Aswan University Journal of Sciences and Technology Volume 3, issue 1, June 2023

Medical Image Segmentation Using Deep Learning: Review

Walaa M. Abd-Elhafiez[1], Aida H. Abu bakr[2], E. A. Zanaty[3] and M. E. Hussein[2]

[1]*Computer Science Department, Faculty of Computers and Artificial Intelligence, Sohag University, Sohag, Egypt.*

[2]*Mathematics Department, Faculty of Science, Aswan University, Aswan, Egypt.*

[3]*Information Technology Department, Faculty of Computers and Artificial Intelligent, Sohag University, Sohag, Egypt.*

Abstract— Medical image segmentation has greatly increased healthcare sustainability. It is currently a major research area in the field of computer vision. The many artefacts inherent in medical images make it a difficult process to segment them. Deep neural models have recently demonstrated their use in a variety of image segmentation applications. With the rapid advancement of deep convolutional neural networks, medical image processing has become a study hotspot development of deep learning. The primary focus of this study is the deep learning-based segmentation of medical images.. This study gives an overview of the literature in the area of deep convolutional neural network-based medical image segmentation. The article looks at several popular medical image datasets, several segmentation task evaluation measures, and the effectiveness of various CNN-based networks. The current study also examines several issues in the area of segmenting medical images and various state-of-the-art solutions accessible in the literature, in contrast to the existing survey and review articles. The paper has several contributions which are as follows: Firstly, the present study provides an overview of the current state of the deep neural network structures utilized for medical image segmentation. Secondly, the paper describes the publicly available techniques of Network Training. Thirdly, it presents the various performance metrics employed for evaluating the deep learning segmentation models and the medical image segmentation datasets. Finally, the paper also gives an insight into the major challenges faced in the field of image segmentation and their state-of-the-art solutions

Keywords: — Medical image, Segmentation, Deep learning, CNN.

1. Introduction

One of the most difficult problems in medical image analysis is medical image segmentation, which involves separating organ or lesion-related pixels from background medical images like MRI or CT scans in order to provide crucial information on the shapes and sizes of these organs. The process of segmenting a image entails dividing it into sections that are strongly correlated along with the original image's area of interest (RoI). The goal of medical image segmentation is to provide an input image in a meaningful format for anatomical study, identifying the region of interest (RoI), measuring the volume of tissue to determine the size of the tumor, and assisting with calculating the radiation dose, planning the course of treatment before applying radiation therapy [1]. Numerous researchers have suggested various automated segmentation systems using currently available technologies. Early systems were built using standard methods, such as edge recognition filters and mathematical methods. Then, for a considerable amount of time, machine learning techniques for extracting manually created characteristics were the norm. The primary focus of creating such a system has always been designing and extracting these properties, and the complexity of these approaches has been seen as a substantial barrier to their use. Deep learning approaches entered the scene in the 2000s as a result of hardware advancement and began to show off their impressive powers in image processing applications.Deep learning techniques have emerged as the top choice for image segmentation, especially for medical image segmentation, due to their promising capabilities. Deep learning-based image segmentation has attracted a lot

Corresponding author E-mail: aida.hessain@aswu.edu.eg

Received April 11, 2023, received in revised form April 23, 2023, accepted April 28, 2023.

(ASWJST 2023/ printed ISSN: 2735-3087 and on-line ISSN: 2735-3095) https://journals.aswu.edu.eg/stjournal

of attention, especially over the past several years, which emphasises the need for a thorough analysis of it. To the best of our knowledge, there hasn't been a a comprehensive analysis of medical image segmentation based on deep learning. [2, 3]

At first, segmentation was carried out using mathematical models (such as Fuzzy C-mean clustering and K-means clustering) and basic image processing (such as edge detection techniques and watershed) [4]. In order to create a model for segmentation, At the turn of the century, methods for pattern detection and machine learning (ML) were employed. These supervised techniques continue to be widely used, and many commercial programmes for medical image analysis are built on them.In these methods, the extracted characteristics are created or chosen by researchers and are referred to as handcrafted features. They are typically understood by humans, however they might not be the best attributes for segmentation. Recently, computers have been able to extract representative features from images thanks to the advancement of deep learning and neural networks (DL).The foundation of many deep neural networks used for image processing is the idea that by learning high-level properties, networks with many layers can transform input pictures into output labels. The convolutional neural network is the most common type of image processing model to date (CNN). The CNN, a significant model, has addressed numerous important commercial applications and demonstrated its aptitude in numerous competitions. At numerous workshops, competitions, and conferences, CNN and other DL methods in medical image analysis began to demonstrate their effectiveness [5].

Several review articles are attempting to condense these apps since the number of studies grows quickly. The journals [6, 7] have written comprehensive reviews on DL in medical image analysis. Review papers on DL-based medical image segmentation have just been presented by [8–11]. Some review articles have more specialised objectives. For instance, the work in [12, 13] summarised cardiac image segmentation networks; [14] covered medical imaging's use of generative adversarial networks (GANs); [15] examined the biomedical image analysis limited sample issue; [16] looked into the issue of noisy labels in medical image analysis; [17] looked into DL solutions for medical image segmentation with imperfect datasets.; and [18] surveyed not-sosupervised networks using semi-supervised, multi-instance, and Medical image segmentation high-level prior-based loss functions are categorised according to the prior's properties. in [19] as shape, size, topology, and inter-region constraints. Shape-constrained DL for medical image segmentation was reviewed in [20]. Some review articles concentrate on certain segmentation tasks in particular modalities. Artificial intelligence in stroke imaging was discussed in [21]. While [23] concentrated on Segmenting tumors of liver in computed tomography (CT) images, [22] reviewed DL techniques for isointense baby brain segmentation in MRI. [24] investigated how organs at risk were segmented for head and neck radiation planning. [25] provided a summary of the segmentation techniques used on the cartilage of the knee. Brain tumour segmentation is a hot topic, hence [26-30] published a number of review papers on the subject. Small rodent brain segmentation techniques were reviewed in [31], and brain segmentation techniques from multi-atlas to DL were reviewed in [32]. [33] summarised the methods for segmenting the thyroid lobe and thyroid nodules for ultrasound imaging. [34, 35] have recently published two reviews on the segmentation of retinal blood vessels. We can see These review papers show that there are many potential directions for this type of research and that various review focuses have different implications.

Algorithms for image segmentation based on deep learning have seen success in the field of image segmentation due to the rapid growth of artificial intelligence, particularly DL (DL) [36-38]. Deep learning provides some advantages over conventional machine learning and computer vision techniques in terms of segmentation accuracy and speed. The application doctors effectively by verifying the extent of a diseased tumors by using deep learning to segment medical images., comparing the impact before and after treatment quantitatively and reducing their workload considerably.

2. Structures of Deep Neural Network

a majority crucial method for artificial intelligence is deep learning. A deep learning technique builds an artificial neural network using a number of layers. There are three layers in input, buried layer(s), and output of an artificial neural network (ANN). The information is gotten by the input layer of the network, the output layer decides how to use the input, and hidden layers between the input and output layers carry out computations (shown in Figure 1). Between the input and output layers, a deep neural network has numerous hidden layers.

A review of various deep learning neural networks used for image segmentation is provided in this section. As illustrated in Figure 2, the various deep neural network topologies typically used for image segmentation can be categorised.

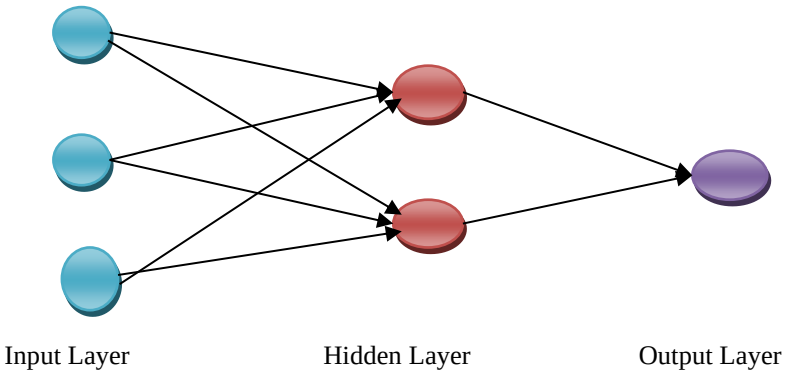

Figure 1:Artificial neural network (ANN) model.

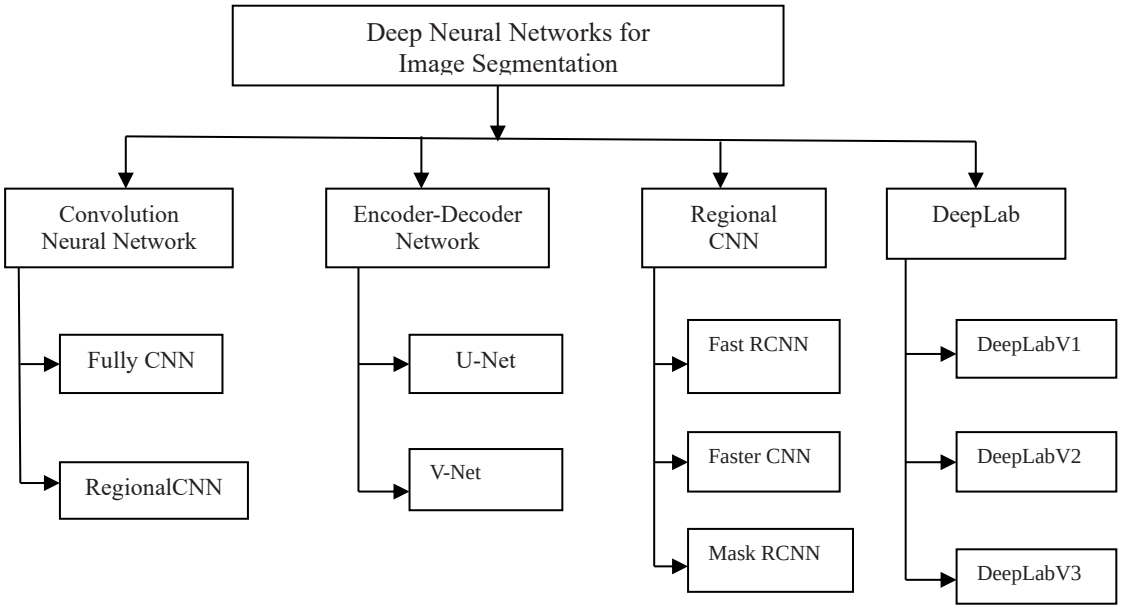

Figure 2: Different types of deep neural network architectures for image segmentation

2. 1. Convolutional Neural Network

Convolutional layer, pooling layer, and fully connected layer are the three basic neural layers that make up Convolutional neural network, or CNN (see Figure 3). Each layer plays a distinct role. The convolution layer recognises unique aspects in an image, such as edges or other visual components. With kernels, the convolution layer multiplies a pixel's immediate neighbours to produce an image. For the purpose of creating its feature maps, CNN convolves the given image using several kernels. The input data's spatial measurements (width, height) are decreased by the pooling layer for the next layers of the neural network.The depth of the data remains unchanged. Subsampling is the name of this procedure. The computing demands on subsequent layers are reduced by this size reduction. In NN, high-level reasoning is done by the fully linked layers. The final results are provided by these layers, which combine the numerous feature responses from the provided input image.Numerous CNN models, including as AlexNet [39], GoogleNet [40], VGG [41], Inception [42], SequeezeNet [43], and DenseNet [44], have been published in the literature. Each network in this situation employs a distinct count of pooling and convolutional layers, with significant process blocks positioned in between. The majority of the time, CNN models are used for classification tasks. SqueezeNet and GoogleNet have been used to divide three kinds of brain MRI images in [45]. The following factors limit the performance of the CNN segmentation models:

i. CNN's completely connected layers are unable to handle varying input sizes.
ii. An object segmentation convolutional neural network with a fully connected layer cannot be used because the number of items of interest in the image segmentation job is not constant. As a consequence, the output layer's length cannot be constant.

