Research article

Reviewing 3D convolutional neural network approaches for medical image segmentation

Ademola E. Ilesanmi [a],[*], Taiwo O. Ilesanmi [b], Babatunde O. Ajayi [c]

[a] *University of Pennsylvania, 3710 Hamilton Walk, 6th Floor, Philadelphia, PA, 19104, United States*
[b] *National Population Commission, Abuja, Nigeria*
[c] *National Astronomical Research Institute of Thailand, Chiang Mai 50180, Thailand*

A B S T R A C T

Background: Convolutional neural networks (CNNs) assume pivotal roles in aiding clinicians in diagnosis and treatment decisions. The rapid evolution of imaging technology has established three-dimensional (3D) CNNs as a formidable framework for delineating organs and anomalies in medical images. The prominence of 3D CNN frameworks is steadily growing within medical image segmentation and classification. Thus, our proposition entails a comprehensive review, encapsulating diverse 3D CNN algorithms for the segmentation of medical image anomalies and organs.
Methods: This study systematically presents an exhaustive review of recent 3D CNN methodologies. Rigorous screening of abstracts and titles were carried out to establish their relevance. Research papers disseminated across academic repositories were meticulously chosen, analyzed, and appraised against specific criteria. Insights into the realm of anomalies and organ segmentation were derived, encompassing details such as network architecture and achieved accuracies.
Results: This paper offers an all-encompassing analysis, unveiling the prevailing trends in 3D CNN segmentation. In-depth elucidations encompass essential insights, constraints, observations, and avenues for future exploration. A discerning examination indicates the preponderance of the encoder-decoder network in segmentation tasks. The encoder-decoder framework affords a coherent methodology for the segmentation of medical images.
Conclusion: The findings of this study are poised to find application in clinical diagnosis and therapeutic interventions. Despite inherent limitations, CNN algorithms showcase commendable accuracy levels, solidifying their potential in medical image segmentation and classification endeavors.

1. Introduction

The proliferation of medical imaging technologies has led to a significant increase in the generation of medical images within healthcare institutions. Daily, a substantial volume of medical images is generated across global medical facilities. Diverse forms of medical imagery, including microscopy, X-ray, ultrasound (US), computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET), find extensive applications in imaging systems. The foremost advantage of medical images lies in their utility for asymptomatic conditions. Furthermore, they serve to identify injuries, conditions, and diseases in their incipient stages.

Among the various facets of medical image processing and analysis, medical image segmentation holds an impressive position. Image segmentation involves the recognition and demarcation of objects. Recognition pertains to the localization of anomalies and organs, while delineation encompasses outlining the procedure of recognition [1]. Accurate segmentation of medical images contributes to radiotherapy, disease quantification, and image-guided surgery. Put simply, segmentation tasks aid in the early detection, diagnosis, and treatment monitoring of ailments.

A diverse array of segmentation techniques exists, which encompasses traditional methods, graph-based approaches, semantic segmentation methods, classification and clustering methods, deformable methods, and other modalities [2,74]. Fig. 1(A–C), shows the diagram of different medical images.

Segmentation methods can be classified into three categories: manual, semi-automatic, and automatic procedures. Manual segmentation, deemed the gold standard, entails an expert radiologist's hand-drawn ground truth. Semi-automatic segmentation employs advanced algorithms to address manual segmentation issues, though some human intervention is still required. Fully automatic segmentation, on the other hand, doesn't rely on human involvement and utilizes both learning and non-learning-based techniques, representing the latest trend. Although manual segmentation is the benchmark in medical practice, it demands substantial time and resources, prompting the need for automated approaches [3]. Automatic segmentation leverages computer-aided diagnosis (CAD) systems to assist radiologists in clinical diagnosis, treatment planning, abnormal structure delineation, and region of interest identification.

The central objective of CAD is swift, accurate, automated segmentation to facilitate communication. Crucially, the segmentation procedure varies based on image modalities, body region localization, artifact types, and noise. There isn't a one-size-fits-all method for all medical images; hence, different algorithms must be considered [4,98]. The emergence of convolutional neural networks (CNNs) has revolutionized pattern recognition and computer vision, offering researchers a rapid, accurate segmentation solution. This study concentrates on the utilization of three-dimensional (3D) CNNs for the segmentation of abnormalities and organs in medical images. While two-dimensional (2D) CNNs were previously used for medical image segmentation, the recent trend involves the development of 3D CNNs tailored for effectively segmenting three-dimensional images or 2D slices. Notably, the key distinction lies in learning representations: 3D CNNs employ input volumes with filters to generate a 3D output mask. This review contributes to the body of knowledge by evaluating diverse methods employing 3D CNNs for MRI, CT, ultrasound, and other medical image segmentations.

The study's goal is to present the outcomes of various 3D CNN algorithms employed in investigating abnormalities and organ segmentation in medical images. This analysis delves into the latest trends, which include, assessment of the performance, preprocessing techniques, and validation metrics of 3D CNN methodologies. Furthermore, it discusses pertinent roadmaps and challenges for abnormalities and organ segmentation in MRI and ultrasound images.

The structure of this paper is as follows: Section 2 outlines the review's organization, search criteria, and methodologies. Section 3, discussed different databases and preprocessing techniques. Section 4 revealed statistical analyses of diverse algorithms, while Section 5 concludes the paper.

1.1. Review organization

This review's methodological framework is categorized into three main sections: (1) Encoder/Decoder, (2) Deep Convolutional Neural Networks (DCNNs) and Fully Connected Networks (FCNs), and (3) other. The comprehensive examination of CNN segmentation within these three domains will be presented. The "other" category encompasses attention models, combinations of various CNN, and similar cases. While 2D models have been employed for segmenting multiple medical images, this review will concentrate primarily on 3D images and Networks. A block diagram portraying the various 3D CNN approaches is illustrated in Fig. 2.
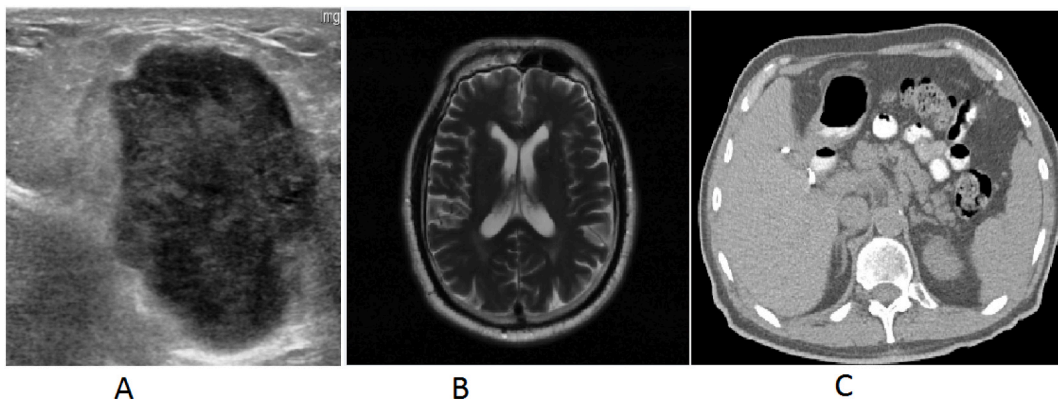

Fig. 1. Different medical images: a) Breast Ultrasound, b) MRI, 3) CT image.

1.2. Search criteria

Relevant research publication databases, including Springer, IEEE, Pubmed, Sciencedirect, Researchgate, and Google Scholar, were meticulously searched for pertinent articles. These selected papers were subsequently screened to identify those specific to the study area of 3D CNN segmentation in medical images. The compiled list of medical images returned indicated that MRI, CT and ultrasound were the most prevalent. To conduct the paper search, pertinent keywords were employed. After completing the keyword search process, a total of 500 papers were chosen. Keywords like 3D CNN, Medical images, segmentation of organs, segmentation of tumors, and abnormalities were utilized. A subsequent step involves search based on paper titles, leading to the selection of an additional 100 papers. In total, 600 papers were acquired through a combination of keyword and title searches. To streamline the scope, papers published before 2018 were excluded, and only those utilizing 3D CNN were retained for review. Papers focussed on animal parts and review papers were omitted from consideration.

1.3. Review of methods

This section provides an introduction to the prominent methodologies involving 3D CNNs utilized for the segmentation of medical images. The chosen publications are organized according to distinct medical image categories. The subsequent discussion delves into various 3D CNN methods, focusing on their specific CNN network types.

1.3.1. Encoder/decoder network

1.3.1.1. U-net. Originally tailored for biomedical and medical image segmentation, the U-Net framework [5,34] amalgamates low and high-level features to yield exceptional outcomes. Researchers have adapted the U-Net methodology to address abnormalities/organ segmentation within medical images. Comprising an encoder-decoder structure, the U-Net incorporates convolution layers, max-pooling (for downsampling reduction), deconvolution layers, and skip connections (refer to Ronneberger et al. [5] for comprehensive insights). However, due to the evolution of complex images, the 2D U-Net proved inadequate for 3D image segmentation, necessitating the creation of the 3D U-Net. Gcek et al. [6] improved the convention U-Net to a 3D model. The model consists of 4 3 × 3 convolutions, 2 deconvolutions, and 4 ReLU [7]. Several variants of 3D U-Net have been used for segmenting abnormalities/organs in medical images. For example, the research by Qamar et al. [8] combines the dense connection [79], residual connection [81], and the inception module [80] in a U-Net framework for segmentation. Yee et al. [19] proposed a U-Net variant that consists of 2 convolutions and 8 convolutions + dilations in the encoding stage. In the decoding stage, the network consists of 8 convolutions + dilations, 1 upsampling, and 1 convolution. Meyer et al. [21] proposed a variant of the U-Net that uses the anisotropic 3D multi-stream CNN architecture. The network consists of 27 convolutions + BN [78] + ReLU, 7 max-pooling, 12 upsampling, 4 dropouts, and 1 sigmoid.

Guo et al. [27] proposed the filterNet which is a variant of the U-Net for segmenting muscle in MRI. The network consists of 3 blocks
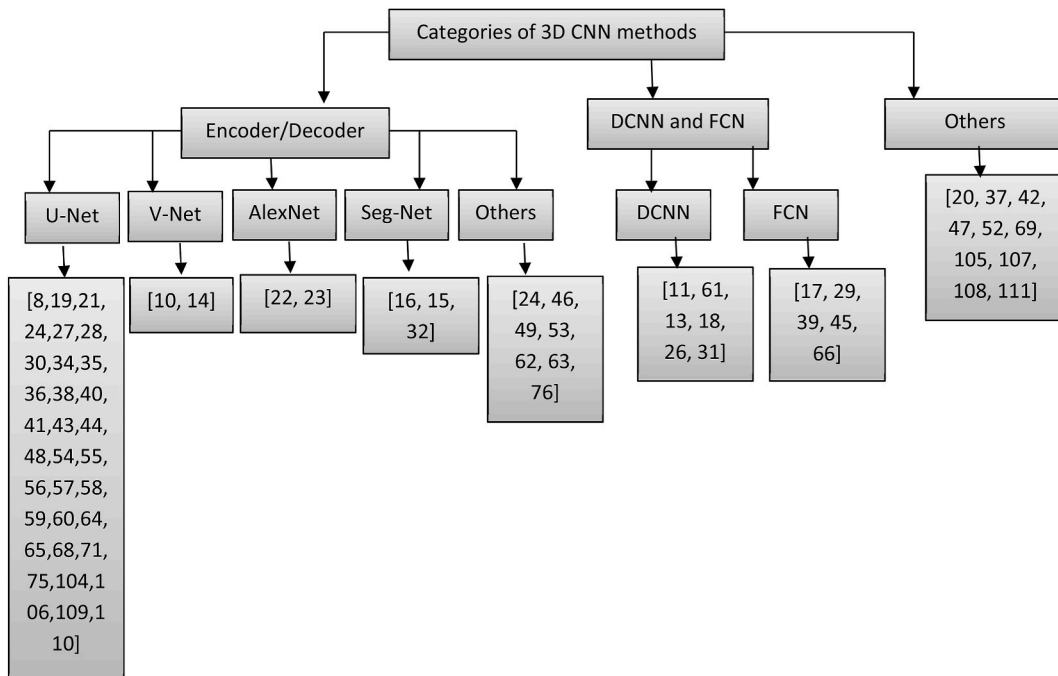

Fig. 2. Categories of 3D CNN methods for medical images.

with 7 convolutions, 4 BN + ReLU, 1 hardtanh, 1 element-wise addition [84], and short skip connections [83]. Recently, MRI has been used for detecting tumors in the breast. Qiao et al. [28] used the multi-label attention-guided joint phase network for segmenting tumors in the breast using MRI. This network consists of 5 encoders with 1 input intensity map and 4 post-contrast inputs. The input images are fed into each encoder and processed with the guided self-attention network [82] and concatenated for processing to the next layer in the decoder. Each encoder consists of 5 layers, with a convolution layer acting as a buffer at the end of each encoder. The decoder has 5 single layers with a self-attention module and a concatenation module. Jiang & Guo [30] used the improved 3D U-net for segmenting the brain in MRI. The network consists of 14 convolutions + BN + ReLU layers, 3 max-pooling, 3 upsampling layers and 3 concatenate and crop layers. The fully connected conditional random fields (CRF) [85] are used to produce the label maps (see Fig. 3).

