

## Review Article

Open Access, Volume 4

# Machine learning-based anatomical segmentation: A systematic review of methodologies and applications

Evan Russell<sup>1,2\*</sup>; Adrian Boyd<sup>1</sup>; Dewar Finlay<sup>1</sup>; Luis Trindade<sup>2</sup>

<sup>1</sup>School of Engineering, University of Ulster, Belfast BT15 1ED, UK.

<sup>2</sup>Machine Learning Department, Axial Medical Printing Limited, Belfast BT2 8HD, UK.

**\*Corresponding Author: Evan Russell**

School of Engineering, University of Ulster, Belfast BT15 1ED, UK.

Email: he.russell@ulster.ac.uk

Received: Apr 25, 2024

Accepted: May 27, 2024

Published: Jun 03, 2024

Archived: [www.jclinmedimages.org](http://www.jclinmedimages.org)

Copyright: © Russell E (2024).

## Abstract

Image segmentation has established itself as a cornerstone of modern medical imaging practices, achieving state-of-the-art accuracy and playing a pivotal role in critical applications such as pathology diagnosis, treatment planning, tumour segmentation and organ delineation to name a few. The integration of deep learning and neural network techniques has sparked a dynamic and constantly evolving landscape, propelling rapid advancements in image segmentation. This work aims to provide an all encompassing review of the general state-of-the-art, with a focus on developments made within the last few years. Over 100 papers presenting development in the field have been considered, primarily published between 2020-2023, separated into their architectural type. Within each type, significant works and conclusions have been highlighted. We conclude the paper with a discussion of the challenges facing the segmentation landscape as a whole, and how these challenges have been best circumnavigated to date.

**Keywords:** Medical segmentation; Machine learning anatomical segmentation medical image processing; Convolutional neural networks.

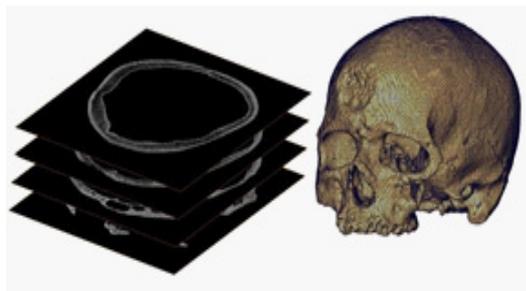
## Introduction

The field of medical imaging has seen great advancements within the last decade. Constant exponential growth in technological capability in both processing power and storage capabilities has led to an increased potential for integration within many work streams, and thus a stronger reliance on said technology to support the ever-growing industrial landscape. In particular, scientific development holds significant potential for substantial advancement in medical imaging. Internal body imaging plays a vital role in the diagnosis and surgical planning processes, and the required medical staff undergo years of extensive training to gain the expertise and insight necessary to interpret the intricacies within medical images. However due to

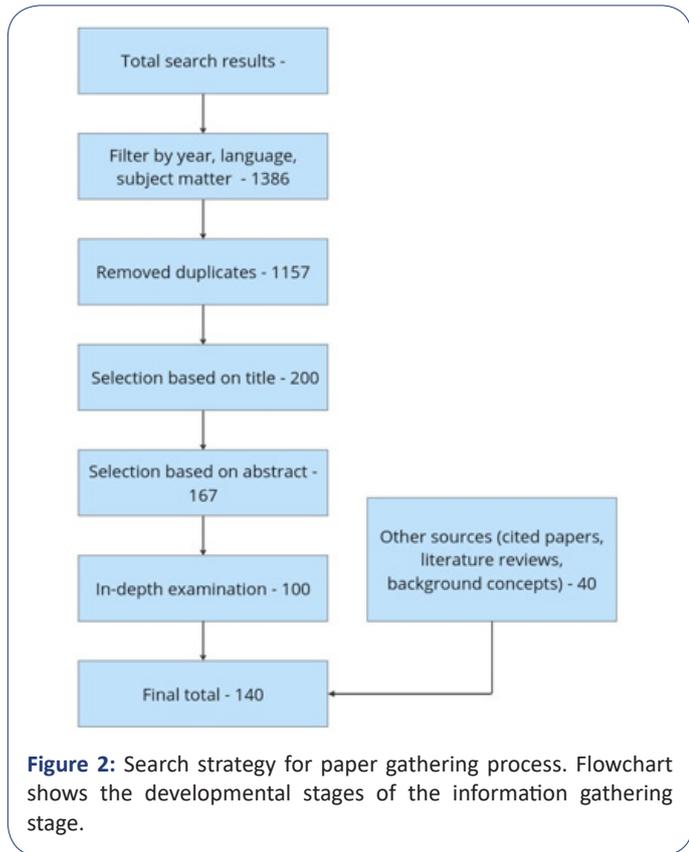
the increasing demand and workload on the required medical professionals, diagnosis time increasingly acts to bottleneck the full imaging process. Therefore any and all advancements that can be made to alleviate this constraint are greatly sought after.

In particular, this applies to the field of medical segmentation, a time-intensive process that requires extensive effort on the part of the trained professional, as well as niche, specific and lengthy training procedures of its own. Segmentation refers to the process of defining specific key structures within the medical image to a pixel level, such as tumours, organs or tissue. This process allows for the 3D modelling of isolated organs or internal bodily systems by the stacking of 2D layers (see Fig-

ure 1), often obtained from pixel-level selection of Computed Tomography (CT) scans, Magnetic Resonance Imaging scans (MRIs) or ultrasound technology. Its application has demonstrated substantial utility in computer-aided healthcare in an ever-increasing number of applications such as patient specific instrument design processes [2], disease diagnosis and monitoring, adaptive radiotherapy and surgical preparation. Given the considerable time that must be dedicated to the segmentation process, it has remained a particularly key area of focus in the field of automation. A range of algorithms have been developed making use of several methodologies within the realm of machine learning and Artificial Intelligence (AI) to perform classification on a pixel level. These advancements aim to streamline and expedite the segmentation process, allowing for efficient and accurate organ definition and anatomical modelling while reducing the burden placed on the surgeon. Fu et al. [3] explores a comparison between from-scratch segmentation compared to manually corrected automatic segmentation. The results revealed a significant time-saving advantage, with the utilization of the semi-automatic system requiring a quarter of the time as fully manual segmentation.



**Figure 1:** Diagrammatical representation of multi-stacking of 2D planes to form 3D voxel-based models [1].



**Figure 2:** Search strategy for paper gathering process. Flowchart shows the developmental stages of the information gathering stage.

This field is witnessing rapid and frequent advancements with the continuous publication of newer developments and increasingly streamlined algorithms. Consequently, it is crucial to stay informed of any recent progress in order to maintain a relevant understanding of the landscape as a whole. This paper aims to cover the general state-of-the-art within the field of deep learning techniques applied to medical image segmentation as of the time of writing, with a particular focus on widely adopted stratagems and algorithms. Any necessary background knowledge within the field is also detailed, and developments of specific importance are identified that may indicate promising future developments.

Figure 2 shows the search process through which the material discussed in this paper was gathered. Initial searches and filtering criteria yielded an excessive number of results thus leading to more robust filtering processes. Upon more in-depth analysis of paper contents, 100 papers presenting novel segmentation processes were identified. An additional 40 papers collected presenting appropriate background information or relevant work.

The rest of this paper is structured as follows: Section 1 covers the landscape of segmentation as a whole, and the dominant topics within the imaging landscape. Section 2 details deep learning processes utilised within the field, including Generative Adversarial Networks (GANs), Convolutional Neural Networks (CNNs) and their derivative - Fully Convolutional Networks (FCNs), as well as specific models of interest. Lastly, section 3 covers the general distribution of models, highlighting popular frameworks and areas under particular focus. We conclude the paper with coverage of relevant technical challenges facing the medical imaging landscape, and discuss potential future developments in the field.

### Segmentation classifications

Segmentation processes can typically be divided into 2 main categories, being semantic segmentation and instance segmentation respectively. The emergence of a third category, panoptic segmentation, has also been reported in recent literature [4].

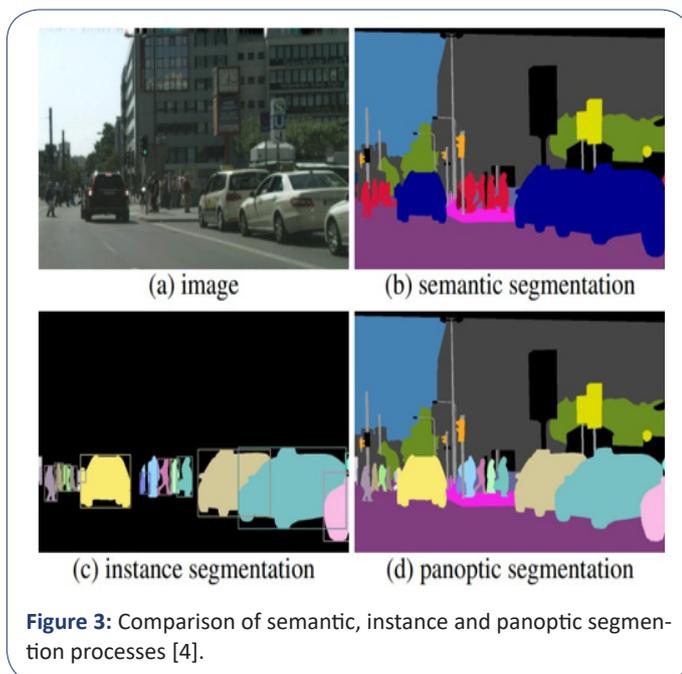
#### Semantic and instance segmentation

Semantic segmentation [5-7] involves the labelling of every individual pixel within a target image with its corresponding semantic class. Any pixel within an object is deemed a part of the same collective, causing no differentiation between two separate instances belonging to the same class. This can cause ambiguity in regions where two objects are overlapping or in close proximity with one another, often “bridging the gap” or wrongly categorising a significant area, as well as difficulty distinguishing individual objects (Figure 3).

This complication is confronted with instance segmentation [8-10], which goes beyond semantic segmentation by introducing a unique independent class for each object within the image. In this case, neighboring cases of the same object will be classified separately, enabling a more detailed understanding of the image landscape as a whole. Heavily occluded objects within the image are still vulnerable to misclassification, and the increased computational demand in some cases limits instance segmentations applicability to real-time segmentation processes and hinders the scalability for large-scale applications.