2.1.1. Fully Convolutional Network

Only convolutional layers exist in a fully convolutional network (FCN). The last layer of fully connected of a CNN can be changed into a fully convolutional layer to transform the various existing CNN architectures into FCN. Instead of making patch-wise forecasts, the model created by [46] can produce a full-size input picture and dense pixel-wise predictions from a spatial segmentation map. The model makes use of skip connections, which upsample feature maps from the top layer and fuse them with feature maps from lower layers. Thus, the model generates a thorough segmentation in a single pass. However, the typical FCN paradigm includes the following drawbacks[47]:

It accepts global context data poorly and is too slow for real-time inference. The output feature maps' resolution is reduced as a result of propagation across various convolution and pooling levels in FCN. FCN forecasts consequently have a poor resolution and fuzzier object boundaries.

ParseNet [48] is a reported example of a sophisticated FCN that uses global average pooling to achieve global context. It has also been reported on methods for integrating models like DL design incorporates Markov random fields and conditional random fields.

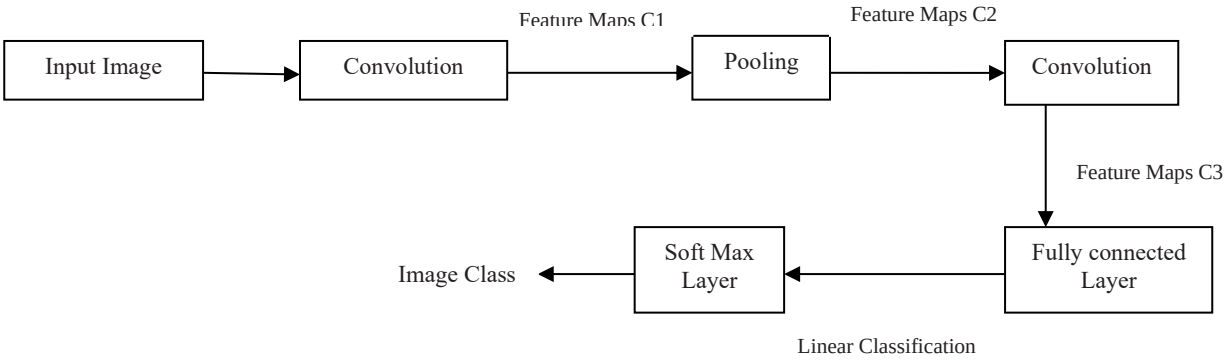

Figure 3: Architecture of Convolutional neural network.

2.2. Encoder-Decoder Models

In order data elements must be mapped from the input domain to the output domain., encoder-decoder based models use a two-stage paradigm. The encoder step transforms the input, x, into a representation of latent space., which the decoder then uses to forecast the output. The following sections address the various Typical encoder-decoder-based segmentation algorithms for medical images include the following:

2.2.1. U-Net.

There are downsampling and upsampling components in the U-Net model [49]. The FCN-like architecture used in the downsampling portion using 3 3 convolutions for extracting features to capture context. To reduce the amount of computed feature maps, deconvolution is performed during the upsampling phase. To prevent any information loss, the maps of feature produced by the downsampling are provided as input for the part of upsampling. Precise localisation is provided by the symmetric upsampling component. The map of segmentation is created by the model that assigns every pixel of the image a category.

The following benefits are provided by the U-Net model:

i. U-Net model can segment images well with a small number of tagged training images.
ii. To forecast a fair segmentation map, Location data from the downsampling route and contextual data from the upsampling channel are combined using the U-Net architecture..

A few restrictions that apply to U-Net models are following:

i. The maximum size for the input image is 572 x 572.
ii. The learning typically slows down in the intermediate layers of deeper UNET models, causing the network to ignore the characteristics in the layers with abstract layers.
iii. Connection skips in the model enforce a constrained fusion approach that results in the accumulation of feature maps with the same scale in the encoder and decoder networks.

As remedies for these drawbacks, the U-Net++ [50], Attention U-Net [51], and SD-UNet [52] versions of the U-Net design have all been proposed in the literature..

2.2.2. VNet.

The segmentation of medical images is also done using an FCN-based model [53]. Compression and decompression networks are the two components of the VNet architecture. At each stage of the compression process, convolution layers with residual function are used. The volumetric kernels were used in these convolution layers. The decompression network enhances low resolution feature maps' spatial depiction by extracting features. Both the foreground and the background regions are given two-channel probabilistic segmentation.

2.3. Regional Convolutional Network

The task of object recognition and segmentation has been carried out using a regional convolutional network. Using a selective search technique, the R-CNN design described in [54] creates region proposal networks for bounding boxes. These regional suggestions are sent to a CNN, which will yield a feature vector map after being distorted to standard squares. To categorize the objects contained in the region proposal network, features from the image are extracted into the output dense layer and input into a classification algorithm. To increase the precise suggestion for the region's level or bounding box, the program also forecasts the offset values. The operational steps of In Figure 4, the R-CNN design is displayed.. The employment of the fundamental model of RCN is constrained by the next factors:

i. Because it takes the network about 47 seconds to train for a classification assignment involving 2000 area suggestions in a test image, it cannot be used in real time.
ii. A predetermined method was used for selective search. As a result, no learning is happening at that time. This could lead to the establishment area of unfavourable candidate suggestions.

The literature has suggested various R-CNN variants, such as fast R-CNN, faster R-CNN, and mask R-CNN, to overcome these disadvantages.

2.3.1. Fast R-CNN

The proposed regions of the image overlap in R-CNN are carried out repeatedly along with the identical CNN computations. An input image and a list of object suggestions are fed into the rapid R-CNN described by [55]. Following that, the CNN produces convolutional feature maps. Each object suggestion is then reshaped into a vector of feature with a defined size by the ROI pooling layer. The model's final fully linked layers get the feature vectors. Finally, the Softmax layer receives the ROI feature vector calculated in order to predict the proposed region's class and offset values [56]. The fast R-CNN is delayed because a selective search algorithm is being used.

2.3.2. Faster R-CNN

The suggested regions were developed through a prolonged, selective search procedure in R-CNN and fast R-CNN. A single convolutional network was therefore used in the faster R-CNN architecture described by [57] to complete the region proposal and classification tasks. A region proposal network (RPN) is used by the model to pass the CNN feature map's sliding pane at the top as a whole. K different possible boundary boxes are generated for each window, with each box's score corresponding to the position of the object. The precise categorization boxes are produced by Rapid R-CNN received these bounding frames.

2.3.3. Mask R-CNN

He and his colleagues expanded quicker R-CNN will show Mask R-CNN for example segmentation in [58]. The model is capable of identifying an image's components and creating a superior segmentation mask for every item. To preserve the precise spatial coordinates of the provided image, RoI-Align layer is used. Using a CNN, the region proposal network (RPN) produced several RoIs. Multiple warped bounding boxes with fixed size are produced by the RoI-Align network. using a softmax layer to generate classification, the previously calculated warped features are sent to a fully linked layer. The model comprises three branches of output, the first of which computes the bounding box coordinates, the second of which establishes the related classes, and the third of which assesses the binary mask for each RoI. The model unifies the training for all branches. Regression modelling is used to enhance the bounded boxes. For each RoIa binary mask is produced by the mask algorithm.

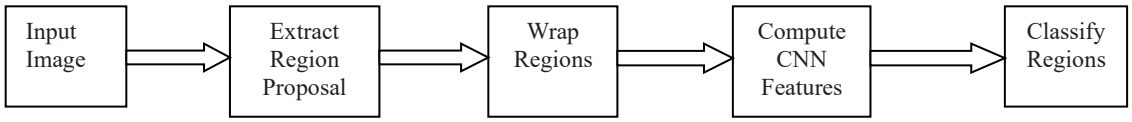

Figure 4: R-CNN architecture

2.4. DeepLab Model.

For extracting the features from an image, the DeepLab model uses the ResNet-101/VGG-16 is a pretrained CNN model with arous convolution [59]. When atrous convolutions are used, the following benefits occur:

i. It regulates how CNNs' feature replies are resolved.
ii. It does so without requiring the learning of any additional parameters, turning the a dense feature extractor from image classification network.
iii. Uses conditional random field (CRF) to create output that is finely segmented.

In the literature, a number of DeepLab variations DeepLabv1, DeepLabv2, DeepLab3, and DeepLabv3+ have all been suggested.

The source image is routed via a one or two deep convolution layers with a deep CNN layer in DeepLabv1 [60]. (Figure 5). A rough map of feature is produced as a result. After that, the feature map is upsampled. using the bilinear interpolation procedure to the size of the source image. To create the interpolated data is applied to a conditional random field that is completely connected to create the final segmented image.

There are numerous atrous convolutions are performed to the input map of feature in the DeepLabv2 model at various dilation rates. The outputs are combined in a fuse. The objects are divided into scale-dependent segments using atrous spatial pyramid pooling (ASPP). The atrous convolution with various rates of dilation was employed in the ResNet model. Atrous convolution allows for the acquisition of data from a vast effective field with a minimal Parameter count and computational complexity.

Enhancing the atrous spatial pyramid pooling (ASPP) module with image-level characteristics makes DeepLabv3 [61] an improvement over DeepLabv2. Additionally, batch normalisation is used to make it simple to train the network. The module of ASPP from DeepLabv3 is combined with an encoder and decoder framework in the DeepLabv3+ paradigm. For feature extraction, the model employs the Xception model. In order to compute more quickly, the model additionally used atrous and depth-wise separable convolution. The structural details and semantic knowledge that the low-level and high-level features, respectively, stand for are

Encoding and decoding modules make up DeepLabv3+ [62]. Using extensive Convolution and the backbone networks like PNASNet, ResNet, Xception, and MobileNetv2, the encoding path obtains the necessary data from the input image. With the help of the data from the encoder path, The output is rebuilt with the required dimensions by the decoding path.