Zhang et al. [24] proposed the triple intersecting U-Net from segmenting brain tumors. The network consists of 3 U-Net intertwined to produce a segmentation mask. The first U-Net consists of 9 layers each in both the encoder and decoder. The output of the first U-Net is concatenated with the original image to act as the input for the next U-Net. The second U-Net consists of 15 layers each in both the encoder and decoder. The last U-Net involves compatibility with the second to produce the final output. Su et al. [35] apply the multi-pathway architecture using the convolution and transpose convolution to segment tumors in MRI. This network was inspired by the U-Net and consists of 4 dilated convolutions, 6 convolutions, and 10 max-pooling in the encoder. While the decoder consists of 4 transpose convolutions, 4 concatenations, and 4 convolutions. Tang et al. [36] proposed the Dual Attention-based Dense SU-Net for segmenting tumors in head and neck regions. The network consists of 3 convolutions, 10 dense layers, 5 max-pooling, 4 PAM blocks, 5 unpooling layers, and 4 channel attention modules. This network combines the attention and the dense network with the U-Net as the backbone.

Calisto & Lai-Yuen [38] used the encoder/decoder framework to search for the best architecture for the segmentation of MRI. This network utilized two components 1) Search Procedure, and 2) Macro search structure [86]. The macro-search consists of 5 max-poolings, 5 transpose convolutions, 5 summation layers, and 1 convolution + softmax. Zhou et al. [40] proposed an efficient residual neural network (ERV-Net) for the segmentation of brain tumors in MRI. The network was first made computationally efficient with a subnetwork in the encoder phase. Then, another subnetwork uses the residual network to avoid degradation at the decoder phase. The network uses an encoder/decoder medium that consists of 13 convolutions, 13 shuffle unit strides, 5 concatenation layers, and 5 skip networks. Zhou et al. [41] proposed the iterative localization refinement for segmentation of the prostate using MRI. The encoder consists of 3 convolutions, 4 residual IBN blocks, and one 2D convolution. Meanwhile, the decoder consists of 3 residual Instance Batch Normalization (IBN [87]) blocks, 1 IBN block, 4 upsampling layers, and 1 single convolution. Overall, there are 17 layers with different convolutions, factors, and strides.

The U-Net method has also been used directly for segmentation without much modification. For example, Huo et al. [43] used the U-Net with several other networks for segmentation of the brain in MRI. Similarly, Coupe et al. [44] used the U-Net for segmenting the whole brian in MRI images. Luo et al. [64] used the U-Net for segmenting multiple organs using multiple modalities. Herrera et al. [68] used the U-Net for segmenting the brain in volumetric MRI. Niyas et al. [48] used the Res-U-Net to segment lesions in MRI. The network replaced the conventional U-Net block with two consecutive residual blocks. The encoder/decoder network consists of 4 double consecutive blocks, and 4 max-pooling. The decoder part consists of 4 convolutional layers, and 4 3D transpose convolutional layers. The encoder and decoder are concatenated with the concatenation layer, the same way it was performed in the traditional U-Net.

Cao et al. [56] proposed the dilated densely connected U-Net for segmentation of mass in breast ultrasound images. The network consists of 4 dense blocks, 3 upsampling layers, 3 up-samplings, 4 convolution + BN + ReLu layers, and 1 max-pooling. The dense block consists of feature maps, 1 transition layer, and 5 BN + ReLU + convolution. Yang et al. [57] proposed the patch-of-interest FuseNet with the hybrid loss for segmentation of tumors in ultrasound images. The network use information in the 2D U-Net and 3D U-Net for the prediction of the final segmentation. The 3D U-Net is joined with the pyramidal input and output patches. The network consists of 14 convolution + Bn/ReLU layers, 2 max-pooling, 4 deconvolution layers, 4 concatenation layers, and 2 convolution + sigmoid layers (see Fig. 4).

Zhou et al. [58] proposed a lightweight multi-scale network for the segmentation of tumors in ultrasound images. The lightweight network is an encoder-decoder network that consists of 9 stages. The network contains 4 down-sampling layers, 4 up-sampling layers,
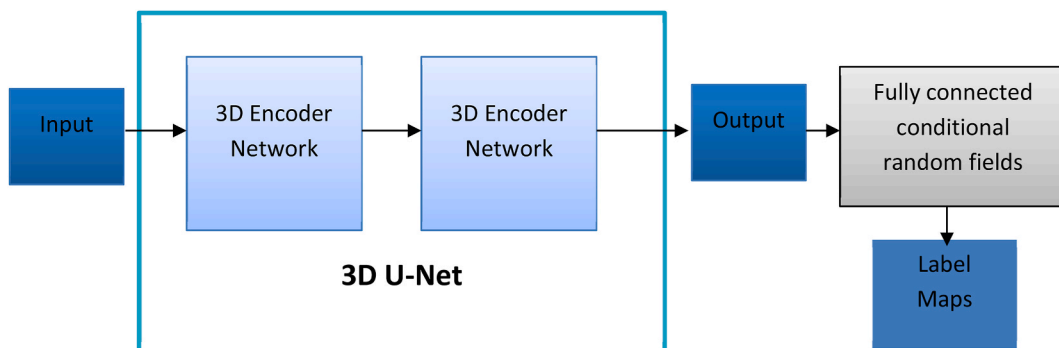
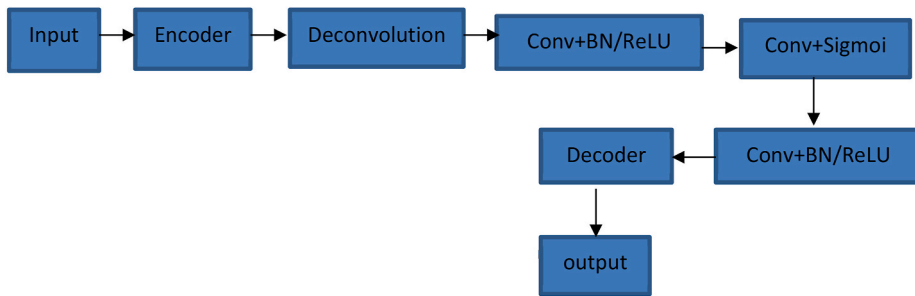

Fig. 3. 3D U-net with CRF [30].

Fig. 4. PIO FuseNet [57].

18 Conv + BN + ReLU, and 1 Conv + BN + ReLU. Yang et al. [59] proposed the contrastive rendering (C-Rend) framework to segment ovary and follicles in ultrasound images. The C-Rend is incorporated with a semi-supervised learning framework to prove reliable superior segmentation performance. The C-Rend is a modified U-Net that consists of an encoder-decoder, render [88], upsamplers, and identity layer [89]. Wang et al. [60] proposed the multi-scale context for lesion segmentation in breast ultrasound. The network used a mixed 2D and 3D convolution module for segmentation. The multi-scale block consists of 8 convolutions, 4 concatenation layers, 4 additional layers, and the U-Net network. Shirokikh et al. [65] propose the adaptive small-scale target localization for the segmentation of abnormalities in MRI. The network is a U-Net-based framework that consists of 3 average pooling [90] layers, 3 trilinear [91] blocks, and a single loss function calculation layer. Yang [71] proposed an improved U-Net framework for the segmentation of the brain in MRI. The traditional U-Net was improved, through the addition of the squeeze-and-excitation block [72], and a dilated convolution procedure [73]. Qin et al. [104] introduced a Node Growth Neural Architecture Search, termed NG-NAS, for the purpose of 3D medical image segmentation. This architecture adopts a U-shaped design, comprising a collection of nodes. Unlike the conventional approach of exploring within a restricted search space from a supernet, NG-NAS commences with a basic architecture containing just 5 nodes. It then progressively expands the optimal candidate node through a greedy procedure until satisfying the imposed constraint. The network was meticulously structured with two distinct generations of nodes, contributing to diverse network topologies, and it incorporates a selection of four candidate operators for each node. Liu et al. [106] present an approach, referred to as the "Prior-based 3D U-Net," which leverages prior knowledge for enhanced performance. The methodology unfolds in the following stages: 1) Initially, training data is utilized to train the 3D U-Net model and create an average shape model (ASM) of the knee joint, encompassing components like cartilage, distal femur, and proximal tibia. Subsequently, the trained 3D U-Net is employed to segment the distal femur and proximal tibia from new input knee MRI data. Then the newly input knee joint structures (distal femur and proximal tibia) are registered with the ASM via the construction of a registration transformation (RT). Next, the RT transformation then maps the ASM to the newly input MRI images, leading to the emergence of a predicted cartilage model. Finally, the predicted cartilage model, along with the distal femur and proximal tibia previously segmented by the initial U-Net step, serves as constraints for the subsequent 3D U-Net segmentation of the knee cartilage.

1.3.1.2. V-net. The V-Net is a 3D deformable structure proposed by Milletari et al. [9] and has been used to segment abnormalities and organs in medical images. Han et al. [10] cascaded the V-Net for whole heart segmentation of abnormal connections. This network consists of two V-Net frameworks joined and cascaded for heart segmentation. Jin et al. [14] used the improved V-Net (PBV-Net) to segment the prostate gland from MRI images. The network consists of 5 convolutions, 5 concatenations, 5 deconvolutions, and a single softmax and deconvolution layer.

1.3.1.3. Seg-Net. The Seg-Net [16] is a pixel-wise semantic segmentation framework that upsamples features in the decoder procedure with lower resolution input feature maps. Seg-Net provides a memory against accuracy tradeoff with good segmentation performance. Several modifications have been proposed for the segmentation of medical images. For example, Yamanakkanavar & Lee [15], proposed the M-Seg-Net that used the global attention mechanism [109] as a skip connection between the decoder and encoder. The network consists of 7 convolutions, 7 deconvolutions, 6 max-pooling, and 6 up-sampling layers. The global attention process the class localization details. Qin et al. [32] proposed the multi-scale discriminative network with the pyramid attention module (PAM [92]) and residual refinement [93] block (RRB) for segmentation of prostate in MRI. The Res layer [33] was used as the backbone encoder, while the PAM and RRB are used as the decoder. Each PAM consists of 2 summation layers, 1 multiplication and concatenation layer, and different convolutions. Meanwhile, each RRB consists of 3 convolutional layers, 1 passthrough, and 1 summation layer.