## Combinative techniques

Semantic instance segmentation [11-14], a hybrid approach, takes on qualities of both segmentation types, of which a handful of notable strategies have emerged over the last few years. This methodology combines the individual instance labelling approach with the class-based object labels employed in semantic segmentation. The result is the generation of pixel-level maps akin to those produced in instance segmentation, but utilising a multi-ID classification system characterising both object type and number. With this change in place, multiple classes of the same object can be classified as more inherently similar to one another than classes representing entirely different objects. Zhou et al. [12] proposed a novel semantic segmentation algorithm that combines aspects of both proposal-based and proposal-free models. They leverage the strengths of both methods by providing a proposal-free model with a Discriminative Deep Supervision (DDS) module, incorporating aspects of proposal-based methodology and introducing instance-sensitive structural information to the system, forming a general instance semantic segmentation algorithm. The promising results of this algorithm demonstrate the effectiveness of combinative semantic instance segmentation and combined proposal-free and proposal-based methods over previous methods utilising a more limited ability set.



**Figure 3:** Comparison of semantic, instance and panoptic segmentation processes [4].

A similar approach was taken by Kirillov et al. [4] who first proposed the concept of panoptic segmentation (Figure 3). Panoptic segmentation adopts a thought process as defined in Adelson et al., “On seeing stuff” [15], which defines two all-encompassing categories within computer vision, that of “things”, distinct, separable objects within the scene that can be recognised, and “stuff”, the amorphous background regions of the image that are not individually identifiable as objects. Panoptic approaches classify individual instances of separate objects with unique labels, as with instance segmentation, but assign the background detail (the stuff) a shared common category label that does not allow separation, as with semantic segmentation approaches. Panoptic segmentation was devised in order to bridge the widening gap between semantic and instance based approaches, and has since been adopted into a range of applications including nucleus segmentation via neural networks [16], lung ultrasound processing [17] and many non-medical industries.

While combinative semantic instance applications are emerging in the current landscape, they have not been widely adopted in the industry and have found limited utilisation in comparison to the long-withstanding and well established semantic and instance segmentation methodologies.

## Dimensionality

Building on the foundation of semantic and instance segmentation methodologies, another pivotal consideration in medical image analysis is the dimensionality of the data being processed. Unlike their 2D counterparts which operate on individual, independent planar slices, 3-dimensional segmentation algorithms [18] leverage the contextual depth information provided by volumetric data to achieve a higher spatial understanding.

Crepes et al. [19] proposed a comparison of performance for models utilising 3D and 2D CNN architectures based on the model deep residual UNet (ResUNet) [20]. For this comparison the 2D data was sampled randomly from a set of CT scans, and the 3D data was taken as a whole-volume. The results showed a higher dice accuracy score for the 2D model but posited that this is likely due to a greater number of samples available when training on slices compared to volumetric data samples. This leads us to a heavily significant limitation in 3D segmentation modelling; the scarcity of annotated volumetric data. While the results indicated a higher dice score for the 2D model, the disparity in sample availability between slice-based and volumetric training data can skew this comparison, and is often the case in 2D and 3D comparative approaches.

Additionally, 3D architectures suffer from heavily increased processing requirements. The need to process volumetric data entirely and simultaneously, significantly increases the processing power necessary to perform such network modelling, often leading to greatly increased training and inference times while also increasing the likelihood of overfitting to the data and requiring more robust regularisation techniques. While present in all aspects of CNN development, the issue of model complexity vs computational efficiency is particularly relevant for 3D models. The results of [19] can be contrasted with the work of Avesta et al. [21] which compares 3D, 2D and 2.5D model architectures trained on 3430 and 60 MRIs separately, across three model styles, capsule networks, UNets, and nnUNets respectively. While the 3D models required up to twenty times as much processing power, the results of all 6 total experiments unanimously agreed that the 3D architectures were most effective, maintaining the highest dice score and often quicker convergence during training which stands in stark contrast to the results of [19]. This indicates that, given sufficient data, 3D architectures maintain their validity as a viable candidate for segmentation processes, in situations where the necessary memory storage and processing power is attainable.

Frequently researchers and practitioners adopt a 2.5D approach as a viable compromise striking a balance between the enhanced accuracy of 3D methodologies without taking on the inevitable computational and memory expenses associated with them. Recent research by Zhang et al. [22] and Ou et al. [23] suggests that for anisotropic volumes with a high discontinuity between slices, direct application of 3D CNNs will yield suboptimal results due to learning of irrelevant features. In contrast, other sources [24,25] suggest that adopting a 2.5D structure is heavily restricting by confining the many possible 2D viewpoints possible from volumetric data to only the 3 cardinal axes, thus

not utilising its full potential. In light of these contrasting viewpoints, it becomes evident that selecting appropriate segmentation dimensionality necessitates careful consideration.

Zhang et al. [26] proposes a series of 2.5D model methodologies; multi-view fusion, fusing 2D/3D features and incorporating inter-slice information. Multi-view fusion involves generating slice-by-slice predictions using 2D CNNs from the 3 cardinal planes and fusing the results through a majority-voting process. Incorporating inter-slice information focuses on leveraging the spatial correlation between consecutive slices by treating them as a time series sequence and using techniques such as Recurrent Neural Networks (RNNs). Lastly, fusing 2D/3D features involves combining features that have been extracted from 2D and 3D CNNs respectively to achieve a higher efficiency while still leveraging spatial information, a structure that would later be incorporated into the well-established 2.5D model design, D-UNet [27]. It was found that all three of the conceived methods surpassed the compared standard 2D network, however, the exact effectiveness of each method is dependent on the segmentation task at hand. Overall it was concluded that 3D CNNs are the best choice for maintaining accuracy alone, however consideration of the time and processing power required is once more entirely subject to the discretion of the reader. The paper highlights the importance of considering the specific characteristics of any medical data in use, and how, due to the sheer number of factors for consideration, choosing the most optimal model is a more nuanced, procedural question with no definitive answer.

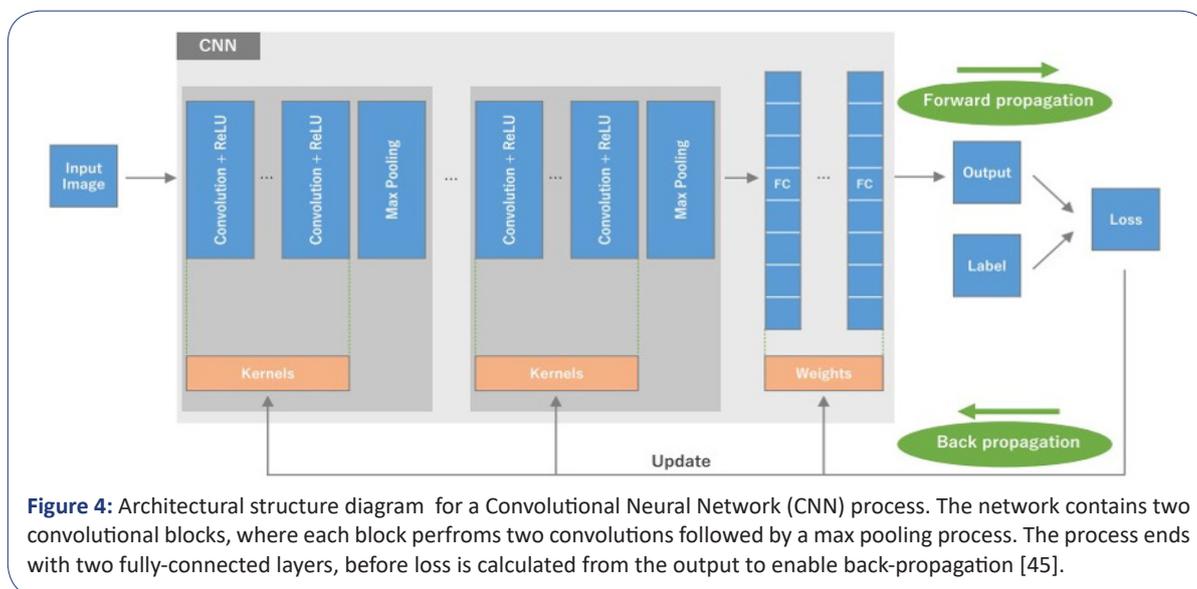
### Segmentation approaches

This section details some overarching approaches to medical segmentation processes. In practice, it is rare for the described methodologies to be utilised in such a discrete, well defined manner. Instead it is commonplace to see combinative

approaches of the following methodologies, often integrating aspects of a range of techniques leading to highly situation-specific algorithms developed to tackle segmentation of a specific organ, region or modality. Nevertheless, the overall topical definitions are described here.

### Atlas-based segmentation

Many framework methodologies have been developed for accurate anatomical segmentation, often yielding reliable results [28,29]. One prominent early developed technique is atlas-based segmentation, which consists of using pre-defined, manually contoured segmentations of similar scans, the titular “atlas”, to act as a reference guide for future segmentation of unseen scans. It is based on the principle idea that key anatomical areas will mostly maintain consistency in shape and size across patients, and makes use of alignment registration algorithms to align the contoured source image with the newly acquired test image. Atlas based segmentation has been developed for automatic computer vision-based segmentation for both singular and multi-organ segmentations, [28] however it is important to acknowledge that this approach is not without limitation. The success of atlas-based segmentation relies heavily on the precision and generality of the atlas contours provided, leading to a decreased reliability in cases where a diverse range of demographics are not considered. Atlas-based segmentation processes have additionally been shown to have decreased viability for application to pathological deviations or trauma-based scans due to an obvious deviation from the typical shape, having a strongly observed tendency to under-segment target images [30]. As a result, they have largely been superseded by more advanced methodologies. Modern approaches have integrated the atlas-based approach with more advanced neural networks to better capture aspects of the organ shape deviating from the norm [31-34] and heighten generalizability to promising results.



**Figure 4:** Architectural structure diagram for a Convolutional Neural Network (CNN) process. The network contains two convolutional blocks, where each block performs two convolutions followed by a max pooling process. The process ends with two fully-connected layers, before loss is calculated from the output to enable back-propagation [45].

### Generative Adversarial Networks (GANs)

While atlas-based segmentation techniques benefit from simplicity, they quickly became overshadowed by more advanced methodologies capable of much more precise segmentation. One such method that has seen such utilisation is Generative Adversarial Networks (GANs) [35,36] which, while typically used for image generation, translation and data augmentation, have seen a recent rise in application to the medical segmentation industry in recent years.

A GAN is built around a two-network system, the discriminator and the generator networks respectively, which are trained together in a competitive style process. The generator network is responsible for mapping a path from a set of simulated noise data to a target distribution. Meanwhile, the discriminator is trained to distinguish between the real data and the simulated data of the generator. The competitive, adversarial training process allows both networks to gradually improve their respective abilities, the intended end point being for the generator to create simulated data that is indistinguishable from actual data.