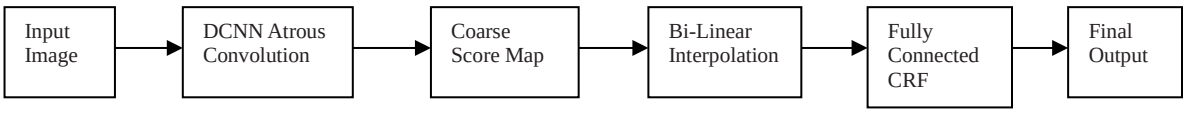

Figure 5: DeepLab architecture.

2.5. Comparison of Different Deep Learning-Based Segmentation Methods.

The different deep neural networks discussed in the above sections are employed for different applications. Each model has its own advantages and limitations. Table 1 gives a brief comparison between different deep learning-based image segmentation algorithms. Table 2 gives a summary on deep learning-based medical image segmentation methods

Table 1: Comparison between different image segmentation algorithms.

Deep learning algorithm	Algorithm description	Advantages	Limitations
CNN	It consists of three main neural layers, which are convolutional layers, pooling layers, and fully connected layers	(a) It is simple (b) It involves feeding segments of an image as input to the network, which labels the pixels	(a) It cannot manage different input sizes (b) Fixed size of output layer causes difficulty in segmentation task
FCN	All fully connected layers of CNN are replaced with the fully convolutional layers	The model outputs a spatial segmentation map instead of classification scores	It is hard to train a FCN model to get good performance
U-Net	It combines the location information obtained from the downsampling path and the contextual information obtained from upsampling path to predict segmentation map	It can perform efficient segmentation of images using limited number of labelled training images	(a)Input image size is limited to 572 × 572. (b) (e skip connections of the model impose a restrictive fusion scheme causing accumulation of the same scale feature maps of the network
VNet	It uses selective search algorithm to extract 2000 regions from the image called region proposals	(a) It predicts the presence of an object within the region proposals (b) It also predicts four offset values to increase the precision of the bounding box	(a) Huge amount of time is needed to train network to classify 2000 region proposals per image (b) It cannot be implemented in real time (c) Selective search algorithm is a fixed algorithm
Fast R-CNN	It uses selective search algorithm which takes the whole image and region proposals as input in its CNN architecture in one forward propagation	It improves mean average precision (mAP) as compared to R-CNN	There is high computation time due to selective search region proposal generation algorithm
Faster R-CNN	It uses region proposal network	It generates the bounding boxes of different shapes and sizes	There is lower computation time
Mask R-CNN	It gives three outputs for each object in the image: its class, bounding box coordinates, and object mask	a) It is simple and flexible approach b) It is current state-of-the-art technique for image segmentation task	There is high training time
DeepLabv1	a) It uses atrous convolution to extract the features from an image b) It also uses conditional random field (CRF) to capture fine details	a) (ere is high speed due to atrous convolution b) Localization of object boundaries improved by combining DCNNs and probabilistic graphical models	Use of CRFs makes algorithm slow
DeepLabv2	It uses *atrous spatial pyramid pooling (ASPP)* and applies multiple atrous convolutions with	Atrous spatial pyramid pooling (ASPP) robustly segments objects at multiple scales	There are challenges in capturing fine object boundaries

	different sampling rates to the input feature map and fuses them together		
DeepLabv3	It uses atrous separable convolution to capture sharper object boundaries	It can segment sharper targets	It still needs more refinement for object boundaries
DeepLabv3+	It extends DeepLabv3 by adding a decoder module to refine the segmentation results along the object boundaries	There is better segmentation performance as compared to deepLabv3	It is a large model with number of parameters to train. So, while training on higher resolution images and batch sizes, it needs large GPU memory.

Table 2: Summary on deep learning-based medical image segmentation methods.

Organ	Segmented area	Model utilized	Dataset	Modality	Remarks	Accuracy
Cardiac	Cardiac, left, and right ventricular cavities and myocardium [63]	2D/3d CNN	ADC2017	Cardiac MR images	--	95.0 (LV), 89.3 (R V), and 89.9 (Myo)
	Heart [64]	RFCN	MICCAI2 2009 challenge dataset	Cardiac MR images	RFCN reduces computational time, simplifies segmentation, and enables real time applications	93.5(Dice)
	Heart [65]	U-Net	--	DT-CMR images	U-Net automated the DT-CMR postprocessin g, supporting real time results	93(Dice)
Brain	Brain [66]	Binary UNet	LGG	MRI	--	99.7
	Brain [67]	Transformer Backbone	LGG	MRI	--	99.6
	Brain tumor [68]	DeepLab		CT images	DeepLab with conditional random fields produces high accuracy	85.74
Lungs	Pneumonia [69]	Iyke-Net	CXR	Chest X-ray images	--	87
	Pneumothora x segmentation [70]	FC-DenseNet with SCSE module	PACS	Chest X-ray image	Spatial weighted cross-entropy loss function improves	93

					precision at boundaries	
	Pneumothorax segmentation [71]	Mask R-CNN	SIIM-ACR	Chest X-ray images	Bounding box regression helps in improving classification	82(Dice)
Liver	Liver and tumor segmentation [72]	Cascaded FCN	DIRCAD dataset	CT and MRI images	Separate set of filters applied at each stage improves segmentation	91,94.3
	Liver segmentation [73]	HED-mask R- CNN	CHAOS challenge	CT and MR images	High segmentation accuracy obtaine	94, 89 and 91
Digestive system	Pancreas [74]	Recurrent NN (LSTM)	NIH-CT-82, uflmri-79	Abdominal CT and MRI images	RNN performs better than HNN and UNET	91.0, 90.5
Breast	Breast masses [75]	DBN + CRF/ SSVM	DDSM-BCRP, INbreast databases	Mammograms	CRF model is faster than SSVM	90
Eyes	Retinal blood vessels [76]	U-Net with modifications	DRIVE/STARE	Retinal images	Modification allows precise and faster segmentation of blood vessels	95.9, 96.1
	Retinal blood vessels [77]	U-Net, LadderNet	DRIVE/STARE/ CHASE	Retinal images	--	97.06, 97.77, and 97.73

3. Techniques of Network Training

3.1. Deeply Supervised

The fundamental principle of deep supervision is to directly supervise the hidden levels and disseminate that supervision down to lower layers as opposed to only doing so at the output layer. By including the complementary objective role of hidden layers, this idea has been used for non-medical purposes in [78]. A 22 layer network's two hidden levels were also overseen in GoogleNet [79].

Deeply supervised techniques were used by Dou et al. in [80] in order to divide the 3D liver CT volumes. Such was accomplished by utilising deconvolution layers to upsample the classification output is densely packed using the lower and intermediate level features and the softmax layer. In addition to higher convergence, their results also display lower training and validation error.

Three classifiers were injected in a similar manner [81] to categorise the mid-level output characteristics from a U-Net-like structure's contracting section. At training stage, the classification outcomes served as a controller. The multi-level environmental data of the network aided in enhancing localization and discriminating skills. Additionally, the Additional classifier improved the gradient's reverse expansion flow during training.

3.2. Weakly Supervised

Existing supervised methods for segmenting medical images automatically call for pixel level (or voxel level in the case of 3D) annotation, which isn't always feasible. Making such annotations will also be time-consuming and

costly [82]. This issue was solved in general image processing by employing services for outsourced labelling similar to Amazon MTurk, which is clearly unsuitable for medical image production. An innovative solution to this problem is to use image-labeled data, such as data with a binary identifier that denotes whether a pattern is present or not.

In order to lessen the system's reliance on fully labelled images, this concept was implemented in [83] by using "point labels," which are simply single pixels indicating the existence of a nodule. Using the statistical data about the nodules, They used the position of that pixel and its surroundings as the positive sample for training by retrieving the volume around it. For instance, the nodules are typically displayed in slices that are 3–7 slices apart and range in width from 3 to 28 pixels. With samples that were only weakly labelled, the method had a respectable sensitivity of 80%.

Segmenting pulmonary nodules entirely automatically from data with poor labelling was achieved by Feng et al. [84] using a CNN. Their approach is based on a finding from [85] that showed CNN's aptitude at locating discriminative zones. In order to do this, they utilised a CNN categorization to identify the pieces of that included nodules, and they also utilised discriminatory area using features to separate the slice's discriminatory areas, creating what is known as a diagram of nodule activation (NAM). A multi-GAP CNN was also created to take advantage of NAMs from shallower layers with better spatial precision, following the idea of [86].The outcome, a 0.55 Dice score, was comparable to completely supervised methods but less accurate. Deeply supervised approaches' dominance was expected because they use pixel-level annotation, which provides essential information to deal with a variety of intensity patterns, especially at the edges. Contrary to [83], which relied more on strict hypotheses about the size and shape of the nodules that were derived from statistical data, the proposed method aids in the more automatic extraction of the nodule-containing areas.

3.3. Transfer Learning

Transfer learning is the ability of a system to recognise and use information obtained in one domain to another [87,88].

Transfer learning can be accomplished in two ways: by fine-tuning a network that has already been trained on generic images [89] or by A network that has already been trained on medical images is being fine-tuned for a new target organ or job. Transfer learning occurs when the duties of the source and target networks are more similar Transfer learning occurs when the duties of the source and target networks are more similar. Transfer learning occurs when the duties of the source and target networks are more similar. has been shown to perform better, but even transferring the weights of responsibilities that are widely apart has outperformed random initialization[90]. The weights in [91] are adjusted for prenatal ultrasound image segmentation using weights from a generic network (VGG16).Similar to [92], the original weights were added to the polyp detection process after being retrieved from a remote application. The layers had to all be adjusted as a consequence. They discovered that increasing the sensitivity by 25% required fine-tuning all levels as opposed to just the top layer. Nevertheless, some experiments that were created from start also produced better results than those that involved perfecting an already learned network [93].

Three main layers can be used to implement transfer learning:

(1) Complete network adaptation entails updating all of the weights throughout training while initialising them using a pre-trained network (as opposed to a random initialization) [94, 95].
(2) Partially adapted networks, which entails initialising the network parameter coming from a previously trained network while the weights freezed for the first layers and updating the last layers while training [96-98].
(3) Zero adaptation entails starting the network's overall weights from a pre-trained model and making no other changes. Due to the enormous variance in how an organ (or target) looks, it is generally not advised to use a medical network adopted a zero adaptation approach.

When providers have received training in generic images, it is especially not advised. Furthermore, because the objects in biomedical images may differ greatly in size and look from one another, segmentation results may not be affected by transfer learning from models with vastly different organ appearances.

3.3.1. Network Structure

However, the network structure also affects the technique that is used. Full adaptation improves performance for shallower networks, whereas partial adaptation will speed up convergence and lighten the computational strain for deeper networks [92].

3.3.2. Organ and Modality

The intended organ and images of it modalities are another essential component of transfer learning. For instance, in [99] they used partial transfer for the T2 modality and complete weight transfer for the T1 MRI. Since the modality

has many noises and the organ has a wide range in appearance, results of [100] demonstrate that the completely adaptation method's average Dice score is higher (ADS) [101] compared to no adaptation and limited adaptation in kidney segmentation by ultrasonography.