1.3.1.4. Alex-Net. The Alex-Net [22] is an 8-layer network with learnable parameters. The input format of the data source is in 3 dimensions with the network having 5 convolutions, 3 max-pooling, 1 dropout, and 5 ReLU layers. Chen et al. [23] proposed an improved Alex-Net to segment prostate tumors in MRI images. This network adds the BN and the global maximum pooling [94] algorithm to the conventional Alex-Net.

1.3.1.5. PSP-net. Pyramid Scene parsing network (PSPNet) [25] is a semantic segmentation model that learns and trains modules that can segment and classify pixels. Yan et al. [24] used the PSP-Net for segmenting prostate in MRI. The network is an encoder/decoder

framework that consists of 9 convolutions, global average pooling, 3 upsampling, and 4 skip connections.

1.3.1.6. Others. Ke et al. [46] proposed the encoder/decoder dense network for segmentation of carcinoma in MRI. The network is a self-constrained method with an end-to-end pattern. It consists of 9 convolutions, 21 upsampling layers, 4 denseblock_down layers, 8 denseblock_down + transition block down layers, 5 concatenation layers, 1 multiplication layer, global average pooling, and 1 fully connected layer (See Fig. 5).

Celik & Talu [49] combine the generative adversarial network (GAN) [50] with the atrous convolution for segmentation of the brain in MRI. The network uses the collaborative effects of the generator and the discriminator for segmentation. The generator is an encoder/decoder U-shaped mechanism that consists of 11 convolutions, an atrous convolution feature pyramid (ACFP) with a position attention module (PAM) [52], and 2 concatenation layers. The ACFP and PAM act as the connection between the encoder and the decoder and consist of 3 convolutions, 2 upsampling, and 2 concatenation layers. The discriminator was used to classify the output of the generator. Cui et al. [53] proposed an end-to-end learning-based method (TSegNet) for the segmentation of teeth in dental models. The network consists of 18 fully connected layers, and 3 channel concatenation in the encoder. While the decoder consists of 9 fully connected layers and 2 channel concatenation. Organs segmented include, incisors, canines, premolars, and molars for the upper and lower teeth.

Dong et al. [62] proposed the deep atlas network for segmentation of ventricle in echocardiography. The network consists of a transformer and deformable networks. The transformer network contains 5 convolutions, 4 BN + ReLU, 4 max-pooling layers, and a single FC layer. The deformable network used trilinear sampling [63] to produce the final output. Meng et al. [75] proposed the 3D CNN with dual parts for the segmentation of the liver in MRI. The network has two parts 1) local pathway and 2) global pathways. Each pathway consists of 8 blocks and 3 concatenation layers. Each block contains 1 convolution, 1 BN, and 1 ReLU. At the end of the pathways, 2 fully connected dropouts with a softmax are used to produce the final output. The summary of methods in this section is available in Table 1.

1.4. Summary of encoder/decoder methods

Researchers have explored the possibility of amalgamating two or more encoder/decoder frameworks. For instance, Martins et al. [55] harnessed insights from both U-Net [110] and V-Net architectures to segment cerebral ventricles in ultrasound images. This emerging trend yielded impressive accuracy, while also addressing associated intricacies common with medical image segmentation. Notably, the 3D U-Net emerged as the most widely employed network for segmenting ultrasound images. In breast ultrasound segmentation, detecting common abnormalities like tumors and lesions is crucial, prompting authors to tailor the U-Net framework accordingly. Intriguingly, a combination of 2D and 3D approaches holds promise for enhancing accuracy in medical image segmentation. Mlynarski et al. [75] skillfully merge both 2D and 3D techniques for brain segmentation in MRI scans. This synergy between distinct CNN types has the potential to mitigate redundancy and overfitting. Fig. 6 illustrates a graph detailing the frequency of segmentation utilization for each framework.

The encoder/decoder network exhibits certain common limitations, such as the potential loss of neighboring or local information due to successive pooling operations or convolution striding, resulting in decreased feature map resolution. Encoders serve as feature extraction units, while decoders function as fusion and processing components. A valuable takeaway from this section underscores the efficacy of encoder and decoder networks for medical image segmentation. In summary, the U-Net, AlexNet, PSPS-Net, Seg-Net, V-Net, and GAN emerge as dependable 3D segmentation frameworks for medical images.

1.5. DCNN and FCN methods

Deep Convolutional Neural Networks (DCNNs) are specifically crafted to acquire knowledge about filters and systematically
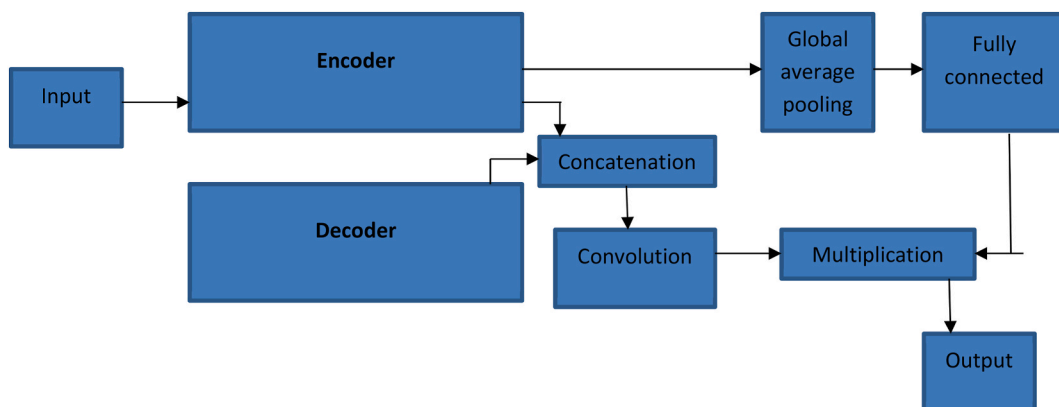

Fig. 5. 3D self-constrained DenseNet [46].

Table 1
Summary of Encoder/Decoder method.

CNN Type	Ref	Common image type	Popular Abnormalities/ Organ	Coverage publication year	Average (DICE)	Average (Hausdorff)
U-Net	See Fig. 2	Brain, tumor, prostrate	MRI, CT, US	2019, 2020, 2021, 2022, and 2023	0.96	2.04
V-Net	See Fig. 2	Heart, prostate	MRI	2020, 2021	0.94	1.52
Seg-Net	See Fig. 2	Brain, prostate	MRI	2020, 2021	0.96	3.16
GAN	See Fig. 2	Ventricle, Brain	MRI, Echocardiography	2020, 2022	0.96	1.15
AlexNet and PSP-Net	See Fig. 2	Prostate	MRI	2021	0.97	1.02
Others	See Fig. 2	Liver, teeth, Carcinoma	MRI	2020, 2021	0.96	2.02
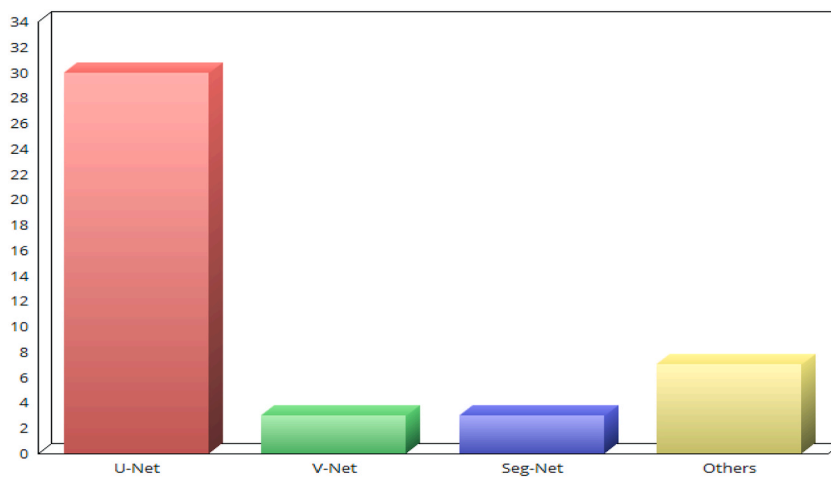
Fig. 6. Number of the framework for encoder/decoder.

amalgamate these acquired filters. Their predominant application lies within segmentation and classification endeavors, particularly when handling extensive datasets. Conversely, Fully Connected Networks (FCNs) encompass architectures wherein layers establish localized connections, thereby sidestepping the need for dense parameters. FCNs predominantly find application in semantic segmentation tasks, demonstrating remarkable efficiency in their training process.

1.5.1. DCNN

Charron et al. [11], used the 3D convolutional neural network (DeepMedic [12]) to segment and detect brain metastases on MRI images. The network uses a multi-scale CNN [109] with 2 convolutional pathways [54] and 11 layers (see Fig. 7).

Zhou et al. [13] proposed a DCNN framework that consists of a backbone, atrous convolution [95], and classification layers. Interestingly, the DeepMedic architecture is the inspiration for this research. Wei et al. [18] proposed the cascaded nested network (CasNet) for segmentation of brain MRI. The CasNet is a twin nested segmentation network that consists of 5 side blocks, 3 core blocks + pooling, 3 core blocks + deconvolution, 2 core blocks, and 1 convolution. Each core block consists of 2 convolutions, 2 ReLU, 2 BN, and a single dropout. Each side block consists of 1 core block, 1 convolution, and 1 deconvolution. Dolz et al. [26] proposed the ensemble 3D deep CNN for segmenting infant brain in MRI. This network was built on the DeepMedic and uses a combination of 2D on 3D models. The network consists of convolutional layers, fully connected layers, and classification layers.

Zhou et al. [31] used the atrous-convolution feature pyramid method as the backbone for feature learning to segment tumors in MRI. The CRF was used as the post-processing method to obtain structural segmentation. The network consists of 4 parts 1) backbone, 2) feature pyramid, 3) classification, and 4) processing. The backbone consists of 9 convolutions and 3 atrous convolution layers. While the feature pyramid has 9 atrous convolutional layers and 2 upsampling layers. The classification part consists of 6 convolution layers and 1 softmax. Finally, the processing part consists of a fully connected layer and CRF. Qui & Ren [61] proposed a deeply supervised network for automatic segmentation of Liver and heart volumetric images. The network conducts volume-to-volume learning and eliminates overfitting and gradient vanishing problems. This network consists of 6 convolutions, 2 max-pooling, 5 deconvolutions, and 3 softmax layers. The output from different softmax layers is fused to produce the final label.
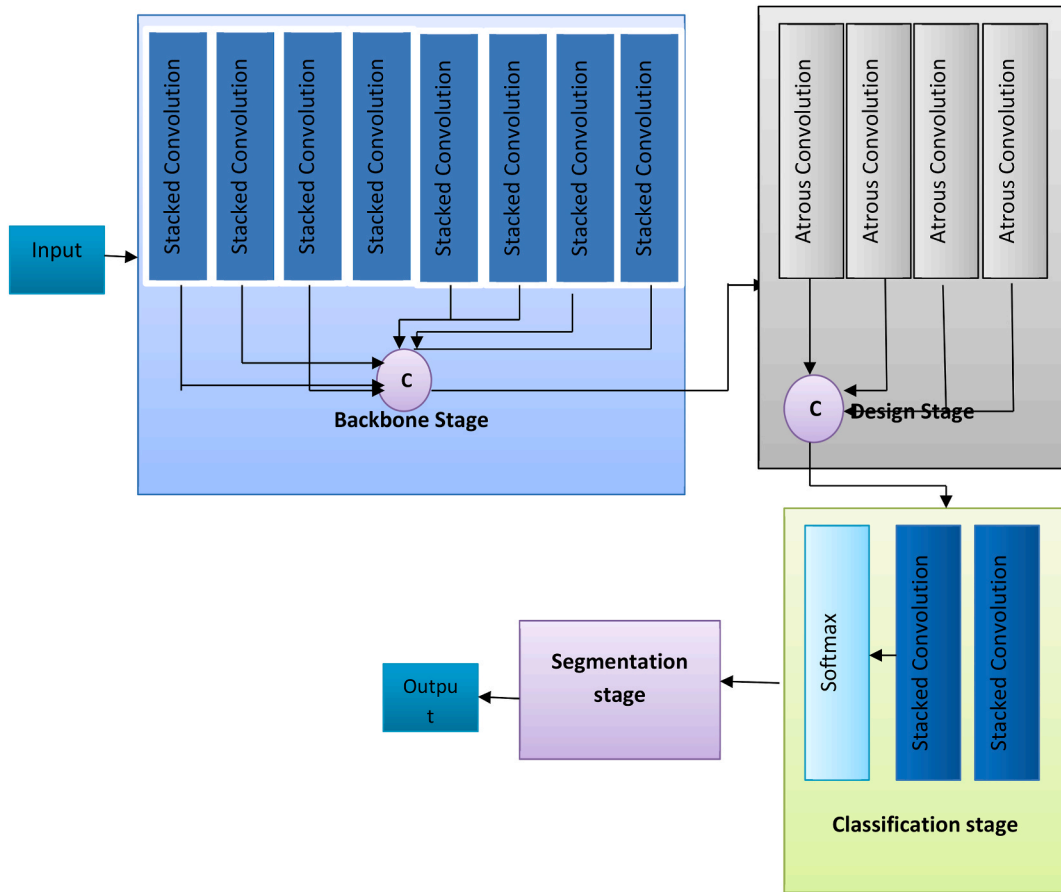
Fig. 7. DenseAFPNet [11].