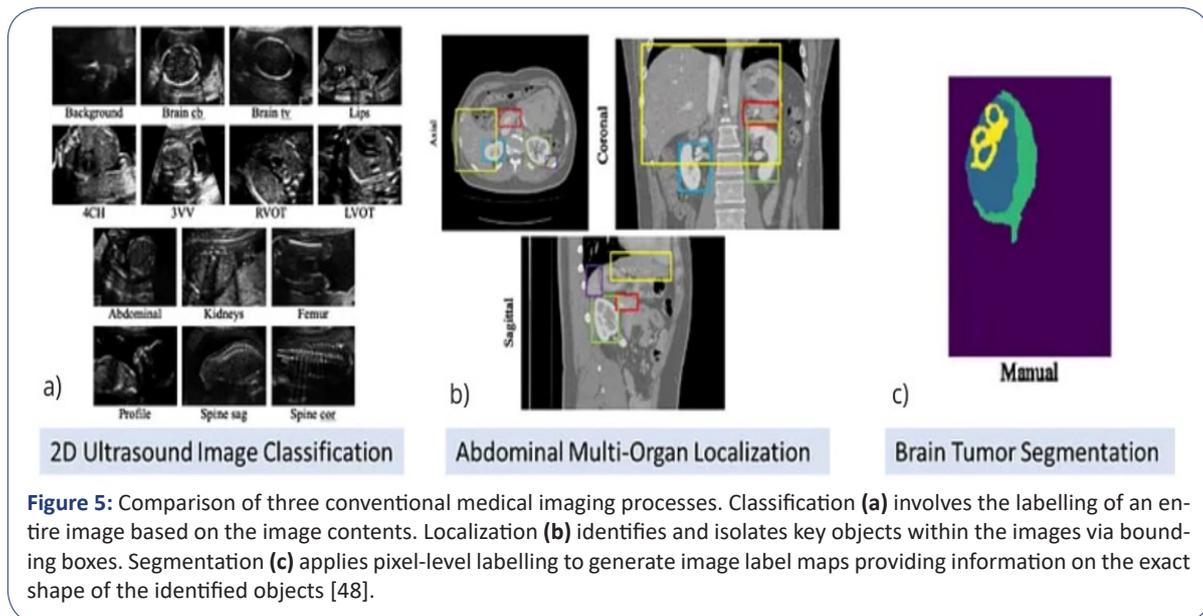
GANs have been applied to the segmentation problem, able to generate pixel-wise segmentation masks that can accurately semantically classify the key region of interest within an acceptable error [37,38].

As a whole, the community has adopted a semi-supervised approach to GAN-based segmentation. Li et al. [39] recently proposed a GAN-based system to perform segmentation on a set of hippocampus data, where two models were contrasted, adopting fully-supervised and semi-supervised learning models respectively across four experiments. It was concluded that the semi-supervised GAN approach trained on unlabeled data maintained a consistent improved segmentation accuracy over the compared fully-supervised model, having an average accuracy increase of 0.4%. It is suggested that semi-supervision

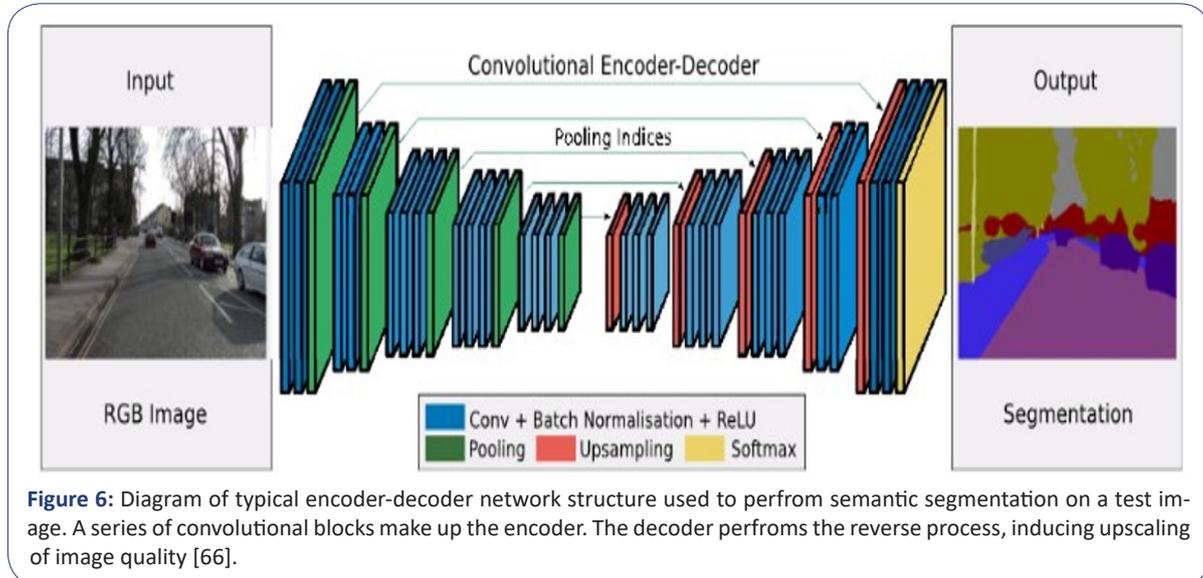
makes more efficient use of a limited dataset, hence its high applicability within the specific application. With that said, avenues of unsupervised models, particularly self-supervised models [40-42], are still regularly explored for specific applications.

### Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) first proposed by [43] have emerged as the prevailing methodology for medical image segmentation, gaining widespread adoption and recognition within the field. Similar to GANs, CNNs leverage deep learning to facilitate accurate and efficient segmentation, however CNNs are built around the concept of feature extraction. Features of the provided training images are automatically identified and learned by the network through application of convolutional operations.



**Figure 5:** Comparison of three conventional medical imaging processes. Classification (a) involves the labelling of an entire image based on the image contents. Localization (b) identifies and isolates key objects within the images via bounding boxes. Segmentation (c) applies pixel-level labelling to generate image label maps providing information on the exact shape of the identified objects [48].



**Figure 6:** Diagram of typical encoder-decoder network structure used to perform semantic segmentation on a test image. A series of convolutional blocks make up the encoder. The decoder performs the reverse process, inducing upscaling of image quality [66].

This allows the CNN to capture local spatial feature information and exploit spatial hierarchies, enabling the discernment of relevant patterns within the provided images. Convolution is enabled by passing a kernel, typically a 3x3 matrix across the image in a sliding-window fashion, performing matrix multiplication at each step to generate a feature map, where higher pixel values correspond to the presence of that feature. By convolving the image with multiple kernels, a range of features are extracted. This process replaces the fully connected layers present in typical neural networks, a necessary adaptation

to prevent impractical processing time when handling image-based data [44-47].

Figure 4 [45] shows a diagrammatic representation of a typical CNN architecture. After a series of convolution processes, max pooling is performed to reduce spatial resolution and extract the most salient features. The process typically ends with one to two fully connected layers to integrate the extracted features and perform high-level segmentation. Back propagation allows the CNN to learn and refine the weights and biases of the network by iteratively propagating the error from the output

layers back through the network, adjusting the values to better minimise the loss function.

### Architectural structure

CNNs have seen substantial use, as well as countless deviations adapted for specific use cases within the field. This section covers notable structural adaptations of CNNs aimed at optimising their suitability for the task of image segmentation.

#### Region-based CNNs

The natural development of the CNN leads us to Region-based CNNs (R-CNNs), an application of CNNs to object detection (Figure 5). In the seminal paper, Girshick et al. [48] proposed it as a novel approach to leverage the versatility of CNNs in a more region based framework that showed a 30% improvement over the standard object detection systems at time of publication.

R-CNNs introduce a region proposal stage, which isolates objects of interest via bounding boxes, followed by an object detection system making use of conventional CNNs. Utilising a region proposal stage efficiently reduces the search space, minimising the computational burden placed on the system. Each individual region can then be analysed by a separate CNN process. R-CNNs built upon the success of CNN-based object detection systems with a generally reduced processing time, increased robustness when handling differently sized objects, better handling of partially obscured regions and an improved receptive field. Processing time was reduced with an improved accuracy by Girshick et al. [49] by introducing Region of Interest (ROI) pooling layers for more efficient feature extraction and sharing of convolutional features across multiple region proposals. This was further improved by Ren et al. [50] who integrated a Region Proposal Network (RPN) capable of generating region proposals without the need for external algorithms.

Pixel-wise instance segmentation ability for each bounding box region was later incorporated into the R-CNN process by He et al. [51]. Building upon the foundation of faster R-CNNs, mask R-CNNs introduce a binary mask channel output alongside the pre-established class and box-offset prediction outputs. The parallelization of the channels simplifies the process compared to the standards at the time which used prediction for classification [52], while preserving the advantages of faster R-CNN. The paper [51] presents a comparison of methods showing that mask R-CNN outperforms all state-of-the-art approaches at time of publication. Owing to its consistently reliable segmentation accuracy and preferable processing time, R-CNN has maintained relevance with minimal changes in design since its conceptualization, seeing many recent applications in literature, with applications to tumour detection [53,54], lung segmentation for covid-oriented [55] and general purpose use cases [56,57], and blood vessel segmentation [58] to name a few. Studies such as that by Felfeliyan et al. [59] and Dandil et al. [55] confirm the validity of this approach with evaluation metrics rivalling other established methods. In comparison, a study by Dogan et al. [60] proposed a hybrid solution utilising Mask R-CNNs localization ability integrated into a novel UNet architecture for segmentation. Compared against 16 state-of-the-art models, the proposed system unanimously achieved higher segmentation accuracy on pancreatic data, once again suggesting hybrid methodologies as a promising approach for improving segmentation performance beyond what standalone algorithms can perform.

### Fully Connected Networks (FCNs)

While CNNs have seen extensive use within medical image segmentation tasks and have demonstrated remarkable performance and versatility, their inherent limitation lies in their inability to capture fine-grained details caused by the down-sampling processes. As a result, alternatives have been developed to combat these issues. Fully Convolutional Networks (FCNs) [61,5] have recently emerged as a specialised architecture designed explicitly for development of high-accuracy segmentation on the pixel level. FCNs replace the late-stage fully-connected layers present within a conventional CNN with additional convolutional layers, preserving a higher level of spatial information throughout the network. In order to address the reduction in resolution caused by the convolutional and max pooling layers, transposed convolutions or upsampling techniques are often utilised to restore the output to match the resolution of the input image, a necessity for segmentation mapping. This also allows for efficient mapping of both local and global contextual information, heightening the segmentation ability further.