3.3.3. Dataset Size

A parameter for role-playing used for determining the degree of transfer learning is the size of the intended dataset. Overfitting could occur if complete adaptation is used with a limited target dataset and many parameters (deeper networks). Therefore, limited adaptation is preferable. On the other hand, overfitting won't be a problem and full adaptation can function properly if the target dataset is comparatively small in size larger. The impact of dataset size on a full adaptation approach was examined by Tajbakhsh et al. in [92]. By expanding the training dataset from a quarter to its entire size, the results show a 10% improvement in sensitivity (from 62 to 72%).

4. Metrics for Segmentation Evaluation and Data Sets

4.1. Evaluation Metrics

You need a reliable objective indication to judge an algorithm's quality. Hand-drawn annotations by physicians are often the gold standard for medical segmentation algorithms (ground truth(GT)). The predictions are another another outcome of the algorithm segmentation (Rseg, SEG for short). Both pixel-based and overlap-based techniques are used to evaluate segmentation of medical images.

Dice index: The formula for determining similarity is called the dice coefficient. It is typically applied to determine how identical or overlapped two samples are. It's also the one that gets used the most. Its possible values are 0 to 1. The segmentation impact improves as the value approaches 1. The metrics is defined as follows given two groups A and B:

$$Dice(A,B) = 2\frac{|A \cap B|}{|A + |B|} \quad (1)$$

Jaccard index: The Jaccard index and die coefficient are comparable. The metrics are described as: given two sets A and

$$Jaccard(A,B) = \frac{|A \cap B|}{|A \cup B|} \quad (2)$$

Segmentation Accuracy (SA): The percentage of the actual region in the GT image is represented by the area of accurate segmentation. Among them, Rs stands for the segmented image's reference region that was painstakingly drawn by the expert. Ts indicates the actual area of the image as determined by the segmentation algorithm. The quantity of incorrectly segmented pixels is denoted by |Rs - Ts|.

$$SA = \left(1 - \frac{|R_s - T_s|}{R_s}\right) \times 100\% \quad (3)$$

Oversegmentation Rate: The following equation is used to determine how many pixels are split into the GT image's reference area:

$$OR = \frac{O_s}{R_s + O_s} \quad (4)$$

In contrast to the theoretical segmented image Rs, in Os, the pixels are present the divided image the divided image that was actually created. R stands for the segmented image's reference area that was by the professional, manually drawn.

Undersegmentation Rate: the proportion of the segmentation's output to the GT image's missing pixels. figured out as follows:

$$UR = \frac{U_s}{R_s + O_s} \quad (5)$$

Although the segmented image Rs contains the pixels in the split image Us, the segmented image Us does not. Both Rs and Os fall under the same classification.

Hausdorff distance: The distance between the two boundaries of the ground truth and the segmentation result input to the network is described here as a measure of the degree of similarity between two groups of points. sensitive to the

dividing line

$$H = \left(\max_{i \in seg}\left(\min_{j \in gt}(d(i,j))\right), \max_{j \in gt}\left(\min_{i \in seg}(d(i,j))\right)\right) \quad (6)$$

where, i and j are points belonging to different sets. d represents the distance between i and j.

4.2. *Medical Image Segmentation Data Sets*

It is essential to gather enough data for the data set before using any deep learning model segmentation. The experts' high-quality image data and data set with matching label standards, which allows for fair system comparison, are what determine the segmentation algorithm's quality[81]. The public data sets that are frequently used in the segmentation of medical images will be introduced in this section.

Medical Segmentation Decathlon (MSD):A sizable, a collection of different anatomical sections' open source, manually annotated medical images was produced by Simpson et al. [102]. Through thorough benchmarks, Using this data set impartially assess segmentation generally techniques while opening up availability of medical images to the general public. The data set includes 2633 three-dimensional medical images from various sources and real-world clinical uses of different anatomical parts. There are ten categories in total:

1. Task01, BrainTumor: In total, there are 750 labels, which are divided into two categories: edoema and glioma (necrotic/active cancer). It is an MRI scan that was obtained in the course of routine therapeutic practise.
2. Task02, Heart: There are thirty in total, and the placard specifically mentions the left atrium. These figures are taken from the Left Atrial Segmentation Challenge. (LASC). Images were created using a 1.5T Achieva scanning and had voxel resolution of 1.25 1.25 2.7 mm3.
3. Task03, Liver: There are 201 sheets in total, with markings divided into sections for the liver and tumours. The diagnostic technique is known as CT. For the pictures, in-plane resolution varied from 0.5 to 1.0 mm and slice thickness from 0.45 to 6.0 mm.
4. Task04, Hippocampus: With the labels hippocampus, cranium, and body, there are 394 in total. The technique is called MRI imaging. The data set included MRI scans from 105 people with nonaffective psychotic disorders and 90 adults in good health.
5. Task05, Prostate: There are 48 in total, and the markings read: Central gland of the prostate, peripheral zone. The technique is MRI imaging. 48 multiparametric MRI scans from Radboud University (The Netherlands) were part of the prostate data set and had earlier been distributed and segmented.
6. Task06, Lung: Lung tumour is the designation for all 96 of them. The technique is CT radiography. Patients from Stanford University with non-small-cell lung cancer made up the lung data collection. On a typical CT cross section, an expert thoracic radiologist marked the tumour position using OsiriX.
7. Task07 Pancreas: has 420 labels total, with the pancreas and pancreatic mass labels being separated. CT imaging is the method. Patients whose pancreatic masses were removed made up the pancreas data set.
8. Task08 HepaticVessel: The labels for the 443 total liver vessels are "hepatic vessels." CT imaging is the method. Patients with various primary and metastatic liver tumours made up the second liver data set.
9. Task09 Spleen: The spleen is the label for 61 different items. CT imaging is the method. Patients at Memorial Sloan Kettering Cancer Center receiving chemotherapy for liver metastases made up the spleen data set.
10. Task10 Colon: The label for all 190 cases is "colon cancer." CT imaging is the method.

Segmentation in Chest Radiographs (SCR):The JSRT database serves as the source for all chest radiographs. The SCR database was developed to make it easier to compare the segmentation of the heart, clavicle, and lung field in conventional posterior chest radiographs [103]. The database's entire data set has been carefully split to offer benchmarks. With a grayscale of 12 bits and a spatial resolution of 0.175 mm/pixel, the image is scanned from film to a resolution of 2048 2048 pixels. There are no lung nodules in any of the other 93 images, but all 154 images contain at least one lung nodule.

Brain Tumor Segmentation (BRATS): This data set is a combination of the MICCAI conference and a brain tumor data set for the segmentation challenge [104]. Since 2012, it has been held annually to assess the most effective techniques for segmenting brain tumours and to contrast various techniques. The data set is published as a result. Necrotic region, healthy brain tissue, edoema area, tumor areas for enhancement and non-enhancement are the five different sorts of labelling. Every year, There are new training sets introduced.

Digital Database for Screening Mammography (DDSM): The field of mammography image analysis study makes

extensive use of the resource DDSM [105]. The database has over 2500 studies in it. Two breast images from each study are included, along with some pertinent patient and image data.

Ischemic Stroke Lesion Segmentation (ISLES):This offers MRI images with several precise stroke samples and associated clinical data. To assess predicting clinical outcomes and stroke histology in precise MRI scan images, this challenge is set up.

Liver Tumor Segmentation (LiTS):For the segmentation of the liver and liver tumours, various clinical locations from around the world have contributed these data and segmentations. 130 CT scans are included in the training data set, while 70 scans are included in the test data set [106].

Prostate MR Image Segmentation (PROMISE12): For prostate segmentation, the data collection is used. Patients with prostate cancer and benign illnesses such benign prostatic hyperplasia are included in these statistics. These examples feature a transverse T2-weighted MR prostate image.

Lung Image Database Consortium Image Collection (LIDC-IDRI):This data set consists of lesion labels that correlate to diagnosis results for chest medical image files (such as CT and X-ray). Studying early cancer detection in high-risk populations is the goal. There are 1018 study examples in all. Four seasoned The diagnostic and annotation process was done in two stages by thoracic radiologists on the images in each scenario [107].

Open Access Series of Imaging Studies (OASIS):The goal of this initiative is to make it possible for the provision of brain MRI data sets by the scholarly community without charging anything. There is now a third generation available. The 30-year-old The WUSTL Knight ADRC's OASIS-3 is a retrospective compilation of data from more than 1000 participants in a number of ongoing projects. OASIS-3 is a longitudinal data collection for Alzheimer's disease and normal ageing that includes neuroimaging, clinical, cognitive, and biomarker information. Participants ranged in age from 42 to 95 [108] and included 609 adults with normal cognitive function and 489 people in varying stages of cognitive decline.

Digital Retinal Images for Vessel Extraction (DRIVE): The data set is used for comparing how blood vessels are segmented in retinal images. The 40 images that made up the DRIVE database were chosen at random from a Dutch experiment to screen for diabetic retinopathy. 33 of them showed no symptoms of diabetic retinopathy, while seven had modest early symptoms in those circumstances. 8 bits per colour plane are used to capture each image at a resolution of 768 by 584 pixels. Each image has a circular field of view with a diameter of roughly 540 pixels [109].

Mammographic Image Analysis Society (MIAS):A British research organisation established the MIAS breast cancer X-ray imaging database in 1995. Eight bits of grayscale make up each pixel. 322 images in total, including 208 images of healthy breasts, 63 benign breast cancer images and 51 malignant breast cancer images, are available in the MIAS database for the left and right breasts of 161 individuals. Experts have calibrated the lesion area's border as well [110].

Sunnybrook Cardiac Data (SCD): The dataset, also referred to as the 2009 cardiac MR left ventricle segmentation challenge data, comprises of 45 cine-MRI images of patients with a range of pathologies, including healthy subjects, patients with hypertrophic heart disease, patients with heart failure with and without an infarction, and patients with heart failure without an infarction [111].

In addition to the numerous data sets that are commonly used for medical image segmentation, as discussed above, there are several competition data sets that support the excellence of the algorithm provided by the renowned medical image challenge competition.

***Grand Challenges in Biomedical Image Analysis*:** It was made to assist people in resolving problems related to international development and health. It addresses every difficulty related to study of medical images, including processing of medical images. The development of many good algorithms has been sparked by this, which also presents the greatest challenge in the area of medical image processing.

Liver Tumor Segmentation Challenge: This contest aims to motivate researchers to investigate liver lesion segmentation techniques. Different clinical sites throughout the world give the data and slices for the challenger tournament. There are 130 CT scans in the training data set and 70 in the test data set.

2019 Kidney and Kidney Tumor Segmentation Challenge (KiTS19):In contrast-enhanced CT images, the semantic

segmentation of kidneys and kidney cancers is the KiTS19 challenge. 300 individuals having preoperative arterial-phase abdominal CTs that had been expertly annotated made up the data set. Of them, 210 (70%) were made available as a training set, and the remaining 90 (30%) were saved for the test set. The data sets for segmenting medical image are shown in Table 3.

Table 3. Medical image segmentation data sets.

