

Fig. 13. Generated 3D model based on the segmentation results of real CT scans.

Table 4. Values obtained by the proposed method in real CT scans.

Dice	Accuracy	IoU
0.9772	0.9970	0.9555

Table 4 presents the qualitative evaluation of the developed segmentation method.

After extracting the liver region from the images, a 3D model of the organ is reconstructed within InVesalius3 software using the generated images from the segmentation step.

5. Conclusion

Augmented reality has the potential to significantly brush up the quality of medical interventions. Technical advances in the field of hepatic surgery, like visual feedback and localization of risky regions, aim to make the operation easier for surgeons and safer for patients. In this paper, a deep learning segmentation method was presented. Our target was to generate segmented 2D images of the liver organ that can be invested to reconstruct a specific biomechanical model of the patient. To validate the

proposed segmentation method, a liver dataset from IRCAD institute was used as input for the model. A pre-processing step was applied to enhance the input images, which therefore yields a better segmentation process. The segmented images, were, eventually, invested to generate a biomechanical model within the SOFA software. As a final note, it is noteworthy that though the obtained findings can be considered as worthwhile results, our research work is a step that may be built upon and taken further as it provides promising research directions and paves the way for further fruitful lines of investigations. Notably, in future research we may tackle deeper and fuller segmented images which serve to generate a patient-specific biomechanical model that represents the elastic and physical behavior of the liver organ. Additionally, the generated model would be promising in terms of handling elastic deformations thanks to its physical characteristics using Young's modulus and Poisson's ratio. At this state of analysis, this synthesis is valuable as it serves as an enlightening guideline for future researchers. In fact, although the liver is the central focus of this study, the elaborated model may be applied to other organs within the abdominal area, such as the kidneys which are potentially largely distorted. Thus, to amplify the scope of study and explore its obscure areas, an accurate deformable behavior is required through defining appropriate physical properties of the organ under analysis.

References

1. J. M. Paik, P. Golabi, Y. Younossi, A. Mishra and Z. M. Younossi, Changes in the global burden of chronic liver diseases from 2012 to 2017: The growing impact of NAFLD, *Hepatology* **72**(5) (2020) 1605–1616.
2. C.-H. Zhang, Y. Cheng, S. Zhang, J. Fan and Q. Gao, Changing epidemiology of hepatocellular carcinoma in Asia, *Liver Int.* **42**(9) (2022) 2029–2041, doi: 10.1111/liv.15251.
3. M. W. Gifari, H. Naghibi, S. Stramigioli and M. Abayazid, A review on recent advances in soft surgical robots for endoscopic applications, *Int. J. Med. Robot. Comput. Assist. Surg.* **15**(5) (2019) e2010.
4. M. Müller, M.-C. Rassweiler, J. Klein, A. Seitel, M. Gondan, M. Baumhauer, D. Teber, J. J. Rassweiler, H.-P. Meinzer and L. Maier-Hein, Mobile augmented reality for computer-assisted percutaneous nephrolithotomy, *Int. J. Comput. Assist. Radiol. Surg.* **8**(4) (2013) 663–675.
5. F. Faure *et al.*, "SOFA: A multi-model framework for interactive physical simulation, in *Soft Tissue Biomechanical Modeling for Computer Assisted Surgery* (Springer, 2012), pp. 283–321.
6. O. I. Alirr, Deep learning and level set approach for liver and tumor segmentation from CT scans, *J. Appl. Clin. Med. Phys.* **21**(10) (2020) 200–209.
7. P. F. Christ *et al.*, Automatic liver and tumor segmentation of CT and MRI volumes using cascaded fully convolutional neural networks, arXiv:1702.05970.
8. M. Masood, T. Nazir, M. Nawaz, A. Mehmood, J. Rashid, H.-Y. Kwon, T. Mahmood and A. Hussain, A novel deep learning method for recognition and classification of brain tumors from MRI images, *Diagnostics* **11**(5) (2021) 744.
9. V. A. Pimpalkhute, R. Page, A. Kothari, K. M. Bhurchandi and V. M. Kamble, Digital image noise estimation using DWT coefficients, *IEEE Trans. Image Process.* **30** (2021) 1962–1972.