Deep supervision is often deployed during the training process of FCNs [62,63]. Building upon the idea of a loss function quantifying performance at the end of a network, deep supervision involves the integration of additional classifiers at regular intervals within the network architecture, allowing for better propagation of gradients and thus more efficient, effective learning.

FCNs quickly became the gold standard within medical image segmentation with the development of UNet [64] a key model architecture that made use of skip connections within an encoder-decoder setup. UNet has received unprecedented attention within the field of medical segmentation, and has served as a catalyst for the development of numerous derivative models. The remarkable impact of UNet and its influence necessitates a dedicated exploration in section 2.2.4.

FCNs have since seen their own wave of development with adaptations specialised for key niche purposes being published regularly. Region-based FCNs, inspired by Mask.

#### Encoder-decoder networks

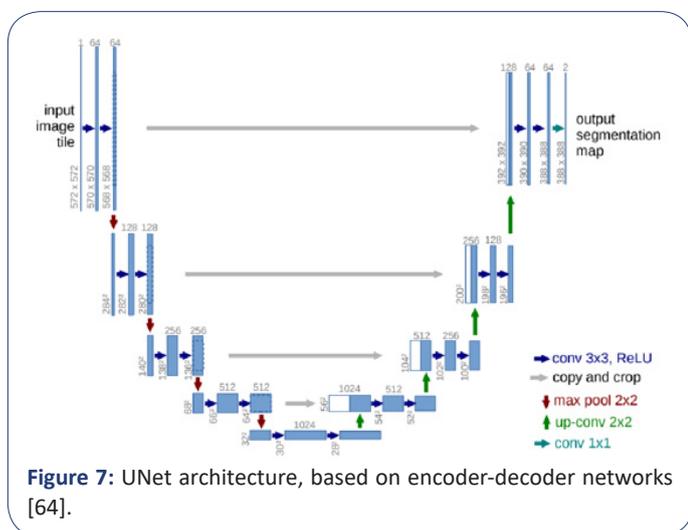
Within the field of FCNs, a multitude of specialised architectures have surfaced, signifying notable advancements for the greater medical segmentation landscape. Encoder-decoder networks [65] are one such advancement, adopting a specific convolutional structure. Figure 6 [66] shows the full encoder-decoder architecture with each individual layer's function labelled. The lack of any non-convolutional layers defines the encoder-decoder architecture as an FCN.

These networks are distinguished by their division into two distinctive components; namely the encoder and decoder stages. During the encoder stage, image dimensionality is progressively reduced through a series of convolution and pooling layers, allowing the capture of hierarchical features at different levels of abstraction. These features therefore capture both global and local information, enabling the network to learn rich representations of the input image. Typically the decrease in resolution is met with an increase in the number of feature maps generated at that layer, motivated by the need to capture increasingly abstract features of the image.

The bottleneck layer consequently acts as a bridge between the encoder and decoder sections. The purpose of the decoder is to transform this representation back into a full-resolution output matching the original input domain. The decoder operates in a reverse-manner to counteract the encoder, performing a series of up sampling deconvolutions via the use of transposed convolutional layers. This gradually increases the spatial resolution, recovering the spatial details lost during the down sampling process. The decoder outputs a dense, pixel-level segmentation map having learned inherent spatial and contextual features from the up sampling/down sampling process, and combining them with higher level semantic information from the encoder. This allows the network to incorporate global context and semantic understanding into the segmentation map. Encoder-decoder based networks have shown superior performance compared to traditional CNN architectures, particularly during segmentation of smaller objects with noisy boundaries [67].

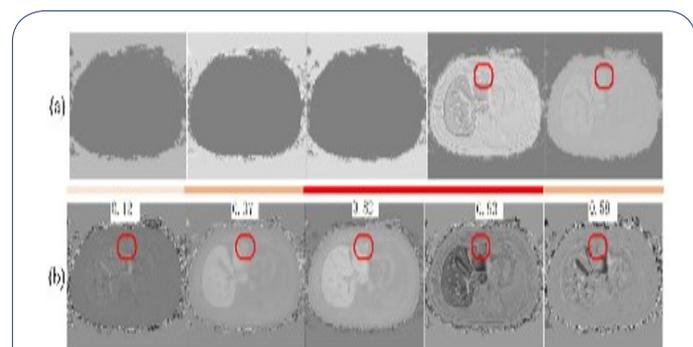
### UNet

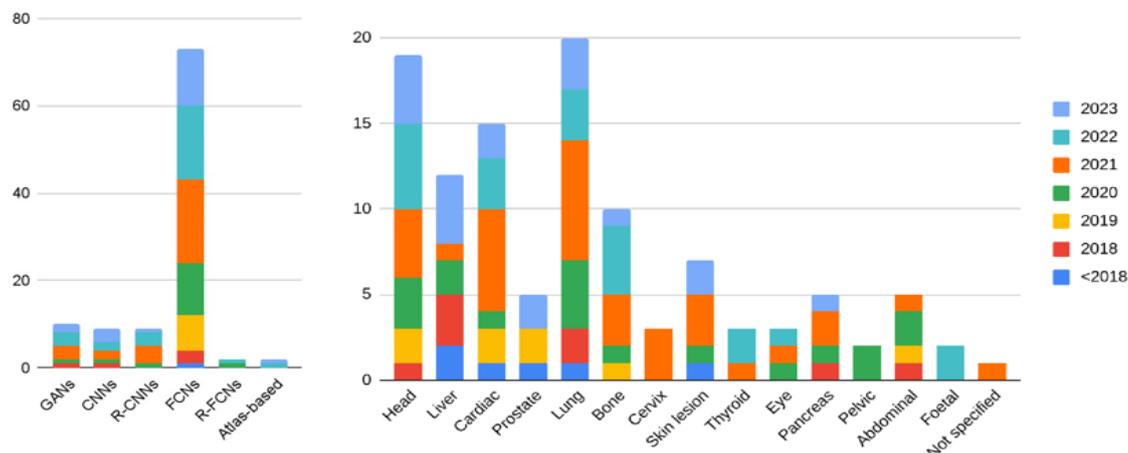
Encoder-decoder networks have maintained relevance within the genre of specialized medical segmentation, thanks in large part to the emergence of the UNet algorithm [64]. UNet has seen significant attention and has been extensively researched within the medical imaging community owing to its exceptional performance in a range of applications including tumour segmentation, lesion detection and single or multi-organ processing, as well as its remarkable ability to successfully segment smaller objects [68]. Figure 7 shows the typical UNet architecture as presented in [64]. The architecture makes use of the encoder-decoder setup as described in 2.2.3, where the encoder extracts hierarchical features from the input image, and the decoder gradually up samples the features to reconstruct a full-resolution output map. UNets key innovation to the field is the integration of skip-connections between corresponding layers in the encoder and decoder branches. These skip-connections enable the transfer of information between the encoder and decoder at multiple levels of resolution and abstraction. By introducing this method of information transfer, the UNet architecture addresses the issue of data loss during the down sampling section, and facilitates the integration of fine-grained details from the encoder with high-level semantic representations from the decoder.



In a recent study by Baccouch et al. [67] the performance of various segmentation methods including UNet, Faster R-CNNs, CNN and 3D-FCNs were evaluated for the automatic segmentation of cardiac MRI scans. Dice score was utilised as the evaluation metric. The results clearly showed that models utilising UNet architecture, be it entirely or in conjunction with other methods, achieve a significantly higher dice score than alternate methodologies, with UNet based methods achieving dice scores of ~94 and non-Unet methods reaching ~84 on average. The worst performing UNet method achieves an equal dice score to the best performing alternative method. UNet models additionally showed more promising results for the additional metrics, Hausdorff distance and accuracy, achieving consistently higher results than other models. Overall, this paper confirms the validity of UNet and UNet hybrid approaches as suitable inter-organ segmentation techniques. The paper posits that this significant increase in ability may be due to the circumventing of an inherent limitation in conventional CNNs, the restriction on input image size as detailed in [69].

UNet has established continued relevance in the field of medical imaging thanks in large part to its inherent flexibility and ease of customization of architecture and pre-processing. UNet's architecture allows for easy modifications and integration of domain specific knowledge to address unique novel challenges such as identification of distinct pathologies or segmentation of non-conventional organ shapes. In recent years, the development of UNet-based adaptations has shown no signs of deceleration. UNet++ [70,71] has shown itself as an adaptation of particular notoriety, introducing nested skip pathways building upon UNets established skip connections. Nested pathways enable multiple connections to be made across resolutions between different tiers of the encoder and decoder, thus allowing features to be captured at multiple scales more effectively, leading to an improved performance in capturing of fine details. The hierarchical nature of nested skip pathways additionally facilitate better feature reuse and gradient flow through the network, enabling more efficient training processes and quicker convergence. In its seminal paper, Zhou et al. [70] demonstrated the increased accuracy of UNet++ with deep supervision over UNet and wide UNet across 4 different medical imaging datasets. Results showed an average increase in Intersection over Union (IoU) of 3.4 and 3.9 respectively. A later study also by Zhou et al. [71] provided further comparisons of effectiveness for both semantic and instance segmentation cases, using both 2D and 3D architectures. For semantic segmentation, UNet++ was compared against standard and wide UNet, with UNet++ being seen to consistently outperform them. For instance segmentation, Mask R-CNNs were compared against a separate version utilising nested skip pathways, dubbed "Mask





**Figure 9:** Published works summarized by year of publication and (a) – Segmentation methodology. (b) – Primary organ(s) used for segmentation.

R-CNN++”, and once again results unanimously showed higher performance on all evaluation metrics. Results suggest that inclusion of nested skip pathways over the original skip connections leads to consistently better feature learning and segmenting ability.

UNet++ has in turn seen its own wave of developmental offshoots. A recent paper by Zhang et al. [72] discusses the application of UNet++ to small-scale precise target segmentation, and proposes an altered algorithm. SUNet++ seeks to increase segmenting ability for fine structures through inclusion of 3-dimensional context, and an attention mechanism [73,74] to better focus on identified key areas of interest. The model was compared against a series of other UNet derivatives including UNet++ UNet3+ and nnUNet. Results derived from segmentation of fine structures within liver scans show SUNet consistently achieving the highest Positive Predictive Value (PPV) compared to other models, however similar metrics such as the sensitivity and Hausdorff distance remain similar across all models. Nevertheless, a higher PPV suggests more confidence in positive predictions, a key requirement particularly for segmentation of fine smaller structures. Figure 8 [72] shows the predictive segmentations from UNet++ (a) vs SUNet++ (b) for 5 similarly placed slices. As can be seen, no structure was detected by the UNet++ model in 3 of the slices, whereas the additional 3D contextual consideration allows SUNet++ to better segment these structures.