Data Set	Modalities	Objects	URL
ISLES	MRI	Brain	http://www.isles-challenge.org/
LiTS	CT	Liver	https://competitions.codalab.org/competitions/17094
PROMISE12	MRI	Prostate	https://promise12.grand-challenge.org/
LIDC-IDRI	CT	Lung	https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI
MSD	MRI, CT	Various	http://medicaldecathlon.com/
BRATS	MRI	Brain	https://www.med.upenn.edu/sbia/brats2018/data.html
DDSM	Mammography	Breast	http://www.eng.usf.edu/cvprg/Mammography/Database.html
OASIS	MRI, PET	Brain	https://www.oasis-brains.org/
DRIVE	Funduscopy	Eye	https://drive.grand-challenge.org/
STARE	Funduscopy	Eye	http://homes.esat.kuleuven.be/~mblaschk/projects/retina/
CHASEDB1	Funduscopy	Eye	https://blogs.kingston.ac.uk/retinal/chasedb1/
MIAS	X-ray	Breast	https://www.repository.cam.ac.uk/handle/1810/250394?show=full
SCD	MRI	Cardiac	http://www.cardiacatlas.org/studies/
SKI10	MRI	Knee	http://www.ski10.org/
HVSMR2018	CMR	Heart	http://segchd.csail.mit.edu/

5. Major Challenges and State-of-the-Art Solutions

Deep learning has benefited the field of medical image segmentation, although it is still difficult to use deep neural networks for the following reasons.

5.1. Data set challenges

The following are some of the several dataset-related challenges:

Limited Annotated Dataset. Large amounts of data are needed for deep learning network models. The training-related data is richly annotated. In many DL-based medical procedures, the dataset is crucial [112]. It can be challenging to gather a lot of annotated medical images for use in medical image processing [113]. Additionally, annotating new medical images is time-consuming, expensive, and complicated. Several sizable datasets are openly accessible. More difficult datasets are still required in order to improve DL model training and support managing dense objects. Larger and more challenging datasets are required because the majority of the existing 3D datasets [114] are not very large and very few of them are generated.

The size of the current medical image libraries can be increased by using image augmentation techniques like cropping and shearing. Other techniques include rotating an image at various angles and flipping it vertically or diagonally. The performance of the system can be enhanced by these techniques. (b) Transfer learning from powerful models can be applied to the problem of insufficient data. (c) Combining data from various sources is the last stage [115].

Class Imbalance in Datasets. There is inherent class inequality in many publicly accessible medical image collections. Where the disease is highly prevalent in patient data uncommon and affects just 10% of patients examined, a severely imbalanced data set makes model accuracy misleading and makes training DL models very challenging. Since most patients are free of the disease and would achieve a local minimum the overall planned model accuracy would be high [116, 117].

Class imbalance can be resolved in one of two ways: (a) oversampling the data, the amount of which is based on the dataset's degree of class inequality. (b) Secondly, the issue of dataset imbalance can be solved by altering the

assessment or performance metric. (c) New data samples can be produced using approaches for data augmentation. (d) By combining minority classes, the issue of dataset class disparity can also be resolved.

Sparse Annotations. Fully annotating 3D images requires work and is not always possible. So, 3D image information segments are partially labelled. It is extremely difficult to train models using these sporadic 3D annotations [113]. The dataset can be subjected to weighted loss function in the case of a sparsely annotated dataset. To learn just from the labelled pixels, The dataset's unlabeled data has all of its weights assigned to zero.

Intensity Inhomogeneities. Inhomogeneities in colour and intensity are prevalent in pathology images [118]. Shading is a result of intensity inhomogeneities over the image. In the segmentation of MR images, it is more precise. Additionally, because nonuniform support films are present, the brightness of the TEM images varies. These variations make segmentation laborious.

In the literature, different nonparametric strategies are proposed and various algorithms are used, to adjust intensity inhomogeneities [118]. To eliminate inhomogeneities, segmentation can be performed prior to prefiltering. Additionally, improvements in scanning technology address intensity inhomogeneities.

Complexities in Image Texture. Various artefacts from image editing may be visible in medical images. Noise is introduced into the image by the various sensors and electronic parts employed to capture it [10,119]. There may be weak image boundaries and grey levels that are quite close to one another in the obtained image. Tissue overlap may be visible in dermoscopic images, as well as abnormalities like skin lines and hair. It is challenging to identify the region of interest in medical imaging because of all these complexity.

Before segmentation, a variety of image enhancing techniques are performed to eliminate various noises and artefacts in the image. The image enhancement approach reduces image noise while maintaining the sharpness of the image's edges.

5.2. DL Models Challenges

The following are the significant difficulties in training DNN to reliable the medical images segmentation:

Overfitting the Model, The model being overfit occurs when it learns Compared to the situation of unprocessed data, the training dataset contains the details and regularities with a high degree of accuracy. The majority of the time, it occurs when training a model with sparse training data[120] .

With the use of augmentation techniques, the size of the dataset can be increased, which can be used to [116] address overfitting. (b) Dropout strategies [121] assist manage overfitting by discarding some of the output from a random group of neurons of network throughout every cycle.

Memory Efficient Models, Models for segmenting medical images need a lot of memory [122]. These models must be simplified in order to be compatible with particular devices, such as mobile phones.

The amount of memory needed for a DL model The amount of memory needed for a DL model The amount of memory needed for a DL model can be decreased by using simpler models and model compression techniques.

Training Time. Deep neural network architecture training takes time. Fast convergence of the deep NN training time is necessary for image segmentation.

Applying batch normalisation is one way to solve this issue [122]. By deducting the pixel values of the image's mean value, it refers to finding the pixels close to 0. It works well at enabling quick convergence. Convergence can also be accelerated by adding pooling layers to decrease parameter dimension.

Vanishing Gradient. Vanishing gradient is a challenge for deep neural networks [123]. It happens because It is impossible to revert to previous layers if the final gradient is lost.. In 3D models, the vanishing gradient issue is more obvious.

The gradient disappearing issue has a number of solutions. (a) The gradient value is strengthened by combining the auxiliary losses and the initial loss of the hidden layers and upscaling the intermediary hidden layer output using deconvolution and softmax [118]. (b) We can also avoid vanishing gradient by carefully initialising the weights [124] for the network.

Computational Complexity. High computing efficiency is required for feature analysis performed by deep learning algorithms. High performance computer systems and GPU are required for these methods [125]. Some of the more

advanced algorithms could need supercomputers for model training, which might not be available. In order to overcome these problems, the researcher must take into account a specified set of factors in order to achieve a certain level of accuracy.

Conclusion

The most common network architectures used for medical image segmentation are firstly listed in this research.Then, we provided a summary of the primary training methods for the segmentation of medical image, together with their benefits and shortcomings.A summary of the numerous medical image datasets used for illness segmentation as well as the various performance measures used to assess the effectiveness of the image segmentation algorithm are also provided.Finally, we concentrated on the primary difficulties with deep learning-based approaches to medical image segmentation. Deep learning is becoming increasingly relevant in image segmentation as a result of technological advancements.The current work will aid scientists in developing neural network architectures for disease diagnosis in the medical field. Additionally, the researchers will get knowledge about the cutting-edge solutions and potential difficulties in the field of segmenting medical images using deep learning.

The paper has several contributions which are as follows: Firstly, the present study provides an overview of the current state of the deep neural network structures utilized for medical image segmentation. Secondly, the paper describes the publicly available techniques of Network Training. Thirdly, it presents the various performance metrics employed for evaluating the deep learning segmentation models and the medical image segmentation datasets. Finally, the paper also gives an insight into the major challenges faced in the field of image segmentation and their state-of-the-art solutions

References

[1] Priyanka Malhotra , Sheifali Gupta , Deepika Koundal , Atef Zaguia , and Wegayehu Enbeyle (2022). *Review Article* :Deep Neural Networks for Medical Image Segmentation, Journal of Healthcare Engineering , Volume 2022, Article ID 9580991, 15 pages, https://doi.org/10.1155/2022/9580991

[2] Mohammad HesamHesamian· WenjingJia· Xiangjian He· Paul Kennedy (2019). Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges, in Journal of Digital Imaging, 32, 582-596.

[3] Kunal Aggarwal, Marina Manso Jimeno, Keerthi Sravan Ravi, Gilberto Gonzalez, Sairam Geethanath (2023). Developing and deploying deep learning models in brain MRI: a review, arXiv:2301.01241v1 [eess.IV], 3 Jan 2023.
https://doi.org/10.48550/arXiv.2301.01241

[4] Van der Velden BHM, Kuijf HJ, Gilhuijs KGA, Viergever MA. (2022). Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. Med Image Anal., 79,102470.

[5] Lu Liu, Jelmer M Wolterink, Christoph Bruneand Raymond N J Veldhuis (2021). Anatomy-aided deep learning for medical image segmentation: areview, Physics in Medicine & Biology,66.

[6] Minaee S, Boykov Y, Porikli F, Plaza A, Kehtarnavaz N, Terzopoulos D (2022. Image segmentation using deep learning: a survey. IEEE Trans Pattern Anal Mach Intell, 44(7), pp. 3523–3542. https://doi.org/10.1109/TPAMI.2021.3059968

[7] Sah M, Direkoglu C ,(2022). A survey of deep learning methods for multiple sclerosis identification using brain mri images. Neural Comput Appl, 34(10),7349–7373. https://doi.org/10.1007/s00521-022-07099-3

[8] Taghanaki S A, Abhishek K, Cohen J P, Cohen-Adad J and HamarnehG (2021). Deep semantic segmentation of natural and medical images: a review, Artif. Intell. Rev., 54,1–42.

[9] Haque I R I and Neubert J (2020). Deep learning approaches to biomedical image segmentation, Inform. Med. Unlocked,18, 100297.

[10] Hesamian M H, Jia W, He X and Kennedy P (2019). Deep learning techniques for medical image segmentation: achievements and challenges, J. Digit. Imaging , 32, 582–596.

[11] Liu, X.; Song, L.; Liu, S.; Zhang, Y. (2021)"A Review of Deep-Learning-Based Medical Image Segmentation Methods". *Sustainability* 2021, *13*, 1224. https://doi.org/ 10.3390/su13031224

[12] Zhuang X et al ,(2019). " Evaluation of algorithms for multi-modality whole heart segmentation: an open-access grand challenge", Med. Image Anal., 58, 101537.

[13] Chen C, Qin C, Qiu H, Tarroni G, Duan J, Bai W and Rueckert D, (2020)" Deep learning for cardiac image segmentation: a review", Front. Cardiovascular Med., 7(25).