1.5.2. FCN

Wu & Tang [17] proposed the multi-atlas 3D FCN ensemble model for the segmentation of the brain in MRI images. The network consists of 8 convolutions, 6 max-pooling, 3 convolutions + ReLU, 3 deconvolutions + ReLU, skip connections and 1 softmax layer. Hu et al. [29] proposed the improved FCN methods for segmenting ventricles in MRI. The network consists of 16 convolutions, 2 max-pooling, 3 upsampling, and 1 softmax layer (see Fig. 8).

Kong et al. [39] proposed the combination of the global attention mechanism (GAM) with the local attention mechanism (LAM) [96] for the segmentation of brain tumors. The GAM segments the discriminate learning with a weight-allocation loss function while the LAM reverses the effective feature guidance associated with the united loss function at different levels. The network consists of 4 LAM, 27 convolutions, 4 max-pooling layers, 1 sigmoid, residual network, 2 concatenations, and 4 skip connections. Li et al. [45] used the multi-scale modality dropout learning network to localize the intervertebral disc in MRI. The network involves three pathways which integrate spatial information into multiple-scale inputs. The first pathway consists of 7 convolutions, 1 max-pooling, and 1
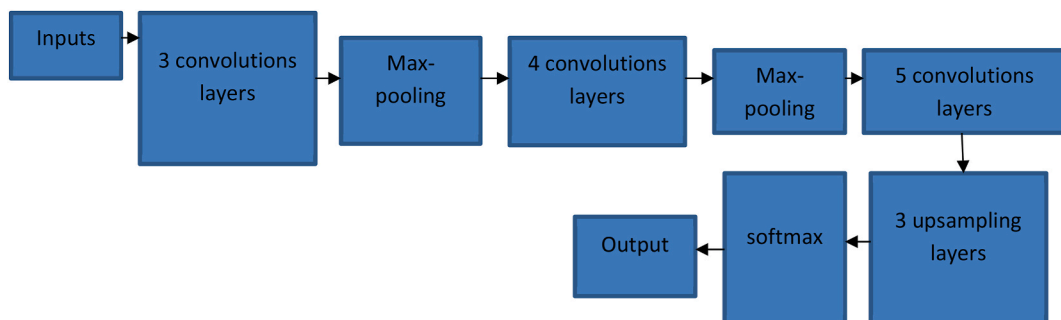

Fig. 8. 3D-ASM [29].

deconvolution. The second pathway consists of 5 convolutions, while the third consists of 4 convolutions, 1 max-pooling, and 1 deconvolution. A merge operation concatenates the three pathways, then a final convolution and softmax are used to produce the final segmentation (See Fig. 9).

Dolz et al. [66] proposed an FCN that consists of 25 convolutions and Fc layers to segment the brain in MRI. The network is an improvement to the baseline FCN [51,67].

1.5.3. Summary of deep CNN and fully CNN methods

A significant challenge associated with CNNs is the vanishing gradient issue. Nonetheless, certain researchers have harnessed the 3D learning and interference approach to mitigate this concern [61]. Employing techniques like cascaded networks, attention pooling, and the Siamese network has also proven to be effective in addressing the vanishing gradient challenge. In this context, the widely recognized DeepMedic [12] framework has been adopted for brain segmentation. This network, in conjunction with other methodologies, enhances segmentation effectiveness and elevates accuracy rates. While the FCN and DCNN architectures may not be inherently tailored for medical image segmentation, they nonetheless yield favorable outcomes when applied to this domain. The summary of methods in this section is available in Table 2.

1.6. Other methods

There are other 3D CNN methods which are not part of the categories mentioned earlier, these methods are used for segmenting abnormalities and organs in medical images. The research by Ye et al. [20] proposed a parallel pathway dense neural network, a variant of the dense network. The network consists of 3 layers; input, convolution, and fully connected layer. The inputs are downsampled into two parts and the outputs of the parts are weighted and combined with the context pathway fusion method. Zhang et al. [37] combined the active learning, attention mechanism, and deep supervision mode to segment the brain in MRI. The network consists of 5 Resblocks, 2 poolings, 3 upsampling, and 2 hybrid attention mechanisms. Each Resblock consists of 2 convolutions, and 2 BN, ReLU layers. The hybrid attention mechanism consists of a channel attention module and a spatial attention module. It has 1 Convolution, BN, ReLU layer, 1 global average pooling, 1 convolution, ReLU, and 1 convolution, Sigmoid. Two concatenation modules are used to connect the layers with the output phase.

Chen et al. [42] proposed the voxelwise residual network (voxResNet) for segmentation of the brain in MRI. The network consists of convolution layers, 4 BN, ReLU layers, and 6 VoxRes layers. The network also includes 4 deconvolution layers, 4 classification layers, and a single fusion layer. Each voxRes module consists of 2 convolutions, 2 BN, ReLU, and 1 concatenation layer. Karayegen & Aksahin [47] proposed the combination of a simplified deep learning network with the semantic network for the segmentation of brain tumors in MRI. The network is composed of 2 convolutions + BN + ReLU, 2 max-pooling, 2 upsampling, 2 transpose convolution + ReLU, and a softmax layer. Semantic segmentation uses the neural network while deep learning uses upsampling and downsampling procedures. Zanjani et al. [52] proposed the masked new computation network (Mask-MCNet) for the segmentation of teeth in intra-oral scans. The network consists of 1) Backbone network 2) Monte Carlo ConvNet [97], 3) Detection Network 4) Localization network and 5) Mask generation network. The network consists of 11 X_convolutions, 9 fully connected layers, 13 BN, 4 Monte Carlo convolution layers, and 3 loss functions. Wang et al. [69] used the 3D P-Net [70] for the segmentation of multiple organs in MRI. The network contained 6 blocks, made up of 18 convolutions, 2 softmax layers, and 1 concatenation layer (see Fig. 10).

Wang et al. (2023) Introduced the Tensorized Transformer Network designed for the segmentation of 3D medical images. This network encompasses three pivotal functions: 1) A multi-scale transformer, featuring layers-fusion, introduced to effectively capture intricate contextual interaction information. 2) The Cross Shared Attention (CSA) module, founded on pHash similarity fusion (pSF), is thoughtfully crafted to extract comprehensive multi-variate dependency features on a global scale. 3) The Tensorized Self-Attention (TSA) module is put forth to address the challenge of handling a substantial number of parameters. Furthermore, this module can be
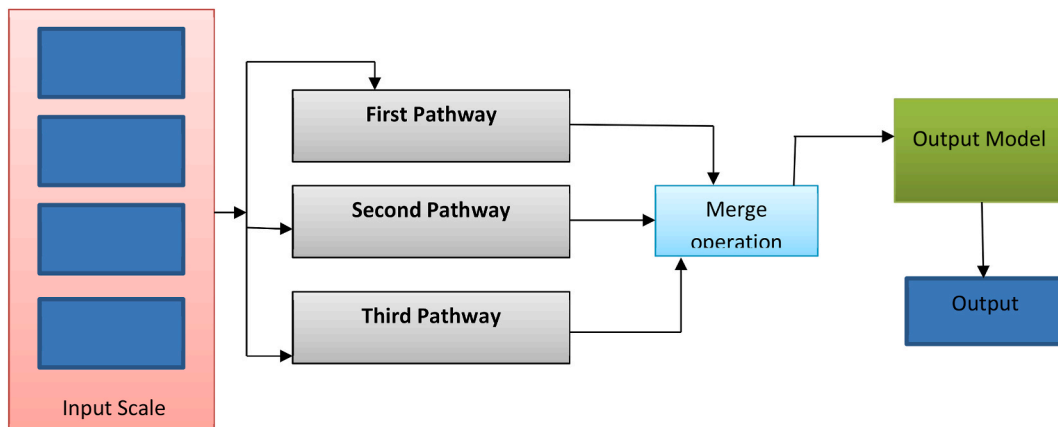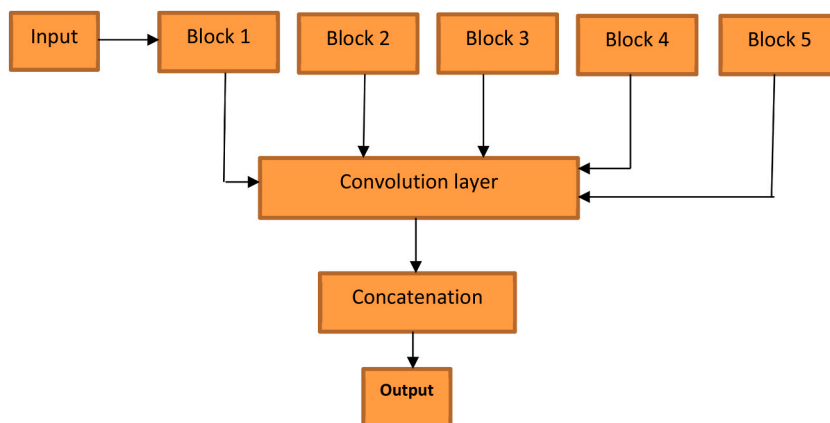

Fig. 9. Multi-path network [45].

Table 2
Summary of DCNN and FCN method.

CNN Type	Ref	Common image type	Popular Abnormalities/ Organ	Coverage publication year	Average (DICE)	Average (Hausdorff)
DCNN	See Fig. 2	Metastases, Tumor, Brain, Liver and Heart	MRI	2018, 2020, 2021	0.91	2.34
FCN	See Fig. 2	Tumor, Ventricles, Brain	MRI	2018, 2021, 2022	0.93	2.62

Fig. 10. 3D P-net [69].

seamlessly integrated into other models. Li et al. [107] harnessed deep learning for the automated segmentation and detection of prostate cancer in MRI scans, introducing a novel framework named "3D Mask R–CNN." This innovative model integrates a 3D convolutional neural network within its Region Proposal Network (RPN) component to effectively extract intricate features from the MRI images. These extracted features play a pivotal role in generating region proposals, a critical step in object detection. The core of the methodology involves leveraging the 3D convolutional neural network within the RPN to extract informative features from the MRI data. These features are subsequently utilized to construct anchor boxes of varying scales and aspect ratios. This anchor box strategy enhances the accuracy of region proposal generation within the RPN, as it aligns the proposals more closely with the characteristics of the extracted features.