While many of the papers discussed within this review claim to achieve previously unprecedented accuracy across evaluation metrics, it is important to consider the environment in which they are developed. Medical image segmentation demands specialised architectures and training schemes to tackle the intricacies of different anatomical features. However this customization often leads to models that are heavily overfitted to the specific problem, limiting the generalizability and applicability to broader scenarios. This is an issue that has become increasingly relevant in modern publications where a myriad of models are presented as superior solutions without adequate validation on diverse datasets, or fair comparison to other state-of-the-art methodologies. Consequently a challenge arises in identifying segmentation algorithms that genuinely excel in generalised cases, and not solely in the limited window of scenarios they are presented with. This establishes the need for more robust, generalised testing frameworks. Recognising the need for more robust frameworks, Isensee et al. [75] developed NNUNet, a framework for UNet based medical segmentation

designed for self-adaption. NNUNet goes beyond traditional methodologies by incorporating self-adaption mechanisms which allow the framework to dynamically adjust the architecture and hyperparameters of the model it is applied to based on the unique input data it is provided with. This adaptive nature allows NNUNet to accurately handle variations in input resolution, geometry and distribution on a procedural, case-by-case basis without the need for manual fine tuning. To name a few, NNUNet allows automatic modification of the model architecture, data augmentation, resolution, and cropping.

To evaluate the ability of NNUNet, experiments were conducted in the context of the Medical Segmentation Decathlon Challenge (MSDC) which provided a set of 10 distinct and diverse datasets encompassing various medical disciplines, image modalities and image sizes. It was found that as of the paper’s publication, models equipped with NNUNet outperformed all other models in the online leaderboard challenge achieving the highest dice score metric in all cases with one class exception. These results demonstrate the outstanding performance and versatility provided by NNUNet in handling complex segmentation in diverse datasets without prior pre-adjustment. The success of NNUNet highlights its potential as a leading framework for medical segmentation tasks, offering a promising solution for overcoming overfitting and heightening a model’s generalizability.

Since its conception NNUNet has been largely accepted into many model frameworks, and its ability scrutinised in a wider range of scenarios. Recently models utilising NNUNet as its base saw 4th and 5th places in the HECKTOR 2022 head and neck segmentation challenge [76] and the FLARE 2022 abdominal organ segmentation challenge [77] respectively.

NNUNet has since seen large-scale adoption within the imaging field thanks to its extraordinary automatic adaptability to a range of use cases, with varying applications in the fields of breast tumour MRI scans [78], foetal ultrasound scans [79], lung tumour diagnosis [80], and brain tumour segmentation [81] to name a few.

### Recent publications

So far we have discussed the significant overarching methodologies and processes used within the medical segmentation field. We will now cover the general landscape of publications over the previous few years.

In total, 100 papers presenting newfound segmentation methodologies have been considered for the analysis, and have been categorised by their overarching method and primary target organ. Figure 9a displays the tallied results for each of these categorization means. FCNs have firmly established themselves as the popular means of segmentation, thanks in large part to the versatility and ease of modification within the UNet architecture.

An increase in publications exploring FCN architectures can be seen year-by-year. The recent emergence of automatic parameter adjustment methodologies such as NNUNet to correct the theorised overfitting issues within personalised UNet models will likely lead to continued and further adoption of these architectures. In contrast, despite recent publications showing strong potential for region-based integration with FCN architectures [12,82], the technique has not been largely adopted within the field, seeing minimal recent attention.

While GAN based segmentation methods have seen noteworthy publication in recent years, it is important to consider that this has been met with similar publication numbers in related fields, such as object detection and diagnosis [83,84], image generation [85,86,87] and data augmentation [88,89] to name a few.

Author	Year	Modality	Anatomy	Model type
Liu et al. [96]			Brain	Multi-task disentanglement framework
Vijendran et al. [97]	2023	Multimodal	Brain	CNN with adaptive Firefly optimization
Xu et al. [98]	2023	Multimodal	Lung, skin lesion	Dual-stream rep. fusion learning
Liu et al. [99]	2022	Multimodal	Brain, cellular	Lightweight dual-domain network
Li et al. [100]	2021	CT scan	Blood vessels	Triple attention network
Gu et al. [101]	2021	MRI	Skin lesion & foetal	Comprehensive attention-based CNN
Zhou et al. [102]	2020	CT scan	Multi-organ	CNN with prior awareness

### CNNs

GANs across various domains is presently the subject of comprehensive investigation. As research endeavours continue to explore the multifaceted applications of GANs, their adoption into a diverse scope of fields holds the promise of wide-ranging transformative impact.

Figure 9b shows the organ distribution across the examined segmentation papers. For the pelvic and abdominal cases, umbrella terms were used as the comprehensive visualisation of the entire area, encompassing both tissue and organs, necessitated the segmentation of a diverse array of structures within the region.

More complex organs such as the pancreas, whose difficulty in segmentation comes from its smaller size and close proximity to other abdominal organs, are yet to see the same attention, however algorithms specifically designed for pancreatic use have emerged. Attention mechanisms, which first define a key area of focus within the image, have been shown to be effective in pancreatic segmentation, largely circumventing the issue presented by abdominal organ proximity [73,90,91]. Indeed, accurate segmentation of the pancreas in particular has proved to be one of the more challenging tasks within the current state-

of-the-art.

Lung segmentation has seen a significant rise in recent years. Influenced by the recent COVID-19 pandemic, investigations into segmentation and automatic diagnosis of lung pathologies have surged [40,55,92-95] to meet the need for better diagnostic methodologies. As anticipated, FCNs emerge as the prevailing methodology within this field demonstrating their widespread adoption, however GAN-based approaches maintain a noteworthy presence in the literature, likely attributed to their utilisation in tackling classification challenges.

Author	Year	Modality	Anatomy	Model type
Fu et al. [3]	2018	MRI	Multi-organ	CNN with correction network

### Technical challenges

While great advancements have indeed been made in the field of medical imaging, many long-standing sources of error within the field have not yet received adequate solutions. This section will cover the main problem areas for medical segmentation, and the processes used to combat these issues.

### The data scarcity problem

Data scarcity can place a significant restriction on the ability of medical segmentation models, heavily impeding their development and performance. This scarcity arises from a range of factors, including limited access to data with appropriate formatting and quality, appropriate screening subjects, and strict rules regarding medical data access and privacy. Additionally, manual annotation of scans is a laborious and time-consuming process requiring high level knowledge and training from medical professionals, further limiting any possible data output. The insufficiency in comprehensive, diverse datasets heavily hampers the training processes of otherwise robust models, leading to reduced accuracy and often enabling overfitting to the data provided. This issue holds particular importance in niche segmentation cases, where access to appropriate scans of specific rare conditions or pathologies is exceptionally limited. This scarcity and limited access to appropriate data has led to a range of work-around solutions. Namely, data augmentation, establishment of collaborative, open source data sharing platforms, and integration of transfer learning strategies.

Data augmentation operations [85,161,162,89] allow for the generation of additional training data via the application of transformation processes to the existing dataset. By applying transformations such as rotations, scaling, contrast adjustments, mirroring or random cropping, distinct training images can be generated that differ enough from the original to provide meaningful additional context and variation.

These processes can also increase the generalizability of a model, better simulating the deviations in imaging conditions and contrast quality present in real images produced from different scanner setups. While augmented data is not wholly new and therefore not as valuable as datasets entirely independent from existing samples, the techniques presented have shown noteworthy and consistent improvements in model robusticity, versatility and accuracy.

Transfer learning [163] has emerged as a powerful technique within the field, providing a feasible solution capable of circumventing the limitations of insufficient data as well as computational resources. Transfer learning relies on neural networks'