[14] Yi X, Walia E and Babyn P,(2019). " Generative adversarial network in medical imaging: a review", Med. Image Anal., 58, 101552.
[15] Zhang P, Zhong Y, Deng Y, Tang X and Li X, (2019). " A survey on deep learning of small sample in biomedical image analysis ", arXiv:1908.00473.
[16] Karimi D, Dou H, Warfield S K and Gholipour A, (2020)" Deep learning with noisy labels: exploring techniques and remedies in medical image analysis", Med. Image Anal., vol. 65, 101759.
[17] Tajbakhsh N, Jeyaseelan L, Li Q, Chiang J N, Wu Z and Ding X, (2020). " Embracing imperfect datasets: a review of deep learning solutions for medical image segmentation ", Med. Image Anal., vol. 63, 101693.
[18] Cheplygina V, de Bruijne M and Pluim J P,(2019). " Not-so-supervised: a survey of semi-supervised, multi-instance, and transfer learning in medical image analysis", Med. Image Anal., vol. 54, pp. 280–296.
[19] Jurdia R E, Petitjean C, Honeine P, Cheplygina V and Abdallah F, (2020)." High-level prior-based loss functions for medical image segmentation: a survey", arXiv:2011.08018v2 22 November 2020.
[20] Bohlender S, Oksuz I and Mukhopadhyay A,(2021). "A survey on shape-constraint deep learning for medical image segmentation", arXiv:2101. 07721 19 January 2021.
[21] Yedavalli V, Tong E, Martin D, Yeom K and Forkert N,(2020). " Artificial intelligence in stroke imaging: Current and future perspectives"Clin. Imaging, vol. 69, pp. 246–254.
[22] Zhang H, Wang Y, Geng X and Fu P,(2020). " Review of deep learning methods for isointense infant brain mr image segmentation", J. Image Graph., vol. 25, pp. 2068–78.
[23] Ma J, Deng Y and Ma Z,(2020) " Review of deep learning segmentation methods for ct images of liver tumors", J. Image Graph., vol. 25, pp. 2024–46.
[24] Vrtovec T, Močnik D, Strojan P, Pernuš F and Ibragimov B, (2020). " Auto-segmentation of organs at risk for head and neck radiotherapy planning: from atlas-based to deep learning methods", Med. Phys., vol. 47, pp. 929–50.
[25] Ebrahimkhani S, Jaward M H, Cicuttini F M, Dharmaratne A, Wang Y and de Herrera A G S, (2020). " A review on segmentation of knee articular cartilage: from conventional methods towards deep learning",Artif. Intell. Med., vol. 106, 101851.
[26] Sun L, Zhang L and Zhang D, (2019). " Multi-atlas based methods in brain mr image segmentation", Chin. Med. Sci. J., 34,110–119.
[27] Hameurlaine M and Moussaoui A, (2019). " Survey of brain tumor segmentation techniques on magnetic resonance imaging", Nano Biomed. Eng, 11, 178–191.
[28] Saman S and Narayanan S J, (2019). " Survey on brain tumor segmentation and feature extraction of mr images", Int. J. Multimedia Inf. Retr., 8, 79–99.
[29] Jiang Z, Lyu X, Zhang J, Zhang Q andWei X, (2020). " Review of deep learning methods for mri brain tumor image segmentation", J. Image Graph., 25, 215–228.
[30] Li Q, Bai K, Zhao L and Guan X, (2020). "Progresss and challenges of mri brain tumor image segmentation", J. Image Graph., 25, 419–31.
[31] Feo R and Giove F, (2019). " Towards an efficient segmentation of small rodents brain: a short critical review", J. Neurosci. Methods, 323, 82–89.
[32] Lin X and Li X, (2019). "Image based brain segmentation: From multi-atlas fusion to deep learning", Curr. Med. Imaging, 15, 443–452.
[33] Chen J, You H and Li K, (2020). " A review of thyroid gland segmentation and thyroid nodule segmentation methods for medical ultrasoundimages ",Comput. Methods Programs Biomed., 185, 105329.
[34] Soomro T A, AfifiA J, Zheng L, Soomro S, Gao J, Hellwich O and Paul M, (2019). " Deep learning models for retinal blood vessels segmentation: areview",IEEE Access, 7, 71696–717.
[35] Samuel P M and Veeramalai T, (2020). " Review on retinal blood vessel segmentation-an algorithmic perspective",Int. J. Biomed. Eng. Technol., 34, 75–105.
[36] Goodfellow, I.; Bengio, Y.; Courville, A.; Bengio, Y. (2016). *Deep Learning*; MIT Press: Cambridge, UK.
[37] Almeida, G.; Tavares, J.M.R.S. (2020). Deep learning in radiation oncology treatment planning for prostate cancer: A systematic review. *J. Med. Syst*, , *44*, 1–15. [CrossRef]

[38] Fengming Lin, Yan Xia, Shuang Song, Nishant Ravikumar, Alejandro F. Frangi, (2023). "Lin, Fengming, et al. "High-Throughput 3DRA Segmentation of Brain Vasculature and Aneurysms using Deep Learning." *Computer Methods and Programs in Biomedicine*, 230, 107355.

[39] A. Krizhevsky, I. Sutskever, and G. E. Hinton, (2012). "Imagenet classification with deep convolutional neural networks," in *Proceedings of the Advances in neural information processingsystems*, pp. 1097–1105, Long Beach, CA, USA, December 2012.

[40] F. Chollet, "Xception: deep learning with depthwise separable convolutions, (2017). " in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1251–1258, Honolulu, HI, USA, July 2017.

[41] K. SIMONYAN AND A. ZISSERMAN, (2014). "VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION,", HTTPS:// ARXIV.ORG/ABS/1409.1556.

[42] C. SZEGEDY, V. VANHOUCKE, S. IOFFE, J. SHLENS, AND Z. WOJNA, (2016). "RETHINKING THE INCEPTION ARCHITECTURE FOR COMPUTER VISION," IN *PROCEEDINGS OF THE IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION*, PP. 2818–2826, LAS VEGAS, NV, USA.

[43] F. N. IANDOLA, S. HAN, M. W. MOSKEWICZ, K. ASHRAF, W. J. DALLY, AND K. KEUTZER, (2016). "SQUEEZENET: ALEXNET-LEVEL ACCURACY WITH 50X FEWER PARAMETERS AND< 0.5 MB MODEL SIZE," , HTTPS://ARXIV.ORG/ABS/1602.07360.

[44] G. HUANG, Z. LIU, L. V. D. MAATEN, AND K. Q. WEINBERGER, (2017). "DENSELY CONNECTED CONVOLUTIONAL NETWORKS," IN *PROCEEDINGS OF THE IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION*, PP. 4700–4708, HONOLULU, HI, USA.

[45] T. HUSSAIN, A. ULLAH, U. HAROON, K. MUHAMMAD, AND S. W. BAIK, (2021). "A COMPARATIVE ANALYSIS OF EFFICIENT CNN-BASED BRAIN TUMOR CLASSIFICATION MODELS," *GENERALIZATION WITH DEEP LEARNING: FOR IMPROVEMENT ON SENSING CAPABILITY*, PP. 259–278.

[46] J. LONG, E. SHELHAMER, AND T. DARRELL, (2015). "FULLY CONVOLUTIONAL NETWORKS FOR SEMANTIC SEGMENTATION," IN *PROCEEDINGS OF THE IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION*, PP. 3431–3440, MA, USA.

[47] Y. Guo, Y. Liu, T. Georgiou, and M. S. Lew, (2018). "A review of semantic segmentation using deep neural networks," *International journal of multimedia information retrieval*, 7(2), 87–93.

[48] W. Liu, A. Rabinovich, and A. C. Berg, (2015). "Parsenet: looking wider to see better," , https://arxiv.org/abs/1506.04579.

[49] O. Ronneberger, P. Fischer, and T. Brox, (2015). "U-net: convolutional networks for biomedical image segmentation," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234–241, Springer, Munich, Germany.

[50] H. CUI, X. LIU, AND N. HUANG, (2019). "PULMONARY VESSEL SEGMENTATION BASED ON ORTHOGONAL FUSED U-NET++ OF CHEST CT IMAGES," IN *PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON MEDICAL IMAGE COMPUTING AND COMPUTER-ASSISTED INTERVENTION*, PP. 293–300, SPRINGER, SHENZHEN, CHINA, OCTOBER .

[51] Q. JIN, Z. MENG, C. SUN, H. CUI, AND R. SU, (2020). "RA-UNET: A HYBRID DEEP ATTENTION-AWARE NETWORK TO EXTRACT LIVER AND TUMOR IN CT SCANS," *FRONTIERS IN BIOENGINEERING AND BIOTECHNOLOGY*, VOL. 8, P. 1471.

[52] C. GUO, M. SZEMENYEI, Y. PEI, Y. YI, AND W. ZHOU, (2019). "SD-UNET: A STRUCTURED DROPOUT U-NET FOR RETINAL VESSEL SEGMENTATION," IN *PROCEEDINGS OF THE 2019 IEEE 19TH INTERNATIONAL CONFERENCE ON BIOINFORMATICS AND BIOENGINEERING (BIBE)*, PP. 439–444, IEEE, ATHENS, GREECE.

[53] F. MILLETARI, N. NAVAB, AND S. A. AHMADI, (2016). "V-NET: FULLY CONVOLUTIONAL NEURAL NETWORKS FOR VOLUMETRIC MEDICAL IMAGE SEGMENTATION," IN *PROCEEDINGS OF THE 2016 FOURTH INTERNATIONAL CONFERENCE ON 3D VISION (3DV)*, PP. 565–571, IEEE, STANFORD, CA, USA.

[54] R. GIRSHICK, J. DONAHUE, T. DARRELL, AND J. MALIK, (2014). "RICH FEATURE HIERARCHIES FOR ACCURATE OBJECT DETECTION AND SEMANTIC SEGMENTATION," IN *PROCEEDINGS OF THE IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION*, PP. 580–587, COLUMBUS, OH, USA.

[55] R. GIRSHICK, "FAST R-CNN, (2015). " IN *PROCEEDINGS OF THE IEEE INTERNATIONAL CONFERENCE ON COMPUTER VISION*, PP. 1440–1448, SANTIAGO, CHILE.

[56] P. MALHOTRA AND E. GARG, (2020). "OBJECT DETECTION TECHNIQUES: A COMPARISON," IN *PROCEEDINGS OF THE 2020 7TH INTERNATIONAL CONFERENCE ON SMART STRUCTURES AND SYSTEMS (ICSSS)*, PP. 1–4, IEEE, CHENNAI, INDIA.

[57] S. Ren, K. He, R. Girshick, and J. Sun, (2015). "Faster r-cnn: towards real-time object detection with region proposal networks," in *Proceedings of the Advances in neural information processing systems*, pp. 91–99, Montreal, Canada.

[58] K. He, G. Gkioxari, P. Doll´ar, and R. Girshick,(2017). "Mask r-cnn," in *Proceedings of the IEEE international conference on computer vision*, pp. 2961–2969, Venice, Italy.

[59] L. C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, (2017). "Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(4), 834–848.

[60] L. C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, (2018). "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Proceedings of the European conference on computer vision (ECCV)*, pp. 801–818, Munich, Germany.

[61] L. C. Chen, G. Papandreou, F. Schroff, and H. Adam, (2017). "Rethinking atrous convolution for semantic image segmentation," , https://arxiv.org/abs/1706.05587.