Guo et al. [108] present the causal knowledge fusion (CKF) framework for cardiac image segmentation. Initially, the CKF framework employs causal intervention techniques to acquire the anatomical factor while discarding the modality factor. Subsequently, the CKF framework introduces a 3D hierarchical attention mechanism to extract multi-scale information from 3D cardiac images. The summary of methods in this section is available in Table 3.

2. Datasets and image preprocessing

Among the scrutinized papers, the BRAT [99–101] and LITs2017 [102] datasets emerged as the most commonly employed datasets. Their widespread utilization can be attributed to their online availability and ease of access. Intriguingly, we noted that these same datasets found application in the segmentation of both abnormalities and organs within medical images. A condensed compilation of image preprocessing methodologies for the segmentation of abnormalities and organs is provided in Table 4.

Table 3
Summary of Other methods.

Author	Ref	Year	Image type	CNN Type	Abnormalities/Organ	Accuracies (DICE)
Ye et al.	[20]	2021	MRI	DenseNet	Tumor	0.88
Zhang et al.	[37]	2021	MRI	Quality-driven deep learning	Brain	0.98
Chen et al.	[42]	2018	MRI	Voxel residual network	Brain	0.91
Karayegen & Aksahin	[47]	2021	MRI	Semantic segmentation and deep learning	Tumor	0.99
Zanjani et al.	[52]	2021	Intraoral scan	Mask-MCNet	Tooth	0.99
Wang et al.	[69]	2018	MRI	P-Net	Multiple	0.90
Wang et al.	[105]	2023	MRI	Self-attention network	Heart, Brain, Uterine	0.93
Li et al.	[107]	2023	MRI	3D Mask R–CNN	Prostate	0.95
Guo et al.	[108]	2023	MRI, CT	Attention mechanism	Cardiac	0.97
Cao et al.	[111]	2023	MRI	Shuttle Attention mechanism	Brain	0.87

Interpolation, rotation, equalization, normalization, and filtering are among the preprocessing techniques that have exerted a notable influence on the successful segmentation of abnormalities and organs in medical images. Research findings underscore that an optimal dataset should possess relevance, usability, and high quality [77]. Consequently, it is our belief that the combination of suitable datasets and proficient preprocessing techniques holds the potential to enhance segmentation accuracy.

3. Summary of data

A total of 63 pertinent research papers have been investigated within this review. Widely adopted metrics such as the Jaccard similarity coefficient, Dice measure, F1 scores, mean surface distance, Hausdorff distance, and those pertaining to the Medical Image Computing and Computer-Assisted Intervention (MICCAI) [8,14,60] serve as prevalent performance indicators for the segmentation of medical images. Prominent organs subjected to segmentation include the brain, prostate, teeth, and ventricle (refer to Fig. 2). As for abnormalities, tumors, lesions, and metastases account for the most frequently segmented anomalies. It is intriguing to observe that the focal points of researchers' segmentation endeavors revolve around the brain and brain tumors. A comprehensive overview concerning the distribution of publications across different years is depicted in Fig. 11.

We restricted our consideration to papers released within the preceding six years. The most substantial volume of publications was registered in the years 2020 and 2021. Across the spectrum, 42 papers were dedicated to encoder/decoder segmentation, while 21 papers were directed towards other segmentation. On the whole, the average dataset proportions employed for testing and training purposes were 175 and 75, respectively. Notably, the requisition of data for Encoder/Decoder segmentation surpassed that necessary for other segmentation. The information concerning the quantity of papers employed for abnormalities and organ segmentation can be found in Fig. 12. The figure illustrates that there is a greater count of papers utilized for organ segmentation compared to those employed for abnormalities segmentation. To be precise, there are 19 papers dedicated to abnormalities segmentation, whereas organ segmentation is represented by 44 papers. In the context of medical images, "abnormalities" pertain to images that deviate from a healthy depiction. Such images may reveal structural impairments, encompassing injuries, lesions, inflammation, swellings, bleeding, and tumors. Body regions under scrutiny encompass the chest, abdomen, pelvis, head, and neck. Organs residing within these regions include the heart, liver, biliary tract, kidneys, spleen, bowels, pancreas, adrenal glands, uterus, ovaries, and prostate gland (see Ref. [112] for details).

4. Conclusions and future directions

Within the ambit of this investigation, we meticulously reviewed a total of 60 articles which harnessed the capabilities of 3D CNN methods for the precise segmentation of both organs and abnormalities within medical images. This classification process systematically categorized the articles into distinct groups, namely encoder/decoder, DCNN, FCN, and other methodologies. The commendable performance demonstrated by the 3D CNN frameworks has underscored their capacity to yield noteworthy segmentation and classification accuracies. Nonetheless, it's noteworthy that these networks remain relatively less utilized in comparison to their 2D CNN counterparts. Operating on the foundational tenets of the CNN framework, the 3D CNN harnesses components such as max-pooling, convolutional kernels, upsampling, and concatenation layers to achieve its segmentation objectives with efficacy.

4.1. Problems of the existing system, and future directions

This study aims to comprehensively review methods pertaining to 3D CNN segmentation of organs and abnormalities in medical images. The performance of various CNN algorithms was systematically evaluated, revealing a spectrum of outcomes ranging from excellent to very good accuracies. Medical images have become an increasingly valuable adjunct in detecting abnormalities, with the current progress of 3D CNN for diagnosis marking significant advancements. CNN algorithms have furnished clinicians with invaluable insights that aid in drawing conclusions regarding specific pathologies. Multiple 3D CNN techniques have effectively contributed to disease diagnosis and the prediction of ailments in patients.

However, certain limitations must be addressed to ensure effective abnormalities and organ segmentation. The first of these constraints is the paucity of available public volumetric databases for comprehensive analysis. The integration of traditional

Table 4
Image preprocessing.

| Reference	Method	Notes
[10,14,20,21,23,28,30,52, 59]	Interpolation and rotation	This method interpolates, flips, and rotates the images
[11,19,28]	Anonymization, reorientation, and resampling	Methodic resampling and rescaling.
[17,24,35,42,47]	Histogram matching, equalization, and smoothing	Accurate identification of the region of interest size and center of the image patch
[27,44]	Bias field correction and inhomogeneity method	Reduces image intensity in-homogeneity by estimating high-frequency content.
[35,40,48,49,55,76]	Normalization, cropping, blurring, registration, and slicing	Shape and edge smoothing of images.
[45,65]	Mean subtraction and intensity clipping	Subtraction of the input data by the mean of intensity of the whole data
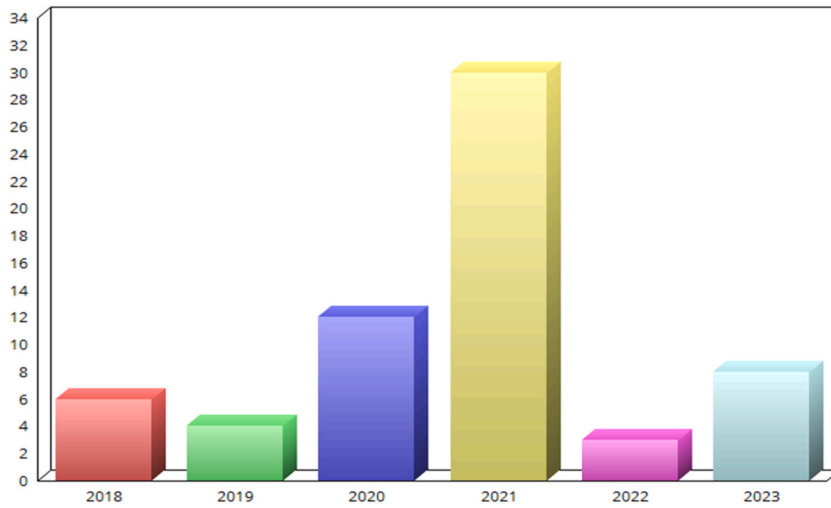
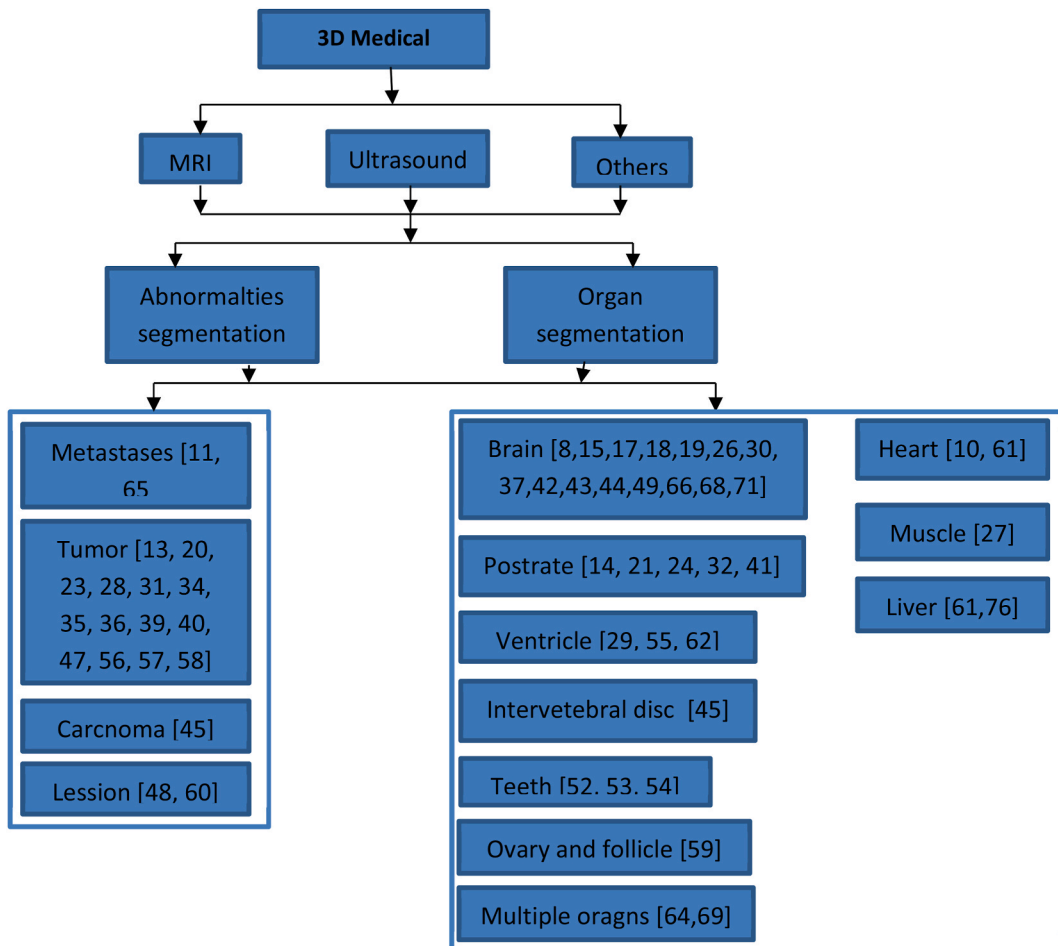**Fig. 11.** Number of papers published yearly.

Fig. 12. Number of papers published relating to organ and abnormalities.

preprocessing methods with CNN models poses another challenge. Furthermore, the limited availability of annotated ground truth images represents an obstacle. The underutilization of 3D CNN frameworks for segmenting dental models stands as another limitation. Additionally, there's a need to develop a comprehensive 3D CNN capable of simultaneously segmenting multiple organs within a specific body region. Lastly, there's a concern regarding over-reliance on specific 3D CNN algorithms for medical image segmentation. These limitations are accompanied by potential solutions and future directions.