Author	Year	Modality	Anatomy	Model type
Cai et al. [103]				3D UNet-based
Zhang et al. [72]				3D UNet-based
Elghazy et al. [104]	2023	MRI	Brain	Dual/Triple stream UNet+ResUNet
Kumari et al. [105]	2023	MRI	Liver	Dual stream ResUNet
Asnawi et al. [106]	2023	Multimodal	Liver	Multiple 3D UNet-based models
Shangzhu et al. [95]	2023	MRI	Brain	Multi-scale UNet adaptation (MSUNet)
Li et al. [107]	2023	CT scan	Lung	UNet-based multi-feature association
Saha et al. [93]	2023	Multimodal	Multi-organ	Attention-based dense UNet
Staybe et al. [108]	2023	Multimodal	Multi-organ	UNet-based hierarchical CNN stack
Hassanzadeh et al. [109]	2023	CT scan	Chest	Ensemble Evolutionary UNet-based
Dang et al. [110]	2023	CT scan	Head	2D UNet-based
Rebaud et al. [76]	2023	MRI	Multi-organ	NNUNet variations
Sun et al. [111]	2022	MRI	Cardiac	2-stage cascaded UNet-based
Abdo et al. [112]	2022	PET/CT	Cardiac	Attention-UNet-based
Hsu et al. [113]	2022	CT scan	Head & neck	2D, 2.5D, 3D and 3.5D UNet variations
Peng et al. [114]	2022	Cone-beam CT scan	Chest/bone	Improved UNet with attention mechanism
Wang et al. [115]	2022	Cone-beam CT scan	Jaw/teeth	Improved UNet with attention mechanism
Wang et al. [115]	2022	CT scan	Jaw/teeth	2D and 3D DeepResUNet
Singh et al. [116]	2022	CT scan	Brain	3D CNN (DeepMedic)
Inkeaw et al. [117]	2022	MRI	Jaw/teeth	Cascaded NNUNet
Alqaoud et al. [78]	2022	CT scan	Cardiac	NNUNet and NNFormer-based
Lee et al. [80]	2022	MRI	Brain	Dual-branch UNet adaptation
Jiang et al. [118]	2022	CT scan	Breast	2D dual-decoder network
Peng et al. [119]	2022	Multimodal	Thoracic	NNUNet, ResNet, VGG models
Schachner et al. [79]	2022	Ultrasound	Multi-organ	Attention-based recalibration network
Guan et al. [120]	2022	Ultrasound	Foetal	UNet with concatenation block
Diniz et al. [121]	2022	MRI	Foetal Torso	Asymmetrical encoder-decoder network
Zhou et al. [12]	2022	CT scan	Brain	Proposal-free FCN
Zhou et al. [122]	2022	Ultrasound	Cardiac	Unified context-refinement network
Sun et al. [123]	2022	CT scan	Pelvic & cellular	V-Net based co-learning network
Xue et al. [124]	2021	Ultrasound	Foetal	UNet and Otsu-based
Li et al. [9]	2021	PET/CT	Eye	2D UNet-based
Alam et al. [125]	2021	CT scan	Wrist	2D UNet-based
Zhang et al. [126]	2021	X-Ray	Bone	FCN with atrous convolution
Xie et al. [127]	2021	CT scan	Lung	3D UNet-based
Zhang et al. [128]	2021	CT scan	Brain	UNet with spatial attention module
Fu et al. [129]	2021	CT scan	Cardiac	Ensemble FCN network
Enshaei et al. [130]	2021	PET/CT	Abdominal	3D UNet-based
Almeida et al. [131]	2021	CT scan	Lung	UNet-based
Luca et al. [132]	2021	CT and MRI	Chest	3D UNet-based
Zopes et al. [133]	2021	Multimodal	Chest	Multi-encoder-decoder network
Mahmud et al. [94]	2021	Multimodal	Cervix	ResUNet-based
Mohammadi et al. [134]	2021	CT scan	Brain	DUNet-based
Zhou et al. [27]	2021	CT scan	Lung	2D UNet-based
Asipong et al. [135]	2021	MRI	Multi-organ	SegNet-based
Xu et al. [136]	2021	CT scan	Brain	Cascaded multi-encoder-decoder network
Xu et al. [136]	2021	PET/CT	Lung Neck	Cascaded 2.5D UNet
Khan et al. [137]	2021	Multimodal	Multi-organ	Cascaded CNN
Zhang et al. [22]	2021	MRI	Lung	FCN with multiscale mixed attention
Ranjbarzadeh et al. [92]	2020	CT scan	Pancreas	3D VNet-based
Zhang et al. [90]	2020	CT scan	Pancreas	UNet++
Li et al. [138]	2020	PET/CT	Lung	2.5D cascaded CNN
Zhou et al. [71]	2020	Multimodal	Whole body	Adaptive fully-dense UNet
Jemaa et al. [139]	2020	PET/CT CT scan	Multi-organ	
Wang et al. [140]	2020	PET/CT CT scan	Multi-organ	
Author	Year	Modality	Anatomy	Model type
Liu et al. [141]	2020	CT scan	Liver	Multi-channel Fusion Net
Wang et al. [142]	2020	Multimodal	Skin lesions	RESNet based with boundary awareness
Zhang et al. [143]	2020	Multimodal	Skin lesions	Dense UNet-based
Jha et al. [144]	2020	Multimodal	Multi-organ	Double UNet architecture
Kim et al. [145]	2020	CT scan	Abdominal	3D patch-based CNN
Zhang et al. [82]	2020	CT scan	Multi-organ	Cascaded V-Net
Moradi et al. [146]	2020	Echocardiogram	Multi-organ	Multi-feature pyramid UNet
Duan et al. [147]	2019	MRI	Cardiac	Segmentation & Landmarking (SSLLN)
Lei et al. [62]	2019	Ultrasound	Cardiac	Deeply supervised VNet
Chan et al. [148]	2019	CT scan	Head and neck	Multi-task learning CNN
Wang et al. [63]	2019	MRI	Prostate	3D Dilated FCN
Weng et al. [149]	2019	Multimodal	Prostate	NAS-UNet
Van Harten et al. [150]	2019	CT scan	Multi-organ	2.5D UNet
Zhang et al. [151]	2019	DCE-MRI	Breast	Mask-guided FCN
Chlebus et al. [61]	2019	CT scan	Breast	2D UNet-based
Isensee et al. [75]	2018	CT & MRI	Liver	NNUNet
Oktay et al. [73]	2018	CT scan	Pancreas	Improved UNet with attention mechanism

## R-FCNs hello

Author	Year	Modality	Anatomy	Model type
Ma et al. [81]	2022	MRI	Brain	Region-based NNUNet
Liang et al. [152]	2020	CT scan	Head & neck	Multi-view fine-grain ROI-based CNN

## GANs

Author	Year	Modality	Anatomy	Model type
Chaturvedi et al. [42]	2023	High-res image	N/A	Cut-and-Paste trained GAN
Raja et al. [153]	2023	MRI	Brain	GAN with K-means clustering & Mobilenet
Li et al. [39]	2022	MRI	Brain	Semi-supervised GAN
Shabani et al. [40]	2022	CT scan	Lung	3D GAN mask generator with 2D UNet
Tyagi et al. [154]	2022	CT scan	Lung	GAN generator with UNet
Wei et al. [155]	2021	CT scan	Liver	GAN with mask RCNN enhancement
Li et al. [91]	2021	CT scan	Pancreas	GAN with attention mechanism
Xu et al. [156]	2021	CT/CTA	Heart	Cycle consistent GAN
Murugesan et al. [157]	2020	MRI	Prostate & heart	SEG-GLGAN with UNet implementation
Costa et al. [158]	2018	Retinal image	Retinal blood vessels	Adversarial Autoencoder-GAN

## R-CNNs

Author	Year	Modality	Anatomy	Model type
Tu et al. [159]	2023	CT scan	Liver	Slice fusion Mask R-CNN
Dandil et al. [55]	2022	CT scan	Lung	Mask R-CNN based
Zhang et al. [53]	2022	MRI	Breast	Parallel region proposal Mask R-CNN
Wei et al. [155]	2022	CT scan	Liver	GAN Mask R-CNN
Dandil et al. [57]	2021	CT scan	Lung	2D Mask R-CNN based
Dogan et al. [60]	2021	CT scan	Pancreas	2.5D Mask R-CNN with UNet
Author	Year	Modality	Anatomy	Model type
Cai et al. [160]	2021	Standard digital image	Miscellaneous	Cascaded R-CNN
UlHaq et al. [54]	2021	CT scan	Liver	ResNet-based Mask R-CNN
Hu et al. [56]	2020	CT scan	Lung	CNN Mask R-CNN

## Atlas-based

Author	Year	Modality	Anatomy	Model type
Xie et al. [33]	2023	MRI/CT	Brain	Multi-Atlas segmentation with DCNN
DIng et al. [34]	2022	MRI/CT	Heart & Liver	BiRegNet atlas-to-target registration

ability to capture generic, transferable features, leveraging the experience gained from pre-trained models and transferring that knowledge to related specific tasks. This in turn reduces the training time needed as well as increasing generalisation for models with limited data. Transfer learning has displayed remarkable success across image classification, object detection and segmentation, and development of new methods of transfer is a prominent field of study. Apostolopoulos et al. [164] proposed the advantage of transfer learning for fast detection of COVID-19, allowing for fast development of capable detection algorithms. While promising results were gleamed, it was proposed that suboptimal amounts of data remained the limiting factor at that time. This work was built upon by Das et al. [165] who proposed the alternate use of x-ray scans over CT due to higher availability and prevalence of X-ray machines over CT scanners. Using transfer learning with a series of CNNs, a classification model was developed that is capable of correctly identifying COVID-19 patients with an accuracy of 99.96% from a dataset of combined pneumonia and healthy patient scans, and 99.92% from tuberculosis, pneumonia and healthy scans respectively. Transfer learning allowed for suitable, successful algorithms to be developed in a timely manner, a vital requirement during the early stages of the COVID-19 pandemic.

GANs have been applied to the data scarcity problem, applying their inherent generative ability and adversarial, combative learning approach to generation of artificial datasets indistinguishable from real data. Mann et al. [166] recently demonstrated the ability of GANs for this application, utilising a deep

convolutional GAN (DCGAN) setup for generation of COVID-19 positive chest scans. Generated image accuracy was measured by testing the images on a classifier algorithm provided by the COVID CT repository, and an accuracy of 40% was seen, showing potential for the effectiveness of procedural data generation methods.

Rejusha et al. [85] recently presented a comparison of the effectiveness of GAN-generated data vs conventional augmentation practices applied to Alzheimer's brain scans. Accuracy held by the DCGAN reached as high as 83%, greatly surpassing the traditional augmentation and baseline Resnet accuracies of 63%. This highlights the potential for GAN-based generation to address and potentially mitigate the data scarcity problem within segmentation practices.

### Class imbalance

Class imbalance poses a significant challenge to segmentation algorithms where the training datasets have a significant difference in positive to negative cases. In the case of medical segmentation, class imbalance presents itself as images with a small number of relevant labels, and a majority background. This is extremely common in many avenues of medical imaging, including tumour detection, pathology identification and organ segmentation. The consequence of class imbalance is biasing towards the negative or background class, leading to suboptimal performance in segmentation of the crucial areas. Most commonly, this issue is tackled by the deployment of specialised weighting functions such as weighted cross entropy

loss, which place a larger emphasis on the identification of the minority class, prioritising reducing false negatives over false positives. Weighted cross entropy loss has seen deployment in a series of recent papers, seeing significant success compared to other state-of-the-art processes [167,168].

Ensemble models, algorithms which train multiple models with differing hyperparameters, have additionally emerged as a method of combating class imbalance. A series of models named “base learners” are trained on the data with a range of deviating hyperparameters, training strategies and random seeds, extracting different feature information from one another. Each individual model generates a prediction mask, which are combined to form a conclusive overall prediction image, often through group-voting procedures or weighted averages. While not highly present within the field of medical segmentation, the cases that do use it have seen promising results [130], and uses within the wider medical imaging field such as detection diagnosis [169] have been reported.

### Boundary complications

Of course, there too exist inherent flaws within the image data itself. Ambiguous or blurred boundaries between objects are commonplace and can lead to reduced accuracy or learning of poor features. A range of factors across different imaging modalities contribute to this process, such as the nucleus relaxation time when using MRI scanners, or deviating attenuation coefficients in objects in the body affecting CT scan quality. Due to the complexity of the image gathering processes and the fact that the flaws inherently lie within the human body, this challenge is not one so easily solved, although Image augmentation techniques have been shown to aid in this issue. Shariful et al. [125] presents a new augmentation approach, providing a comparison of results from a UNet trained on this data vs an identical model trained on conventional augmentation techniques. The model using the custom augmentation process displayed both heightened qualitative and quantitative performance, reducing the effect of overfitting and blurring of boundaries. While this technique does not address the issue at its source, it nonetheless reduces the impact of said issue. Further study will reveal the longevity and wider applicability of this approach.