[62] H. Harkat, J. Nascimento, and A. Bernardino, "(2020). Fire segmentation using a DeepLabv3+ architecture," *Image and Signal Processing for Remote Sensing XXVI*, vol. 11533, Article ID 115330M.

[63] C. F. Baumgartner, L. M. Koch, M. Pollefeys, and E. Konukoglu,(2018). "An exploration of 2D and 3D deep learning techniques for cardiac MR image segmentation," in *Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges. STACOM 2017. Lecture Notes in Computer Science*, M. Pop, Ed., vol. 10663, Cham, Springer.

[64] R. P. Poudel, P. Lamata, and G. Montana, (2016). "Recurrent fully convolutional neural networks for multi-slice MRI cardiac segmentation," in *Reconstruction, segmentation, and analysis of medical images*, pp. 83–94, Springer, Cham.

[65] P. F. Ferreira, R. R. Martin, A. D. Scott et al., (2020). "Automating in vivo cardiac diffusion tensor postprocessing with deep learning-based segmentation," *Magnesium. Resonance in Medicine*, 84, (5).

[66] Nageshwar Rao B, D.Lakshmi Sreenivasa Reddy and Harish Gujja,(2022). "Brain MRI Segmentation Binary U-Net BasedArchitecture using Deep learning algorithm", Springer Nature, DOI: https://doi.org/10.21203/rs.3.rs-1916275/v1

[67] Tashvik Dhamija, Anunay Gupta, Shreyansh Gupta, Anjum, Rahul Katarya and Ghanshyam Singh, (2023) "Semantic segmentation in medical images through transfusedconvolution and transformer networks",Springer Science+Business Media, LLC, part of Springer Nature, Applied Intelligence, 53, 1132–1148,https://doi.org/10.1007/s10489-022-03642-w

[68] X. Gao and Y. Qian, (2018). "Segmentation of brain lesions from CT images based on deep learning techniques," in *Medical Imaging 2018: Biomedical Applications in Molecular, Structural, and Functional Imaging*vol. 10578, Article ID 105782L, 2018.

[69] A. Victor Ikechukwu, S. Murali, R. Deepu, R.C. Shivamurthy, (2021). "ResNet-50 vs VGG-19 vs training from scratch: A comparative analysis of the segmentation and classification of Pneumonia from chest X-ray images", Global Transitions Proceedings, 2, 375–381. https://doi.org/10.1016/j.gltp.2021.08.027

[70] Q. Wang, Q. Liu, G. Luo et al., (2020) "Automated segmentation and diagnosis of pneumothorax on chest X-rays with fully convolutional multi-scale ScSE-DenseNet: a retrospective study," *BMC Medical Informatics and Decision Making*, 20 (14), 1–12.

[71] H. Wang, H. Gu, P. Qin, and J. Wang, (2020). "CheXLocNet: automatic localization of pneumothorax in chest radiographs using deep convolutional neural networks," *PLoS One*, 15(11).

[72] P. F. Christ, F. Ettlinger, F. Gr¨un et al., (2017). "Automatic liver and tumor segmentation of CT and MRI volumes using cascaded fully convolutional neural networks," , https://arxiv.org/ abs/1702.05970.

[73] S. Mulay, G. Deepika, S. Jeevakala, K. Ram, and M. Sivaprakasam, (2019). "Liver segmentation from multimodal images using HED-mask R-CNN," in *Proceedings of the International Workshop on Multiscale Multimodal Medical Imaging*, pp. 68–75, Springer, Shenzhen, China.

[74] J. Cai, L. Lu, F. Xing, and L. Yang, (2018). "Pancreas segmentation in CT and MRI images via domain specific network designing and recurrent neural contextual learning," , https://arxiv. org/abs/1803.11303.

[75] N. Dhungel, G. Carneiro, and A. P. Bradley, (2015). "Deep learning and structured prediction for the segmentation of mass in mammograms," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, 605–612, Springer, Munich, Germany.

[76] T. A. Soomro, A. J. Afifi, A. Shah et al., (2019). "Impact of image enhancement technique on CNN model for retinal blood vessels segmentation," *IEEE Access*, 7, Article ID 158197.

[77] Y. Jiang, H. Zhang, N. Tan, and L. Chen, (2019). "Automatic retinal blood vessel segmentation based on fully convolutional neural networks," *Symmetry*, 11(9), 1112.
[78] LEE CY, XIE S, GALLAGHER P, ZHANG Z, TU Z, (2015). DEEPLY-SUPERVISED NETS. IN: ARTIFICIAL INTELLIGENCE AND STATISTICS, 562–570.
[79] SZEGEDY C, LIU W, JIA Y, SERMANET P, REED S, ANGUELOV D, ERHAN D, VANHOUCKE V, RABINOVICH A, (2015). GOING DEEPER WITH CONVOLUTIONS. IN: PROCEEDINGS OF THE IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, PP 1–9.
[80] DOU Q, CHEN H, JIN Y, YU L, QIN J, HENG PA, (2016). 3D DEEPLY SUPERVISED NETWORK FOR AUTOMATIC LIVER SEGMENTATION FROM CT VOLUMES. IN: INTERNATIONAL CONFERENCE ON MEDICAL IMAGE COMPUTING AND COMPUTER-ASSISTED INTERVENTION. SPRINGER, PP 149–157.
[81] CHEN H, QI X, CHENG JZ, HENG PA ET AL, (2016). DEEP CONTEXTUAL NETWORKS FOR NEURONAL STRUCTURE SEGMENTATION. IN: AAAI, PP 1167–1173
[82] HWANG S, KIM HE, (2016). SELF-TRANSFER LEARNING FOR WEAKLY SUPERVISED LESION LOCALIZATION. IN: INTERNATIONAL CONFERENCE ON MEDICAL IMAGE COMPUTING AND COMPUTER-ASSISTED INTERVENTION. SPRINGER, PP 239–246
[83] ANIRUDH R, THIAGARAJAN JJ, BREMER T, KIM H, (2016). LUNG NODULE DETECTION USING 3D CONVOLUTIONAL NEURAL NETWORKS TRAINED ON WEAKLY LABELED DATA. IN: MEDICAL IMAGING 2016: COMPUTER-AIDED DIAGNOSIS, VOL 9785, P 978532. INTERNATIONAL SOCIETY FOR OPTICS AND IMAGENICS
[84] FENG X, YANG J, LAINE AF, ANGELINI ED, (2017). DISCRIMINATIVE LOCALIZATION IN CNNS FOR WEAKLY-SUPERVISED SEGMENTATION OF PULMONARY NODULES. IN: INTERNATIONAL CONFERENCE ON MEDICAL IMAGE COMPUTING AND COMPUTER-ASSISTED INTERVENTION. SPRINGER, PP 568–576
[85] ZHOU B, KHOSLA A, LAPEDRIZA A, OLIVA A, TORRALBA A, (2016). LEARNING DEEP FEATURES FOR DISCRIMINATIVE LOCALIZATION. IN: PROCEEDINGS OF THE IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, PP 2921–2929
[86] LONG J, SHELHAMER E, DARRELL T, (2015). FULLY CONVOLUTIONAL NETWORKS FOR SEMANTIC SEGMENTATION. IN: PROCEEDINGS OF THE IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, PP 3431–3440
[87] SHIE CK, CHUANG CH, CHOU CN, WU MH, CHANG EY, (2015). TRANSFER REPRESENTATION LEARNING FOR MEDICAL IMAGE ANALYSIS. IN: 2015 37TH ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY (EMBC). IEEE, PP 711–714
[88] P. Celard, E. L. Iglesias, J. M. Sorribes-Fdez, R. Romero, A. Seara Vieira, and L. Borrajo, (2023). "A survey on deep learning applied to medical images: from simple artificial neural networks to generative models", Neural Computing and Applications, 35, 2291–2323. https://doi.org/10.1007/s00521-022-07953-4
[89] HOO-CHANG S, ROTH HR, GAO M, LU L, XU Z, NOGUES I, YAO J, MOLLURA D, SUMMERS RM, (2016). DEEP CONVOLUTIONAL NEURAL NETWORKS FOR COMPUTER-AIDED DETECTION: CNN ARCHITECTURES, DATASET CHARACTERISTICS AND TRANSFER LEARNING. IEEE TRANS MED IMAGING 35(5):1285.
[90] YOSINSKI J, CLUNE J, BENGIO Y, LIPSON H, (2017). HOW TRANSFERABLE ARE FEATURES IN DEEP NEURAL NETWORKS?. IN: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS, PP 3320–3328.
[91] WU L, XIN Y, LI S, WANG T, HENG PA, NI D, (2017). CASCADED FULLY CONVOLUTIONAL NETWORKS FOR AUTOMATIC PRENATAL ULTRASOUND IMAGE SEGMENTATION. IN: 2017 IEEE 14TH INTERNATIONAL SYMPOSIUM ON BIOMEDICAL IMAGING (ISBI 2017). IEEE, PP 663–666
[92] Abdou MA, (20220. " Literature review: efficient deep neural networks techniques for medical image analysis". Neural Comput Appl, 34(8), 5791–5812. https://doi.org/10.1007/s00521-022-06960-9
[93] TRAN D, BOURDEV L, FERGUS R, TORRESANI L, PALURI M, (2016). DEEP END2END VOXEL2VOXEL PREDICTION. IN: PROCEEDINGS OF THE IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION WORKSHOPS, PP 17–24
[94] Pawar K, Chen Z, Jon Shah N, Egan GF., (2022). Suppressing motion artefacts in MRI using an Inception-ResNet network with motion simulation augmentation. NMR in Biomedicine; 35(4). doi:10.1002/nbm.4225
[95] WANG J, MACKENZIE JD, RAMACHANDRAN R, CHEN DZ, (2016). A DEEP LEARNING APPROACH FOR SEMANTIC SEGMENTATION IN HISTOLOGY TISSUE IMAGES. IN: INTERNATIONAL CONFERENCE ON MEDICAL IMAGE COMPUTING AND COMPUTER-ASSISTED INTERVENTION. SPRINGER, PP 176–184
[96] CHEN H, QI X, YU L, HENG PA: DCAN, (2016). DEEP CONTOUR-AWARE NETWORKS FOR ACCURATE GLAND SEGMENTATION. IN: PROCEEDINGS OF THE IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, PP 2487–2496
[97] GIRSHICK R, DONAHUE J, DARRELL T, MALIK J, (2016). REGION-BASED CONVOLUTIONAL NETWORKS FOR ACCURATE OBJECT DETECTION AND SEGMENTATION. IEEE TRANS PATTERN ANAL MACH INTELL 38(1):142–158.