10. J. Shedbalkar, K. Prabhushetty and A. Inchalc, A comparative analysis of filters for noise reduction and smoothening of brain MRIimages, in *2021 6th Int. Conf. Convergence in Technology (I2CT)*, IEEE, 2021, pp. 1–6.
11. M. M. Rahman, M. S. Rana, M. A. Islam, M. M. Rahman and M. H. Talukder, A new filtering technique for denoising speckle noise from medical images based on adaptive and anisotropic diffusion filter, *Int. J. Res. Comput. Commun. Technol.* **2**(9) (2013) 689–693.
12. C. Anam, K. Adi, H. Sutanto, Z. Arifin, W. Budi, T. Fujibuchi and G. Dougherty, Noise reduction in CT images using a selective mean filter, *J. Biomed. Phys. Eng.* **10**(5) (2020) 623.
13. Y. Tounsi, M. Kumar, A. Nassim, F. Mendoza-Santoyo and O. Matoba, Speckle denoising by variant nonlocal means methods, *Appl. Opt.* **58**(26) (2019) 7110–7120.
14. S. Akcay and T. Breckon, "Towards automatic threat detection: A survey of advances of deep learning within X-ray security imaging," *Pattern Recogn.* **122** (2022) 108245.
15. A. Saxena, M. Prasad, A. Gupta, N. Bharill, O. P. Patel, A. Tiwari, M. J. Er, W. Ding and C.-T. Lin, A review of clustering techniques and developments, *Neurocomputing* **267** (2017) 664–681.
16. S. H. Ong, N. Yeo, K. Lee, Y. Venkatesh and D. Cao, Segmentation of color images using a two-stage self-organizing network, *Image Vis. Comput.* **20**(4) (2002) 279–289.
17. K. Sahi, S. Jackson, E. Wiebe, G. Armstrong, S. Winters, R. Moore and G. Low, The value of "liver windows" settings in the detection of small renal cell carcinomas on unenhanced computed tomography, *Canad. Assoc. Radiol. J.* **65**(1) (2014) 71–76.
18. F. Lu, F. Wu, P. Hu, Z. Peng and D. Kong, Automatic 3D liver location and segmentation via convolutional neural network and graph cut, *Int. J. Comput. Assist. Radiol. Surg.* **12**(2) (2017) 171–182.
19. H. Sugimori, Classification of computed tomography images in different slice positions using deep learning, *J. Healthcare Eng.* **2018** (2018) 1–9, https://doi.org/10.1155/2018/1753480.
20. S. K. Siri, P. S. Kumar and M. V. Latte, An improved expectation-maximization algorithm to detect liver image boundary in ct scan images, *IETE J. Res.* (2022) 1–9.
21. A. H. Foruzan *et al.*, Segmentation of liver in low-contrast images using k-means clustering and geodesic active contour algorithms, *IEICE Trans. Inf. Syst.* **96**(4) (2013) 798–807.
22. D. Chi, Y. Zhao and M. Li, Automatic liver MR image segmentation with self-organizing map and hierarchical agglomerative clustering method, in *2010 3rd Int. Congress on Image and Signal Processing* (IEEE, 2010), pp. 1333–1337.
23. N. Nasrullah, J. Sang, M. S. Alam, M. Mateen, B. Cai and H. Hu, Automated lung nodule detection and classification using deep learning combined with multiple strategies, *Sensors* **19**(17) (2019) 3722.
24. B. D. De Vos, J. M. Wolterink, P. A. de Jong, T. Leiner, M. A. Viergever and I. Išgum, Convnet-based localization of anatomical structures in 3D medical images, *IEEE Trans. Med. Imag.* **36**(7) (2017) 1470–1481.
25. P. F. Christ *et al.*, Automatic liver and tumor segmentation of CT and MRI volumes using cascaded fully convolutional neural networks, arXiv:1702.05970.
26. V. Kasapakis, D. Gavalas and E. Dzardanova, Creating room-scale interactive mixed-reality worlds using off-the-shelf technologies, in *Int. Conf. Advances in Computer Entertainment* (Springer, 2017), pp. 1–13.
27. R. Plantefeve, I. Peterlik, N. Haouchine and S. Cotin, Patient-specific biomechanical modeling for guidance during minimally-invasive hepatic surgery, *Ann. Biomed. Eng.* **44**(1) (2016) 139–153.

28. F. Sánchez-Margallo, J. Moyano-Cuevas, R. Latorre, J. Maestre, L. Correa, J. Pagador, L. Sánchez-Peralta, J. Sánchez-Margallo and J. Usón-Gargallo, Anatomical changes due to pneumoperitoneum analyzed by MRI: An experimental study in pigs, *Surg. Radiol. Anatomy* **33**(5) (2011) 389–396.
29. P. Amorim, T. Moraes, J. Silva and H. Pedrini, InVesalius: An interactive rendering framework for health care support, in *Int. Symp. Visual Computing* (Springer, 2015), pp. 45–54.
30. O. Ronneberger, P. Fischer and T. Brox, U-Net: Convolutional networks for biomedical image segmentation, in *Int. Conf. Medical Image Computing and Computer-Assisted intervention* (Springer, 2015), pp. 234–251.
31. X. Li, H. Chen, X. Qi, Q. Dou, C.-W. Fu and P.-A. Heng, H-DenseUNet: hybrid densely connected U-Net for liver and tumor segmentation from ct volumes, *IEEE Trans. Med. Imag.* **37**(12) (2018) 2663–2674.
32. X. Li, R. Bala and V. Monga, Robust deep 3D blood vessel segmentation using structural priors, *IEEE Trans. Image Process.* **31** (2022) 1271–1284.
33. A. Affane, A. Kucharski, P. Chapuis, S. Freydier, M.-A. Lebre, A. Vacavant and A. Fabijanska, Segmentation of liver anatomy by combining 3D U-Net approaches, *Appl. Sci.* **11**(11) (2021) 4895.
34. M. Ahmad, D. Ai, G. Xie, S. F. Qadri, H. Song, Y. Huang, Y. Wang and J. Yang, Deep belief network modeling for automatic liver segmentation, *IEEE Access* **7** (2019) 20585–20595.
35. Ü. Budak, Y. Guo, E. Tanyildizi and A. Şengür, Cascaded deep convolutional encoder-decoder neural networks for efficient liver tumor segmentation, *Med. Hyp.* **134** (2020) 109431.