Multi-modal fusion has also emerged as a promising approach to combat blurred boundaries and partial obscurity of key objects. Partial obscurity can lead to challenges in accurately delineating object boundaries or correct identification of objects entirely, and has proven itself to be a consistent challenge within the field. Multi-modal fusion integrates information from a range of imaging modalities, increasing the overall variation in data gathered and allowing for a more comprehensive image set. Mainly, the complementary nature of combining different modalities allows for increased differentiation of poorly defined or otherwise obscured object boundaries. Multi-modal fusion image generation has been utilised across the field, with significant recent use in the fields of brain segmentation [97,95,105,154] and liver segmentation [124,104].

Multi-modal imaging, despite its advantages, also suffers from a notable drawback when compared to single-mode imaging. The inherent challenge lies in the scarcity of multi-modal data, which necessitates the simultaneous availability of multiple scanning machines. This limitation exacerbates the data scarcity problem already present within medical segmentation, making it even more pronounced for the case of multi-modal imaging.

## Conclusion

This paper has detailed the general landscape of medical segmentation algorithms and relevant processes to the image processing procedure as of the time of writing. Key papers of current relevance or contextual significance have been discussed. While FCNs maintain dominance over the field seeing the vast majority of publications within recent years, advantages still lie within alternative architectures. GANs have particularly seen use in a range of related fields such as classification, diagnosis and image generation, a versatility in application largely unseen by other architectures.

While UNet-based algorithms have been widely adopted for segmentation tasks, individual adaptations of the model are commonplace, seeing tremendous publications of recent years claiming heightened accuracy and lower latency. Thus, fears of overfitting have led to the development of NNUNet, an automatic hyperparameter adjustment algorithm whose versatility has been demonstrated. NNUNet has shown promising results in training accurate segmentation models on limited training datasets, cementing its relevance within the industry.

Despite rapid advancement in the field, segmentation algorithms have not fully resolved the core challenges they face, despite methods of alleviating them becoming commonplace. Issues of data scarcity, quality and consistency continue to be the dominant limiting factor regarding advancement in segmenting accuracy, problems leading to the development of artificial data simulation via augmentation, generation or multi-modality gathering. Advancements continue to be made at a rapid pace to tackle these challenges.

## References

1. Keklikoglou K, S Faulwetter, E Chatzinikolaou. Micro-computed tomography for natural history specimens: A handbook of best practice protocols. 2019; 522: 1-55.
2. Ehlinger M, Favreau H, Murgier J, Ollivier M. Knee osteotomies: The time has come for 3D planning and patient-specific instrumentation. 2023; 109/4.
3. Fu YT, Mazur X, Wu S, Liu X, Chang Y, et al. A novel MRI Segmentation method using CNN based Correction Network for MRI Guided Adaptive Radiotherapy. 2018; 45.
4. Kirillov A, K He, R Girshick, C Rother, P Dollar. Panoptic Segmentation. 2019; 9396-9405.
5. Long J, E Shelhamer, T Darrell. Fully convolutional networks for semantic segmentation, (IEEE Computer Society). 2015; 3431-40.
6. Garcia-Garcia A, S Orts-Escolano, S Oprea. A survey on deep learning techniques for image and video semantic segmentation. 2018; 70: 41-65.
7. Chen LC, G Papandreou, I Kokkinos, K Murphy, A L Yuille. Deep Lab: Semantic Image Segmentation with
8. Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. 2017; 40/4: 834-48.
9. Mueed Hafiz A, G Mohiuddin Bhat. A survey on instance segmentation: state of the art. 2020; 9: 171-79.
10. Li X, Y Peng, M Xu. Edge-enhanced Instance Segmentation of Wrist CT via a Semi-Automatic Annotation Database Construction Method. 2021; 01-08.
11. Cui Z, C Li, W Wang. Tooth Net: Automatic Tooth Instance Seg-

- mentation and Identification From Cone Beam CT Images. 2019; 6361-70.
12. De Brabandere B, D Neven, L Van Gool. Semantic Instance Segmentation for Autonomous Driving. 2017; 478-80.
  13. Zhou S, D Nie, E Adeli, Q Wei, X Ren, et al. Semantic instance segmentation with discriminative deep supervision for medical images. 2022; 82.
  14. Salvador A, M Bellver, M Baradad, F Marques, J Torres, et al. Recurrent Neural Networks for Semantic Instance Segmentation. 2017.
  15. Fathi A, Z Wojna, V Rathod. Semantic Instance Segmentation via Deep Metric Learning. 2017.
  16. Adelson EH. On seeing stuff: the perception of materials by humans and machines. 2001; 1-12.
  17. Chen S, C Ding, M Liu, J Cheng, D Tao. CPP-Net: Context-Aware Polygon Proposal Network for Nucleus Segmentation. 2023; 32: 980-94
  18. Hliboký M, M Bundzel, S Husár. State-of-the-Art in Lung Ultrasound Processing - Brief Review. 2023; 000091-000096.
  19. Dou QL, Yu H, Chen Y, Jin X, Yang J, et al. Heng, 3D deeply supervised network for automated segmentation of volumetric medical images. 2017; 41: 40-54.
  20. Crespi L, D Loiacono, P Sartori. Are 3D better than 2D Convolutional Neural Networks for Medical Imaging Semantic Segmentation?. 2022; 1-8.
  21. Diakogiannis FI, F Waldner, P Caccetta, C Wu. ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data. 2020; 162: 94-114.
  22. Avesta A, S Hossain, M Lin, M Aboian, HM Krumholz, et al. Comparing 3D, 2.5D, and 2D Approaches to Brain Image Auto-Segmentation. 2023; 10(2).
  23. Zhang Y. Cascaded Convolutional Neural Network for Automatic Myocardial Infarction Segmentation from Delayed-Enhancement Cardiac MRI. 2021; 328-33.
  24. Ou Y, Y Yuan, X Huang. Lambda UNet: 2.5D Stroke Lesion Segmentation of Diffusion-Weighted MR Images. 2021.
  25. Hesamian MH, W Jia, X He, P Kennedy. Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges. 2019; 32(4): 582-96.
  26. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. 2017; 36: 61-76.
  27. Zhang Y, Q Liao, L Ding, J Zhang. Bridging 2D and 3D segmentation networks for computation-efficient volumetric medical image segmentation: An empirical study of 2.5D solutions. 2022; 99: 102088.
  28. Zhou Y, W Huang, P Dong, Y Xia, S Wang. D-UNet: A Dimension-Fusion U Shape Network for Chronic Stroke Lesion Segmentation. 2021; 18(3): 940-50.
  29. Isgum I, M Staring, A Rutten, M Prokop, MA Viergever, et al. Multi-atlas-based segmentation with local decision fusion-application to cardiac and aortic segmentation in CT scans. 2009; 28(7): 1000-1010.
  30. Iglesias, J. E. and M. R. Sabuncu, 'Multi-atlas segmentation of biomedical images: A survey. 2015; 24(1): 205-19.
  31. Wachinger C, P Golland. Atlas-Based Under-Segmentation. 2014; 17(01): 315-22.
  32. He Y, S Zhang, Y Luo, H Yu, Y Fu, et al. Quantitative Comparisons of Deep-learning-based and Atlas-based Auto-segmentation of the Intermediate Risk Clinical Target Volume for Nasopharyngeal Carcinoma. 2022; 18(3): 335-45.
  33. Gibbons E, M Hoffmann, J Westhuyzen, A Hodgson, B Chick, et al. Clinical evaluation of deep learning and atlas-based auto-segmentation for critical organs at risk in radiation therapy. 2023; 70: 15-25.
  34. Xie L, LEM Wisse J, Wang S, Ravikumar P, Khandelwal T, et al. Deep label fusion: A generalizable hybrid multi-atlas and deep convolutional neural network for medical image segmentation. 2023; 83.
  35. Ding W, L Li, X Zhuang, L Huang. Cross-Modality Multi-Atlas Segmentation via Deep Registration and Label Fusion. 2022; 26(7): 3104-15.
  36. Goodfellow I, J Pouget-Abadie, M Mirza. Generative Adversarial Networks. 2014; 3(11).
  37. Ashok M, A Gupta. Comparative Study of TRANS - GAN Architecture for Bio-Medical Image Semantic Segmentation. 2022.
  38. Mirza M, S Osindero. Conditional generative adversarial nets. 2014.
  39. Radford A, Metz L. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. 2015.
  40. Li G, N Jamil, R Hamzah. Leveraging Unlabeled Data Using Semi-Supervised Generative Adversarial Network for Medical Image Segmentation. 2022; 67-72.
  41. Shabani SM, Homayounfar M, Koohi-Moghadam V, Vardhanabhuti M, A Nikouei, et al. Self-supervised region-aware segmentation of COVID-19 CT images using 3D GAN and contrastive learning. 2022; 149.
  42. Mu J, S De Mello, Z Yu, N Vasconcelos, X Wang, et al. CoordGAN: Self-Supervised Dense Correspondences Emerge from GANs, (Piscataway: The Institute of Electrical and Electronics Engineers, Inc. (IEEE). 2022.
  43. Chaturvedi K, A Braytee, J Li, M Prasad. SS-CPGAN: Self-Supervised Cut-and-Pasting Generative Adversarial Network for Object Segmentation. 2023; 23(3649): 3649-3649.
  44. Neocognitron: A new algorithm for pattern recognition tolerant of deformations and shifts in position. 1980; 15(6): 455-69.
  45. Yu F, V Koltun. Multi-Scale Context Aggregation by Dilated Convolutions. 2016.
  46. Yamashita R, M Nishio, RKG Do, K Togashi. Convolutional neural networks: An overview and application in radiology. 2018; 9(4): 611-29.
  47. Paszke A, A Chaurasia, S Kim, E Cukurciello. ENet: A Deep Neural Network Architecture for Real-Time Semantic Segmentation. 2016.
  48. Chen L. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. 2018.
  49. Girshick R, J Donahue, T Darrell, J Malik. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. 2014; 580-87.
  50. Girshick R. Fast R-CNN. 2015; 1440-48.
  51. Ren S, K He, R Girshick, J Sun. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. 2017; 39(6): 1137-49