[98] ZENG G, YANG X, LI J, YU L, HENG PA, ZHENG G (2017). 3D U-NET WITH MULTI-LEVEL DEEP SUPERVISION: FULLY AUTOMATIC SEGMENTATION OF PROXIMAL FEMUR IN 3D MR IMAGES. IN: INTERNATIONAL WORKSHOP ON MACHINE LEARNING IN MEDICAL IMAGING. SPRINGER, PP 274– 282

[99] ZENG G, ZHENG G, (2018). MULTI-STREAM 3D FCN WITH MULTI-SCALE DEEP SUPERVISION FOR MULTI-MODALITY ISOINTENSE INFANT BRAIN MR IMAGE SEGMENTATION. IN: 2018 IEEE 15TH INTERNATIONAL SYMPOSIUM ON BIOMEDICAL IMAGING (ISBI 2018). IEEE, PP 136–140

[100] RAVISHANKAR H, SUDHAKAR P, VENKATARAMANI R, THIRUVENKADAM S, ANNANGI P, BABU N, VAIDYA V, (2016). UNDERSTANDING THE MECHANISMS OF DEEP TRANSFER LEARNING FOR MEDICAL IMAGES. IN: DEEP LEARNING AND DATA LABELING FOR MEDICAL APPLICATIONS. SPRINGER, PP 188–196

[101] HUYSKENS DP, MAINGON P, VANUYTSEL L, REMOUCHAMPS V, ROQUES T, DUBRAY B, HAAS B, KUNZ P, CORADI T, B¨ UHLMAN R, ET AL, (2009). A QUALITATIVE AND A QUANTITATIVE ANALYSIS OF AN AUTO-SEGMENTATION MODULE FOR PROSTATE CANCER. ACTA RADIOL ONCOL 90(3):337–345.

[102] SIMPSON, A.L.; ANTONELLI, M.; BAKAS, S.; BILELLO, M.; FARAHANI, K.; VAN GINNEKEN, B.; KOPP-SCHNEIDER, A.; LANDMAN, B.A.; LITJENS, G.; MENZE, B.; ET AL. , (2019). A LARGE ANNOTATED MEDICAL IMAGE DATASET FOR THE DEVELOPMENT AND EVALUATION OF SEGMENTATION ALGORITHMS. *ARXIV* 2019, ARXIV:1902.09063.

[103] VAN GINNEKEN, B.; STEGMANN, M.B.; LOOG, M. , (2006). SEGMENTATION OF ANATOMICAL STRUCTURES IN CHEST RADIOGRAPHS USING SUPERVISED METHODS: A COMPARATIVE STUDY ON A PUBLIC DATABASE. *MED. IMAGE ANAL.*, *10*, 19–40. [CROSSREF] [PUBMED]

[104] MENZE, B.H.; JAKAB, A.; BAUER, S.; KALPATHY-CRAMER, J.; FARAHANI, K.; KIRBY, J.; BURREN, Y.; PORZ, N.; SLOTBOOM, J.; WIEST, R.; ET AL., (2014). THE MULTIMODAL BRAIN TUMOR IMAGE SEGMENTATION BENCHMARK (BRATS). *IEEE TRANS. MED. IMAGING*, *34*, 1993–2024. [CROSSREF]

[105] HEATH, M.; BOWYER, K.; KOPANS, D.; KEGELMEYER, P.; MOORE, R.; CHANG, K.; MUNISHKUMARAN, S., (2000). THE DIGITAL DATABASE FOR SCREENING MAMMOGRAPHY. IN PROCEEDINGS OF THE 5TH INTERNATIONAL WORKSHOP ON DIGITAL MAMMOGRAPHY, TORONTO, ON, CANADA, 11–14, PP. 212–218.

[106] BILIC, P.; CHRIST, P.F.; VORONTSOV, E.; CHLEBUS, G.; CHEN, H.; DOU, Q.; FU, C.-W.; HAN, X.; HENG, P.-A.; HESSER, J.; ET AL., (2019). THE LIVER TUMOR SEGMENTATION BENCHMARK (LITS). *ARXIV*, ARXIV:1901.04056.

[107] ARMATO, S.G., III; MCLENNAN, G.; BIDAUT, L.; MCNITT-GRAY, M.F.; MEYER, C.R.; REEVES, A.P.; ZHAO, B.; ABERLE, D.A.; HENSCHKE, C.I.; HOFFMAN, E.A.; ET AL., (2011). THE LUNG IMAGE DATABASE CONSORTIUM (LIDC) AND IMAGE DATABASE RESOURCE INITIATIVE (IDRI): A COMPLETED REFERENCE DATABASE OF LUNG NODULES ON CT SCANS. *MED. PHYS.*, *38*, 915–931. [CROSSREF] [PUBMED]

[108] MARCUS, D.S.; FOTENOS, A.F.; CSERNANSKY, J.G.; MORRIS, J.C.; BUCKNER, R.L., (2010). OPEN ACCESS SERIES OF IMAGING STUDIES: LONGITUDINAL MRI DATA IN NONDEMENTED AND DEMENTED OLDER ADULTS. *J. COGN. NEUROSCI.*, *22*, 2677–2684. [CROSSREF]

[109] STAAL, J.; ABRÀMOFF, M.D.; NIEMEIJER, M.; VIERGEVER, M.A.; VAN GINNEKEN, B., (2004). RIDGE-BASED VESSEL SEGMENTATION IN COLOR IMAGES OF THE RETINA. *IEEE TRANS. MED. IMAGING* , *23*, 501–509. [CROSSREF]

[110] SUCKLING, J.P. , (1994). THE MAMMOGRAPHIC IMAGE ANALYSIS SOCIETY DIGITAL MAMMOGRAM DATABASE. *DIGIT. MAMMO* , *17*, 375–386.

[111] FONSECA, C.G.; BACKHAUS, M.; BLUEMKE, D.A.; BRITTEN, R.D.; CHUNG, J.D.; COWAN, B.R.; DINOV, I.D.; FINN, J.P.; HUNTER, P.J.; KADISH, A.H.; ET AL., (2011). THE CARDIAC ATLAS PROJECT—AN IMAGING DATABASE FOR COMPUTATIONAL MODELING AND STATISTICAL ATLASES OF THE HEART. *BIOINFORMATICS* , *27*, 2288–2295. [CROSSREF] [PUBMED]

[112] P. N. SRINIVASU, A. K. BHOI, R. H. JHAVERI, G. T. REDDY, AND M. BILAL, (2021). "PROBABILISTIC DEEP Q NETWORK FOR REAL-TIME PATH PLANNING IN CENSORIOUS ROBOTIC PROCEDURES USING FORCE SENSORS," *JOURNAL OF REAL-TIME IMAGE PROCESSING*, 18(5),1773–1785.

[113] O. ÇIÇEK, A. ABDULKADIR, S. S. LIENKAMP, T. BROX, AND O. RONNEBERGER, (2016). "3D U-NET: LEARNING DENSE VOLUMETRIC SEGMENTATION FROM SPARSE ANNOTATION," IN *PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON MEDICAL IMAGE COMPUTING AND COMPUTER-ASSISTED INTERVENTION*, PP. 424–432, SPRINGER, ATHENS, GREECE.

[114] D. N. LEVIN, C. A. PELIZZARI, G. T. CHEN, C. T. CHEN, AND M. D. COOPER, (1988). "RETROSPECTIVE GEOMETRIC CORRELATION OF MR, CT, AND PET IMAGES," *RADIOLOGY*, 169(3), 817–823.

[115] S. BHATTACHARYA, P. K. R. MADDIKUNTA, Q. V. PHAM ET AL.,(2021). "DEEP LEARNING AND MEDICAL IMAGE PROCESSING FOR CORONAVIRUS (COVID-19) PANDEMIC: A SURVEY," *SUSTAINABLE CITIES AND SOCIETY*, VOL. 65, ARTICLE ID P.102589.

[116] J. MERKOW, A. MARSDEN, D. KRIEGMAN, AND Z. TU, (2016). "DENSE VOLUME-TO-VOLUME VASCULAR BOUNDARY DETECTION," IN *PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON MEDICAL IMAGE COMPUTING AND COMPUTER-ASSISTED INTERVENTION*, PP. 371–379, SPRINGER, ATHENS, GREECE.

[117] S. BHATTACHARYA, P. K. R. MADDIKUNTA, S. HAKAK ET AL., (2020). "ANTLIONRE-SAMPLING BASED DEEP NEURAL NETWORK MODEL FOR CLASSIFICATION OF IMBALANCED MULTIMODAL STROKE DATASET," *MULTIMEDIA TOOLS AND APPLICATIONS*.

[118] S. ROY, A. CARASS, P. L. BAZIN, AND J. L. PRINCE, (2011). "INTENSITY INHOMOGENEITY CORRECTION OF MAGNETIC RESONANCE IMAGES USING PATCHES," IN *MEDICAL IMAGING: IMAGE PROCESSING*VOL. 7962, ARTICLE ID 79621F.

[119] Y. GUO AND A. S. ASHOUR,(2019). "NEUTROSOPHIC SETS IN DERMOSCOPIC MEDICAL IMAGE SEGMENTATION," IN *NEUTROSOPHIC SET IN MEDICAL IMAGE ANALYSIS*, PP. 229–243, ACADEMIC PRESS, MA, USA.

[120] D. SHEN, G. WU, AND H. I. SUK, (2017). "DEEP LEARNING IN MEDICAL IMAGE ANALYSIS," *ANNUAL REVIEW OF BIOMEDICAL ENGINEERING*, 19, 221–248.

[121] H. CUI, H. ZHANG, G. R. GANGER, P. B. GIBBONS, AND E. P. XING, (2016). "GEEPS: SCALABLE DEEP LEARNING ON DISTRIBUTED GPUS WITH A GPUSPECIALIZED PARAMETER SERVER," IN *PROCEEDINGS OF THE ELEVENTH EUROPEAN CONFERENCE ON COMPUTER SYSTEMS*, PP. 1–16, LISBON, PORTUGAL.

[122] N. SRIVASTAVA, G. HINTON, A. KRIZHEVSKY, I. SUTSKEVER, AND R. SALAKHUTDINOV, (2014). "DROPOUT: A SIMPLE WAY TO PREVENT NEURAL NETWORKS FROM OVERFITTING," *JOURNAL OF MACHINE LEARNING RESEARCH*, 15(1), 1929–1958.

[123] S. IOFFE AND C. SZEGEDY, (2015). "BATCH NORMALIZATION: ACCELERATING DEEP NETWORK TRAINING BY REDUCING INTERNAL COVARIATE SHIFT," , HTTPS://ARXIV.ORG/ABS/1502.03167.

[124] K. KAMNITSAS, C. LEDIG, V. F. NEWCOMBE ET AL.,(2017). EFFICIENT MULTI-SCALE 3D CNN WITH FULLY CONNECTED CRF FOR ACCURATE BRAIN LESION SEGMENTATION," *MEDICAL IMAGE ANALYSIS*, 36, 61–78.

[125] E. GOCERI,(2019). "CHALLENGES AND RECENT SOLUTIONS, FOR IMAGE SEGMENTATION IN THE ERA OF DEEP LEARNING," IN *PROCEEDINGS OF THE 2019 NINTH INTERNATIONAL CONFERENCE ON IMAGE PROCESSING THEORY, TOOLS AND APPLICATIONS (IPTA)*, PP. 1–6, IEEE, STANBUL, TURKEY.