4.1.1. Insufficient volumetric databases and data availability

The accessibility of publicly available datasets for medical image segmentation remains limited. The utilization and distribution of datasets are often skewed, with a noticeable balance between private and publicly accessible datasets. This observation implies that datasets are often confined to specific institutions such as universities or hospitals, possibly due to privacy policies that safeguard patient data from experimental use. A potential resolution involves revising patient data protection rights to permit the use of anonymized patient data without compromising sensitive information. Another issue encountered is the inadequacy of data within the datasets themselves. Many datasets are found to have constrained quantities of data for both testing and training. It's postulated that datasets containing less than 400 volumetric data points are deemed insufficient. As a recommended course of action, analysts in medical image processing are encouraged to allocate time and resources for collecting data directly from clinicians, potentially involving the recruitment and training of graduate students under the guidance of principal investigators and mentors.

4.1.2. Effective integration of preprocessing and CNN models

Preprocessing methods are a vital component of medical image segmentation pipelines, enhancing both numerical outcomes and visual evaluations. However, effectively harmonizing these preprocessing techniques with 3D CNN models presents a notable challenge. To surmount this challenge, researchers are urged to develop efficient artificial intelligent preprocessing models that seamlessly integrate with CNN segmentation models. The shared domain between the preprocessing and segmentation models could facilitate this integration, streamlining the process.

4.1.3. Scarcity of annotated ground truth images

An integral element of medical image processing lies in the provision of meticulously annotated ground truth data. This serves as the benchmark for learning in CNN models, aiding their training. The production of such ground truth data is labor-intensive, necessitating the involvement of at least three experienced clinicians. The collective input of these clinicians results in a consensus-based ground truth. However, it has been observed that several datasets lack ground truth images due to the challenges involved in obtaining them [103]. To mitigate this issue, researchers are advised to employ a combination of different algorithms with minimal human intervention to generate ground truth markings. Alternatively, adopting deep learning mock-up strategies, such as weakly supervised methods, can be employed for unlabeled data.

4.1.4. Limited usage of 3D CNN frameworks for certain image modalities

Through this review, it becomes apparent that MRI and ultrasound modalities dominate the segmentation landscape. This could be attributed to the research focus on specific organs within the body. MRI demonstrates efficacy in brain tumor detection, while ultrasound excels in detecting breast tumors. Encouragement is extended to medical image analysts to explore the segmentation of organs and abnormalities across diverse modalities, potentially leading to improved accuracies.

4.1.5. Multi-organ and multi-abnormality segmentation using 3D CNNs

While effective segmentation of individual organs within specific body regions is established, further advancements could be realized by extending this capability to simultaneously segment multiple organs within the same region. To address this, the utilization of deep 3D CNN networks, such as the Siamese network, is recommended. Such networks are equipped to accommodate the distinctive characteristics of different organs within a given body region.

4.1.6. Diversification of 3D CNN methods for segmentation

The analysis of reviewed papers indicates a prevalence of 3D CNN methods built upon encoder/decoder networks (e.g., U-Net, attention, and adversarial networks). The performance of these networks, especially U-Net, in medical image tasks could account for this prevalence. While the utilization of U-Net is merited, a single-network reliance is impractical. Researchers are advised to explore a broader spectrum of CNN architectures, employing effective configurations to rival the U-Net model. Furthermore, the DCNN and FCN models also demonstrate the potential for accurate results, urging researchers to develop sophisticated models utilizing these frameworks.

4.1.7. Real-life performance of 3D CNN methods

The segmentation of organs and abnormalities using diverse 3D CNN methodologies is a significant endeavor. However, it's equally crucial to evaluate these algorithms in real-world scenarios. There remains uncertainty over whether certain algorithms will perform optimally when applied to practical applications. To mitigate this concern, analysts and computer vision engineers are encouraged to foster close collaborations with clinicians, ensuring the gradual integration of CNN algorithms into real-life applications.

4.1.8. Standardization of measuring metrics

The adoption of diverse evaluation metrics for quantitative result reporting is a prevalent practice among researchers. This

diversity in metrics can lead to varied assessments, posing challenges in comparative analyses. While this review consolidates two common accuracies from the studies, further steps can be taken to minimize metric proliferation. A suggested approach is to designate MICCAI metrics as the standard for organ segmentation in CT images (as outlined in Section 4).

4.1.9. Addressing high computation space demands of CNNs

It's widely recognized that CNN algorithms demand significant computational resources. The availability of adequate computational space plays a pivotal role in achieving higher segmentation accuracies. To alleviate this concern, the proposal for creating additional cloud computing platforms (like Collab) with substantial computational capacity is made. Such platforms could facilitate researchers in running their applications without incurring substantial costs.

4.2. Conclusion

This study delved into the exploration of 3D CNN methodologies for the segmentation of organs within medical images. Various approaches to segmenting abnormalities within medical images were thoroughly examined. Insights into forthcoming trends have been projected, accompanied by a comprehensive evaluation of the merits and drawbacks of multiple 3D CNN techniques.

In summation, 3D CNNs exhibit promising capabilities in effectively segmenting both organs and abnormalities within medical images. The accrued advantages from employing 3D algorithms hold the potential for widespread adoption in the future. Significantly, collaborative efforts are essential to forge the development of novel and potent algorithms catering to the segmentation needs of abnormalities and organs within medical images.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author.

CRediT authorship contribution statement

Ademola E. Ilesanmi: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Data curation, Conceptualization. **Taiwo O. Ilesanmi:** Writing – original draft, Data curation. **Babatunde O. Ajayi:** Writing – original draft, Visualization, Formal analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Jayaram K. Udupa, Dewey Odhner, Liming Zhao, Yubing Tong, M. Monica, S. Matsumoto, Krzysztof C. Ciesielski, Alexandre X. Falcao, Pavithra Vaideeswaran, Victoria Ciesielski, Babak Saboury, Syedmehrdad Mohammadianrasanani, Sanghun Sin, Raanan Arens, Drew A. Torigian, Body-wide hierarchical fuzzy modeling, recognition, and delineation of anatomy in medical images, Med. Image Anal. 18 (5) (2014) 752–771.

[2] A.E. Ilesanmi, U. Chaumrattanakul, S.S. Makhanov, Methods for the segmentation and classification of breast ultrasound images: a review, J Ultrasound 24 (2021) 367–382, https://doi.org/10.1007/s40477-020-00557-5.

[3] Royal College of Radiologists. The older radiologist. https://www.rcr.ac.uk/clinical-radiology/service-delivery/sustainable-future-diagnostic-radiology/older-radiologist Accessed November 9, 2016.

[4] N. Sharma, L.M. Aggarwal, Automated medical image segmentation techniques, J. Med. Phys. 35 (2010) 3–14.

[5] O. Ronneberger, P. Fischer, T. Brox, U-net: convolutional networks for biomedical image segmentation, in: Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, 5–9 October 2015, pp. 234–241. Munich, Germany.

[6] Ö. Çiçek, A. Abdulkadir, S.S. Lienkamp, T. Brox, O. Ronneberger, 3D U-Net: learning dense volumetric segmentation from sparse annotation, in: Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, 2016, pp. 424–432. Athens, Greece, 17–21 October.

[7] A. Krizhevsky, I. Sutskever, G. Hinton, Advances in Neural Information Processing Systems, 2012, https://doi.org/10.1016/j.protcy.2014.09.007.

[8] Saqib Qamar, Hai Jin, Ran Zheng, Parvez Ahmad, Mohd Usama, A variant form of 3D-UNet for infant brain segmentation, Future Generat. Comput. Syst. 108 (2020) 613–623.

[9] F. Milletari, N. Navab, S.-A. Ahmadi, V-net: fully convolutional neural networks for volumetric medical image segmentation, in: Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV), 25–28 October 2016, pp. 565–571. Stanford, CA, USA.

[10] Tao Han, Roberto F. Ivo, Douglas de A. Rodrigues, Solon A. Peixoto, Victor Hugo C. de Albuquerque, P. Pedro, Rebouças Filho, Cascaded volumetric fully convolutional networks for whole-heart and great vessel 3D segmentation, Future Generat. Comput. Syst. 108 (2020) 198–209.

[11] Odelin Charron, Alex Lallement, Delphine Jarnet, Noblet Vincent, Jean-Baptiste Clavier, Philippe Meyer, Automatic detection and segmentation of brain metastases on multimodal MR images with a deep convolutional neural network, Comput. Biol. Med. 95 (2018) 770–778, 43–54.

[12] Kamnitsask/deepmedic, GitHub. (n.d.). https://github.com/Kamnitsask/deepmedic. (Accessed 28 April 2017).

[13] Zexun Zhou, Zhongshi He, Meifeng Shi, Jinglong Du, Dingding Chen, 3D dense connectivity network with atrous convolutional feature pyramid for brain tumor segmentation in magnetic resonance imaging of human heads, Comput. Biol. Med. 121 (2020) 103766.

[14] Yao Jin, Guang Yang, Ying Fang, Ruipeng Li, Xiaomei Xu, Yongkai Liu, Xiaobo Lai, 3D PBV-Net: an automated prostate MRI data segmentation method, Comput. Biol. Med. 128 (2021) 104160.

[15] Nagaraj Yamanakkanavar, Bumshik Lee, A novel M-SegNet with global attention CNN architecture for automatic segmentation of brain MRI, Comput. Biol. Med. 136 (2021) 104761.

[16] V. Badrinarayanan, A. Kendall, R. Cipolla, SegNet: a deep convolutional encoder-decoder architecture for image segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 39 (12) (2017) 2481–2495.

[17] Jiong Wu, Xiaoying Tang, Brain segmentation based on multi-atlas and diffeomorphism guided 3D fully convolutional network ensembles, Pattern Recogn. 115 (2021) 107904.

[18] Jie Wei, Zhengwang Wu, Li Wang, Toan Duc Bui, Liangqiong Qu, Pew-Thian Yap, Yong Xia, Gang Li, Dinggang Shen, A cascaded nested network for 3T brain MR image segmentation guided by 7T labeling, Pattern Recogn., S0031–3203(21)596-3.

[19] Evangeline Yee, Da Ma, Karteek Popuri, Shuo Chen, Hyunwoo Lee, Vincent Chow, Cydney Ma, Lei Wang, Mirza Faisal Beg, 3D hemisphere-based convolutional neural network for whole-brain MRI segmentation, Computerized Medical Imaging, and Graphics, S0895–6111(21)149-X.

[20] Fangyan Ye, Yingbin Zheng, Hao Ye, Xiaohao Han, Yuxin Li, Jun Wang, Jian Pu, Parallel pathway dense neural network with weighted fusion structure for brain tumor segmentation, Neurocomputing 425 (2021) 1–11.

[21] Anneke Meyer, Grzegorz Chlebus, Marko Rak, Daniel Schindele, Schostak Martin, Bram van Ginneken, Andrea Schenk, Hans Meine, K. Horst, Hahn, andreas schreiber, christian hansen, anisotropic 3D multi-stream CNN for accurate prostate segmentation from multi-planar MRI, Comput. Methods Progr. Biomed. 200 (2021) 105821.

[22] Y. Keizhevsk, I. Stuskever, G. E Hinton, ImageNet classification with deep convolutional neural networks[C], in: Proceedings of the Advances in Neural Information Processing Systems, South Lake Tahoe, US, 2012, pp. 1097–1105.

[23] Jun Chen, Zhechao Wan, Jiacheng Zhang, Wenhua Li, Yanbing Chen, Yuebing Li, Yue Duan, Medical image segmentation and reconstruction of prostate tumor based on 3D AlexNet, Comput. Methods Progr. Biomed. 200 (2021) 105878, 9-11 July 2010.

[24] Lingfei Yan, Dawei Liu, Xiang Qi, Yang Luo, Tao Wang, Dali Wu, Haiping Chen, Yu Zhang, Qing Li, PSP net-based automatic segmentation network model for prostate magnetic resonance imaging, Comput. Methods Progr. Biomed. 207 (2021) 106211.

[25] Liang-Chieh Chen, George Papandreou Iasonas Kokkinos, Kevin Murphy, Alan L. Yuille, Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs (2014) 1–4, arXiv:1412.7062.

[26] Dolz Jose, Christian Desrosiers, Li Wang, Jing Yuan, Dinggang Shen, Ismail Ben Ayed, Deep CNN ensembles and suggestive annotations for infant brain MRI segmentation, Comput. Med. Imag. Graph. 79 (2020) 101660.

[27] Zhihui Guo, Honghai Zhang, Zhi Chen, Ellen van der Plas, Laurie Gutmann, Daniel Thedens, Peggy Nopoulos, Milan Sonka, Fully automated 3D segmentation of MR-imaged calf muscle compartments: neighborhood relationship enhanced fully convolutional network, Comput. Med. Imag. Graph. 87 (2021) 101835.

[28] Mengyun Qiao, Shiteng Suo, Fang Cheng, Hua Jia, Dan Xue, Yi Guo, Jianrong Xu, Yuanyuan Wang, Three-dimensional breast tumor segmentation on DCE-MRI with a multilabel attention-guided joint-phase-learning network, Comput. Med. Imag. Graph. 90 (2021) 101909.

[29] Huaifei Hu, Ning Pan, Haihua Liu, Liman Liu, Tailang Yin, Zhigang Tu, Automatic segmentation of left and right ventricles in cardiac MRI using 3D-ASM and deep learning, Signal Process. Image Commun. 96 (2021) 116303.

[30] Han Jiang, Yanrong Guo, Multi-class multimodal semantic segmentation with an improved 3D fully convolutional networks, Neurocomputing 391 (2020) 220–226.

[31] Zexun Zhou, Zhongshi He, Yuanyuan Jia, AFPNet: a 3D fully convolutional neural network with atrous-convolution feature pyramid for brain tumor segmentation via MRI images, Neurocomputing 402 (2020) 235–244.

[32] Xiangxiang Qin, Yu Zhu, Wei Wang, Shaojun Gui, Bingbing Zheng, Peijun Wang, 3D multi-scale discriminative network with multi-directional edge loss for prostate zonal segmentation in bi-parametric MR images, Neurocomputing 418 (2020) 148–161.

[33] S. Chen, K. Ma, Y. Zheng, Med3D: Transfer Learning for 3D Medical Image Analysis, 2019 00625 arXiv: 1904.

[34] Jinjing Zhang, Jianchao Zeng, Pinle Qin, Lijun Zhao, Brain tumor segmentation of multi-modality MR images via triple intersecting U-Nets, Neurocomputing 421 (2021) 195–209.

[35] Jindong Sun, Yanjun Peng, Yanfei Guo, Dapeng Li, Segmentation of the multimodal brain tumor image used the multi-pathway architecture method based on 3D FCN, Neurocomputing 423 (2021) 34–45.

[36] Pin Tang, Chen Zu, Mei Hong, Rui Yan, Xingchen Peng, Jianghong Xiao, Wuf Xi, Jiliu Zhou, Luping Zhou, Yan Wang, DA-DSUnet: dual Attention-based Dense SU-net for automatic head and neck tumor segmentation in MRI images, Neurocomputing 435 (2021) 103–113.

[37] Zhenxi Zhang, Jie Li, Chunna Tian, Zhusi Zhong, Zhicheng Jiao, Xinbo Gao, Quality-driven deep active learning method for 3D brain MRI segmentation, Neurocomputing 446 (2021) 106–117.

[38] Maria Baldeon Calisto, K. Susana, Lai-Yuen, EMONAS-Net: efficient multiobjective neural architecture search using surrogate-assisted evolutionary algorithm for 3D medical image segmentation, Artif. Intell. Med. 119 (2021) 102154.

[39] Deting Kong, Xiyu Liu, Yan Wang, Dengwang Li, Jie Xue, 3D hierarchical dual-attention fully convolutional networks with hybrid losses for diverse glioma segmentation, Knowl. Base Syst. 237 (2022) 107692.

[40] Xinyu Zhou, Xuanya Li, Kai Hu, Yuan Zhang, Zhineng Chen, Xieping Gao, ERV-Net: an efficient 3D residual neural network for brain tumor segmentation, Expert Syst. Appl. 170 (2021) 114566.

[41] Wenhui Zhou, XingTao, Wei Zhan, Lili Lin, Automatic segmentation of 3D prostate MR images with iterative localization refinement, Digit. Signal Process. 98 (2020) 102649.

[42] Hao Chen, Dou Qi, Lequan Yu, Jing Qin, Pheng-Ann Heng, VoxResNet: deep voxelwise residual networks for brain segmentation from 3D MR images, Neuroimage 170 (2018) 446–455.

[43] Yuankai Huo, Zhoubing Xu, Yunxi Xiong, Katherine Aboud, Prasanna Parvathaneni, Shunxing Bao, Camilo Bermudez, Susan M. Resnick, Laurie E. Cutting, Bennett A. Landman, 3D whole brain segmentation using spatially localized atlas network tiles, Neuroimage 194 (2019) 105–119.

[44] Pierrick Coup, Boris Mansencal, Michael Clement, Remi Giraud, Baudouin Denis de Senneville, Ta Vinh-Thong, Vincent Lepetit, V. Jose, Manjon, AssemblyNet: a large ensemble of CNNs for 3D whole brain MRI segmentation, Neuroimage 219 (2020) 117026.

[45] Xiaomeng Li, Dou Qi, Hao Chen, Chi-Wing Fu, Xiaojuan Qi, L. Daniel, Belavý, gabriele armbrecht, dieter felsenberg, guoyan zheng, pheng-ann heng, 3D multi-scale FCN with random modality voxel dropout learning for intervertebral disc localization and segmentation from multi-modality MR images, Med. Image Anal. 45 (2018) 41–54.

[46] Liangru Ke, Yishu Deng, Weixiong Xia, Mengyun Qiang, Xi Chen, Kuiyuan Liu, Bingzhong Jing, Caisheng He, Chuanmiao Xie, Xiang Guo, Lv Xing, Chaofeng Li, Development of a self-constrained 3D DenseNet model in automatic detection and segmentation of nasopharyngeal carcinoma using magnetic resonance images, Oral Oncol. 110 (2020) 104862.

[47] Gookay Karayegen, Mehmet Feyzi Aksahin, Brain tumor prediction on MR images with semantic segmentation by using deep learning network and 3D imaging of tumor region, Biomed. Signal Process Control 66 (2021) 102458.

[48] S. Niyas, S. Chethana Vaisali, T.G. Iwrin Show, S. Chandrika, Vinayagamani, Chandrasekharan Kesavadas, Jeny Rajan, Segmentation of focal cortical dysplasia lesions from magnetic resonance images using 3D convolutional neural networks, Biomed. Signal Process Control 70 (2021) 102951.

[49] Gaffari Çelik, Muhammed Fatih Talu, A new 3D MRI segmentation method based on generative adversarial network and atrous convolution, Biomed. Signal Process Control 71 (2022) 103155.

[50] I.J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial networks, Commun. ACM 63 (2014) 139–144, https://doi.org/10.1145/3422622.

[51] Z. Zhou, Z. He, Y. Jia, AFPNet: a 3D fully convolutional neural network with atrous-convolution feature pyramid for brain tumor segmentation via MRI images, Neurocomputing 402 (2020) 235–244.

[52] Farhad Ghazvinian Zanjani, Arash Pourtaherian, Svitlana Zinger, David Anssari Moin, Claessen Frank, Cherici Teo, Sarah Parinussa, H.N. Peter, de With, Mask-MCNet: tooth instance segmentation in 3D point clouds of intra-oral scans, Neurocomputing 453 (2021) 286–298.

[53] Zhiming Cui, Changjian Li, Nenglun Chen, Guodong Wei, Runnan Chen, Yuanfeng Zhou, Dinggang Shen, Wenping Wang, TSegNet: an efficient and accurate tooth segmentation network on 3D dental model, Med. Image Anal. 69 (2021) 101949.

[54] Jianda Zhang, Chunpeng Li, Qiang Song, Lin Gao, Yu-Kun Lai, Automatic 3D tooth segmentation using convolutional neural networks in harmonic parameter space, Graphical Models 109 (2020) 101071.

[55] Matthieu Martin, Sciolla Bruno, Michäel Sdika, Philippe Qu'etin, Philippe Delachartre, Automatic segmentation and location learning of neonatal cerebral ventricles in 3D ultrasound data combining CNN and CPPN, Comput. Biol. Med. 131 (2021) 104268.

[56] Xuyang Cao, Houjin Chen, Yanfeng Li, Yahui Peng, Shu Wang, Lin Cheng, Dilated densely connected U-Net with uncertainty focus loss for 3D ABUS mass segmentation, Comput. Methods Progr. Biomed. 209 (2021) 106313.

[57] Hongxu Yang, Caifeng Shan, Bouwman Arthur, Alexander F. Kolen, H. Peter, N. de With, Efficient and robust instrument segmentation in 3D ultrasound using patch-of-interest-FuseNet with hybrid loss, Med. Image Anal. 67 (2021) 101842.

[58] Yue Zhou, Houjin Chen, Yanfeng Li, Qin Liu, Xuanang Xu, Shu Wang, Pew-Thian Yap, Dinggang Shen, Multi-task learning for segmentation and classification of tumors in 3D automated breast ultrasound images, Med. Image Anal. 70 (2021) 101918.

[59] Xin Yang, Haoming Li, Yi Wang, Xiaowen Liang, Chaoyu Chen, Xu Zhou, Fengyi Zeng, Jinghui Fang, Alejandro Frangi, Zhiyi Chen, Ni Dong, Contrastive rendering with semi-supervised learning for ovary and follicle segmentation from 3D ultrasound, Med. Image Anal. 73 (2021) 102134.

[60] Hongyu Wang, Jiaqi Cao, Jun Feng, Yilin Xie, Di Yang, Baoying Chen, Mixed 2D and 3D convolutional network with multi-scale context for lesion segmentation in breast DCE-MRI, Biomed. Signal Process Control 68 (2021) 102607.

[61] Dou Qi, Lequan Yu, Hao Chen, Yueming Jin, Xin Yang, Jing Qin, Pheng-Ann Heng, 3D deeply supervised network for automated segmentation of volumetric medical images, Med. Image Anal. 41 (2017) 40–54.

[62] Suyu Dong, Gongning Luo, Clara Tam, Wei Wang, Kuanquan Wang, Shaodong Cao, Bo Chen, Henggui Zhang, Shuo Li, Deep atlas network for efficient 3D left ventricle segmentation on echocardiography, Med. Image Anal. 61 (2020) 101638.

[63] M. Jaderberg, K. Simonyan, A. Zisserman, et al., Spatial transformer networks, in: Advances in Neural Information Processing Systems, 2015, pp. 2017–2025.

[64] Xiangde Luo, Guotai Wang, Tao Song, Jingyang Zhang, Michael Aertsen, Deprest Jan, Sebastien Ourselin, Tom Vercauteren, Shaoting Zhang, MIDeepSeg: minimally interactive segmentation of unseen objects from medical images using deep learning, Med. Image Anal. 72 (2021) 102102.

[65] Boris Shirokikh, Alexey Shvetsov, Alexandra Dalechina, Egor Krivov, Valery Kostjuchenko, Andrey Golanov, Victor Gombolevskiy, Sergey Morozov, Mikhail Belyaev, Accelerating 3D medical image segmentation by adaptive small-scale target localization, J. Imaging 7 (2021) 35.

[66] J. Dolz, C. Desrosiers, I. Ben Ayed, 3D fully convolutional networks for subcortical segmentation in MRI: A large-scale study, Computer Vision and Pattern Recognition,arXiv:1612.03925. https://doi.org/10.48550/arXiv.1612.03925.

[67] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 3431–3440.

[68] A.M. Herrera, A.J. Cuadros-Vargas, H. Pedrini, Improving semantic segmentation of 3D medical images on 3D convolutional neural networks, in: XLV Latin American Computing Conference, vol. 2019, CLEI), 2019, pp. 1–8, https://doi.org/10.1109/CLEI47609.2019.235102.

[69] G. Wang, W. Li, M.A. Zuluaga, et al., Interactive medical image segmentation using deep learning with image-specific fine tuning, IEEE Trans. Med. Imag. 37 (7) (2018) 1562–1573, https://doi.org/10.1109/tmi.2018.2791721.

[70] G. Wang, et al., DeepIGeoS: a deep interactive geodesic framework for medical image segmentation [Online]. Available: https://arxiv.org/abs/1707.00652, 2017.

[71] Zhuqing Yang, A novel brain image segmentation method using an improved 3D U-net model, Sci. Program. 2021 (2021) 4801077, https://doi.org/10.1155/2021/4801077, 10 pages.

[72] J. Hu, L. Shen, S. Albanie, G. Sun, E. Wu, Squeeze-and-Excitation networks, IEEE Trans. Pattern Anal. Mach. Intell. 42 (8) (2020) 2011–2023.

[73] B. Li, Z.-W. Tan, P.P. Shum, C. Wang, Y. Zheng, L.j. Wong, Dilated convolutional neural networks for fiber Bragg grating signal demodulation, Opt Express 29 (5) (2021) 7110–7123.

[74] Ademola Enitan Ilesanmi, Oluwagbenga Paul Idowu, Stanislav S. Makhanov, Multiscale superpixel method for segmentation of breast ultrasound, Comput. Biol. Med. 125 (2020) 103879.

[75] Pawel Mlynarski, Hervé Delingette, Antonio Criminisi, Nicholas Ayache, 3D convolutional neural networks for tumor segmentation using long-range 2D context, Comput Med Imaging Graph 73 (2019) 60–72, https://doi.org/10.1016/j.compmedimag.2019.02.001.

[76] Meng Lu, Yaoyu Tian, Sihang Bu, Liver tumor segmentation based on 3D convolutional neural network with dual scale, J. Appl. Clin. Med. Phys. 21 (1) (2020) 144–157.

[77] L.M. Koesten, E. Kacprzak, J.F.A. Tennison, E. Simperl, The trials and tribulations of working with structured data: -a study on information seeking behavior, in: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, ACM, New York, NY, USA, 2017, pp. 1277–1289, https://doi.org/10.1145/3025453.3025838.

[78] Sergey Ioffe, Christian Szegedy, Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, 2015 arXiv:1502.03167.

[79] G. Huang, Z. Liu, L. Van Der Maaten, K.Q. Weinberger, Densely Connected Convolutional Networks, 2017, pp. 2261–2269, https://doi.org/10.1109/CVPR.2017.243.

[80] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, in: Proc. IEEE Conf. Comput. Vis. Pattern Recognit, 2015, pp. 1–9.

[81] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, Deep residual learning for image recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Las Vegas, NV, USA, 2016, pp. 770–778, arXiv:1512.03385.

[82] V. Mnih, N. Heess, A. Graves, K. Kavukcuoglu, Recurrent models of visual attention, in: Z. Ghahramani, M. Welling, C. Cortes, N.D. Lawrence, K.Q. Weinberger (Eds.), Advances in Neural Information Pro-cessing Systems 27: Annual Conference on Neural Information Processing Systems 2014, 2014, pp. 2204–2212. December 8-13 2014, Montreal, Quebec, Canada.

[83] C. Liu, et al., Auto-DeepLab: hierarchical neural architecture search for semantic image segmentation, in: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), vol. 2019, 2019, pp. 82–92.

[84] C.-Y. Lee, S. Xie, P. Gallagher, Z. Zhang, Z. Tu, Deeply-supervised nets, in: Artificial Intelligence and Statistics, 2015, pp. 562–570.

[85] Ruizhe Li, Xin Chen, An efficient interactive multi-label segmentation tool for 2D and 3D medical images using fully connected conditional random field, Comput. Methods Progr. Biomed. 213 (2022) 106534.

[86] H.R. Lourenço, O.C. Martin, T. Stützle, Iterated local search: framework and applications, in: M. Gendreau, J.Y. Potvin (Eds.), Handbook of Metaheuristics. International Series in Operations Research & Management Science, vol. 272, Springer, Cham, 2019, https://doi.org/10.1007/978-3-319-91086-4_5.

[87] Hyeonseob Nam, Hyo-Eun Kim, Batch-instance normalization for adaptively style-invariant neural networks, Adv. Neural Inf. Process. Syst. (2018).

[88] Hao Su, Charles R. Qi, Yangyan Li, Leonidas J. Guibas, Render for cnn: viewpoint estimation in images using cnns trained with rendered 3d model views, in: The IEEE International Conference on Computer Vision (ICCV), 2015.

[89] K. He, X. Zhang, S. Ren, J. Sun, Identity mappings in deep residual networks, in: B. Leibe, J. Matas, N. Sebe, M. Welling (Eds.), Computer Vision – ECCV 2016. ECCV 2016. Lecture Notes in Computer Science, 9908, Springer, Cham, 2016, https://doi.org/10.1007/978-3-319-46493-0_38.

[90] Y. LeCun, et al., Gradient-based learning applied to document recognition, Proc. IEEE 86 (11) (1998) 2278–2324.

[91] Heliang Zheng, Jianlong Fu, Zha Zheng-Jun, Jiebo Luo, Looking for the devil in the details: learning trilinear attention sampling network for fine-grained image recognition proceedings of the, in: IEEE Conference on Computer Vision and Pattern Recognition, 2019, pp. 5012–5021.

[92] Hanchao Li, Pengfei Xiong, Jie An, Lingxue Wang, Pyramid Attention Network for Semantic Segmentation, ArXiv, 2018 10180 abs/1805.

[93] Feng Lin, Wengang Zhou, Jiajun Deng, Bin Li, Yan Lu, Houqiang Li, Residual refinement network with attribute guidance for precise saliency detection, Assoc. Comput. Mach. 17 (3) (2021) 19.

[94] Mudau, T. (https://stats.stackexchange.com/users/139737/tshilidzi-mudau), What is global max-pooling layer and what is its advantage over max-pooling layer?, URL (version: 2017-11-10): https://stats.stackexchange.com/q/308218.

[95] Liang-Chieh Chen, Papandreou George, Florian Schroff, Hartwig Adam, Rethinking Atrous Convolution for Semantic Image Segmentation, ArXiv, 2017 05587 abs/1706.

[96] Suriyadeepan Ramamoorthy, Attention mechanism: benefits and applications, Blog, Saama (2018). April 19. (Accessed 13 November 2019).

[97] Hermosilla Pedro, Tobias Ritschel, Pere-Pau Vazquez, Alvar Vinacua, Timo Ropinski, Monte Carlo convolution for learning on non-uniformly sampled point clouds, ACM Trans. Graph. (37) (2018) 1–12.

[98] Juri Yanase, Evangelos Triantaphyllou, A systematic survey of computer-aided diagnosis in medicine: past and present developments, Expert Syst. Appl. 138 (2019).

[99] B.H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, et al., The multimodal brain tumor image segmentation benchmark (BRATS), IEEE Trans. Med. Imag. 34 (10) (2015) 1993–2024, https://doi.org/10.1109/TMI.2014.2377694.

[100] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J.S. Kirby, et al., Advancing the Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features, Nature Scientific Data 4 (2017) 170117, https://doi.org/10.1038/sdata.2017.117.

[101] S. Bakas, M. Reyes, A. Jakab, S. Bauer, M. Rempfler, A. Crimi, et al., Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge, 2018 arXiv preprint arXiv:1811.02629.

[102] Patrick Bilic, Patrick Ferdinand Christ, Eugene Vorontsov, Grzegorz Chlebus, Hao Chen, Qi Dou, Chi-Wing Fu, et al., The Liver Tumor Segmentation Benchmark (LiTS), 2019 arXiv preprint arXiv:1901.04056, https://competitions.codalab.org/competitions/17094.

[103] Jayaram K. Udupa, Dewey Odhner, Liming Zhao, Yubing Tong, M. Monica, S. Matsumoto, Krzysztof C. Ciesielski, Alexandre X. Falcao, Pavithra Vaideeswaran, Victoria Ciesielski, Babak Saboury, Syedmehrdad Mohammadianrasanani, Sanghun Sin, Raanan Arens, Drew A. Torigian, Body-wide hierarchical fuzzy modeling, recognition, and delineation of anatomy in medical images, Med. Image Anal. 18 (2014) 752–771.

[104] Shixi Qin, Zixun Zhang, Yuncheng Jiang, Shuguang Cui, Shenghui Cheng, Zhen Li, NG-NAS: node growth neural architecture search for 3D medical image segmentation, Comput. Med. Imag. Graph. 180 (2023) 102268.

[105] Jing Wang, Aixi Qu, Qing Wang, Qibin Zhao, Ju Liu, Qiang Wu, Tt-Net, Tensorized Transformer Network for 3D medical image segmentation, Comput. Med. Imag. Graph. 107 (2023) 102234.

[106] Hao Liu, Yiran Sun, Xiangyun Cheng, Dong Jiang, Prior-based 3D U-Net: a model for knee-cartilage segmentation in MRI images, Comput. Graph. 115 (2023) 167–180.

[107] Shu-Ting Li, Ling Zhang, Ping Guo, Hong-yi Pan, Ping-zhen Chen, Hai-fang Xie, Bo-kai Xie, Jiayang Chen, Qing-quan Lai, Yuan-zhe Li, Hong Wu, Yi Wang, Prostate cancer of magnetic resonance imaging automatic segmentation and detection of based on 3D-Mask RCNN, J. Radiat. Res. Appl. Sci. 16 (2023) 100636.

[108] Saidi Guo, Xiujian Liu, Heye Zhang, Qixin Lin, Lei Xu, Changzheng Shi, Zhifan Gao, Antonella Guzzo, Giancarlo Fortino, Causal knowledge fusion for 3D cross-modality cardiac image segmentation, Inf. Fusion 99 (2023) 101864.

[109] Hengxin Liu, Guoqiang Huo, Qiang Li, Xin Guan, Ming-Lang Tseng, Multiscale lightweight 3D segmentation algorithm with attention mechanism: brain tumor image segmentation, Expert Syst. Appl. 214 (2023) 119166.

[110] Muhammad Usman Saeed, Bin Wang, Jinfang Sheng, Ghulam Ali, Aqsa Dastgir, 3D MRU-Net: a novel mobile residual U-Net deep learning model for spine segmentation using computed tomography images, Biomed. Signal Process Control 86 (2023) 105153.

[111] Cao Yuan, Weifeng Zhou, Min Zang, Dianlong An, Feng Yan, Bin Yu, MBANet: a 3D convolutional neural network with multi-branch attention for brain tumor segmentation from MRI images, Biomed. Signal Process Control 80 (2023) 104296.

[112] A.E. Ilesanmi, T. Ilesanmi, O.P. Idowu, et al., Organ segmentation from computed tomography images using the 3D convolutional neural network: a systematic review, Int J Multimed Info Retr 11 (2022) 315–331, https://doi.org/10.1007/s13735-022-00242-9.