52. He K, G Gkioxari, P Dollar, R Girshick. Mask R-CNN. 2017; 42(2): 386-97.
53. Jifeng Dai, Kaiming He, Jian Sun, Instance-Aware Semantic Segmentation via Multi-task Network Cascades. (IEEE). 2016; 3150-58
54. Zhang Y, S Chan, VY Park, KT Chang, S Mehta, et al. Automatic Detection and Segmentation of Breast Cancer on MRI Using Mask R-CNN Trained on Non-Fat-Sat Images and Tested on Fat-Sat Images. 2022; 29(Supplement 1): S135-44.
55. UI Haq MNA, Irtaza N, Nida MA, Shah L. Zubair, Liver Tumor Segmentation using Resnet based Mask-R-CNN. 2021; 276-81.
56. Dandil E, MS Yildirim. Automatic Segmentation of COVID-19 Infection on Lung CT Scans using Mask R-CNN. 2022; 1-5.
57. Hu QLF, de F Souza GB, Holanda SSA, Alves FH, Dos S Silva, et al. An effective approach for CT lung segmentation using mask region-based convolutional neural networks. 2020; 103: 101792.
58. Dandil E, MS Yildirim. A Mask R-CNN based Approach for Automatic Lung Segmentation in Computed Tomography Scans. 2021; 1-6.
59. Sakunpaisanwari LN, Yodrabum T, Sirirapisit T, Titijaronroj. Blood Vessels Detection by Regional-based CNN for CT Scan of Lower Extremities. 2022; 14-19.
60. Felfeliyan BA, Hareendranathan G, Kuntze JL, Jaremko, JL Ron-sky. Improved-Mask R-CNN: Towards an accurate generic MSK MRI instance segmentation platform (data from the Osteoarthritis Initiative). 2022; 97: 102056.
61. Dogan ROH, Dogan C, Bayrak, T Kayikcioglu. A Two-Phase Approach using Mask R-CNN and 3D U-Net for High-Accuracy Automatic Segmentation of Pancreas in CT Imaging. 2011; 207: 106141.
62. Chlebus GA, Schenk JH, Moltz B, van Ginneken HK, Hahn, et al. Automatic liver tumor segmentation in CT with fully convolutional neural networks and object-based postprocessing. 2018; 8(1): 1-7.
63. Lei YS, Tian X, He T, Wang B, Wang P, et al. Ultrasound prostate segmentation based on multidirectional deeply supervised V-Net. 2019; 46(7): 3194-3206.
64. Wang BY, Lei S, Tian T, Wang Y, Liu P, et al. Deeply supervised 3D fully convolutional networks with group dilated convolution for automatic MRI prostate segmentation. 2019; 46(4): 1707-18.
65. Ronneberger O, P Fischer. U-Net: Convolutional Networks for Biomedical Image Segmentation. 2015; 234-41.
66. Hinton GE, RR Salakhutdinov. Reducing the dimensionality of data with neural networks. 2006; 313(5786): 504-7.
67. Badrinarayanan VA, Kendall, R Cipolla. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. 2017; 39(12): 2481-95.
68. Baccouch WS, Oueslati B, Solaiman, S Labidi. A comparative study of CNN and U-Net performance for automatic segmentation of medical images: Application to cardiac MRI. 2023; 219: 1089-96.
69. Yin X, XL Sun, Y Fu, R Lu, Y Zhang. U-Net-Based Medical Image Segmentation. 2022; 4189781.
70. Zhang P, Zhong Y. ACCL: Adversarial constrained-CNN loss for weakly supervised medical image segmentation. 2020.
71. Zhou ZMMR, Siddiquee N, Tajbakhsh, J Liang. UNet++: A Nested U-Net Architecture for Medical Image Segmentation. 2018; 11045: 3-11.
72. Zhou ZMMR, Siddiquee N, Tajbakhsh, J Liang. UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation. 2020; 39(6): 1856-67.
73. Zhang LK, Zhang, H Pan, SUNet++: A Deep Network with Channel Attention for Small-Scale Object Segmentation on 3D Medical Images. 2023; 28(4): 628-38.
74. Oktay OJ, Schlemper LL, Folgoc MJ, Lee M, Heinrich K, et al. Attention U-Net: Learning Where to Look for the Pancreas. 2018.
75. Vladimirov NE, Brui A, Levchuk, W Al-Haidri, V Fokin, et al. CNN-based fully automatic wrist cartilage volume quantification in MR images: A comparative analysis between different CNN architectures. 2023; 90(2): 737-51.
76. Isensee FJ, Petersen A, Klein D, Zimmerer P, Jaeger SAA. et al. nnU-Net: Self-adapting Framework for U-Net-Based Medical Image Segmentation. 2018.
77. Rebaud L, T Escobar. Simplicity Is All You Need: Out-of-the-Box nnUNet Followed by Binary-Weighted Radiomic Model for Segmentation and Outcome Prediction in Head and Neck PET/CT. 2023.
78. Huang SL, Mei J, Li Z, Chen Y, Zhang T. et al. Abdominal CT Organ Segmentation by Accelerated nnUNet with a Coarse to Fine Strategy. 2022.
79. Alqaoud MJ, Plemmons E, Feliberti S, Dong K, Kaipa G, et al. Audette, nnUNet-based Multi-modality Breast MRI Segmentation and Tissue-Delineating Phantom for Robotic Tumor Surgery Planning. 2022; 3495-3501.
80. Schachner CD, Gabauer D, Brunner, L Fischer. 2D nnUNet for classification and segmentation of anatomical structures in fetal torso ultrasound. 2022.
81. Lee SD, Wang J, Natarajan N, Yegya-raman T, Kegelman S, et al. Cardiac Substructure Segmentation Using Self-Configuring NnUNet and NnFormer for Cardiac-Sparing Lung Cancer Radiotherapy. 2022; 49(6): E786-87.
82. Ma JJ, Chen. NnUNet with Region-based Training and Loss Ensembles for Brain Tumor Segmentation, Springer Science and Business Media Deutschland GmbH. 2022.
83. Zhang LJ, Zhang P, Shen G, Zhu P, Li X, et al. Block Level Skip Connections Across Cascaded V-Net for Multi-Organ Segmentation. 2020; 39(9): 2782-93.
84. Hussain FR, Ksantini, K Tbarki. GAN-based one-class classification SVM for real time medical image intrusion detection. 2023; 13(1): 625-41.
85. Kuttala DD, Mahapatra R, Subramanian, VRM Oruganti. Dense attentive GAN-based one-class model for detection of autism and ADHD. 2022; 34(10): 10444-58.
86. RRT, VKKS, Artificial MRI Image Generation using Deep Convolutional GAN and its Comparison with other Augmentation Methods. 2021; 1: 1-6.
87. Guo KJ, Chen T, Qiu S, Guo T, Luo T, et al. MedGAN: An adaptive GAN approach for medical image generation. 2023; 163: 107119.
88. Shi J, X Liu, G Yang, G Wang. Enhanced CT Image Generation by GAN for Improving Thyroid Anatomy Detection. 2022; 17: 14-17.
89. Guan QY, Chen Z, Wei AA, Heidari H, Hu XH. et al. Medical image augmentation for lesion detection using a texture-constrained multichannel progressive GAN. 2022; 145: 105444.

90. Hammami MD, Friboulet, R Kechichian. Cycle GAN-Based Data Augmentation for Multi-Organ Detection in CT Images Via Yolo. 2020; 390-93.
91. Zhang XJ, Wu S, Fan M, Li G, Yuan Y, et al. Multi-Scale Hybrid Attention Cascade Network for Pancreas Segmentation. 2020; 277-81.
92. Li M, F Lian, S Guo. An Adversarial Network Embedded with Attention Mechanism for Pancreas Segmentation. 2021; 631-34.
93. Ranjbarzadeh RS, Jafarzadeh Ghouschi M, Bendeche A, Amirabadi MN, Ab Rahman, et al. Lung Infection Segmentation for COVID-19 Pneumonia Based on a Cascade Convolutional Network from CT Images. 2021; 5544742.
94. Saha S, S Dutta. ADU-Net: An Attention Dense U-Net based deep supervised DNN for automated lesion segmentation of COVID-19 from chest CT images. 2023; 85.
95. Mahmud TMA, Rahman SA, Fattah, SY Kung. CovSegNet: A Multi Encoder-Decoder Architecture for Improved Lesion Segmentation of COVID-19 Chest CT Scans. 2021; 2(3): 283-97.
96. Shangzhu J, S Yu. A novel medical image segmentation approach by using multi-branch segmentation network based on local and global information synchronous learning. 2023; 13.
97. Liu Z, J Wei, R Li, J Zhou. Learning multi-modal brain tumor segmentation from privileged semi-paired MRI images with curriculum disentanglement learning. 2023; 159: 106927.
98. Vijendran A, K Ramasamy. Optimal segmentation and fusion of multi-modal brain images using clustering based deep learning algorithm. 2023; 27.
99. Xu R, C Wang, S Xu, W Meng, X Zhang. Dual-stream Representation Fusion Learning for accurate medical image segmentation. 2023; 123: 106402.
100. Liu L, X Fan, X Zhang, Q Hu. Lightweight Dual-Domain Network for Real-Time Medical Image Segmentation. 2022; 396-400.
101. Li Y, J Yang, J Ni, A Elazab, J Wu. TA-Net: Triple attention network for medical image segmentation. 2021; 137: 104836. Gu R, G Wang, T Song, R Huang, M Aertsen, et al. CA-Net: Comprehensive Attention Convolutional Neural Networks for Explainable Medical Image Segmentation. 2021; 40(2): 699-711.
102. Zhou Y, Z Li, S Bai, X Chen, M Han, et al. Prior-Aware Neural Network for Partially-Supervised Multi-Organ Segmentation. 2019; 10671-80.
103. Cai Y, M Liu, Y Zheng, W Yang, Y Long, et al. Swin Unet3D: A three-dimensional medical image segmentation network combining vision transformer and convolution. 2023; 23(1).
104. Elghazy HL, MW Fakhr. Dual- and triple-stream RESUNET (UNET architectures for multi-modal liver segmentation. 2023; 17(4): 1224-35.
105. Kumari KHV, SS Barpanda. Residual UNet with Dual Attention-An ensemble residual UNet with dual attention for multi-modal and multi-class brain MRI segmentation. 2023; 33(2): 644-58.
106. Asnawi MH, AA Pravitarsari, G Darmawan, T Hendrawati, IN Yulita, et al. Lung and Infection CT-Scan-Based Segmentation with 3D UNet Architecture and Its Modification. 2023; 11(1): 213.
107. Li Z, N Zhang, H Gong, R Qiu, W Zhang. MFA-Net: Multiple Feature Association Network for medical image segmentation. 2023; 158.
108. Steybe D, Poxleitner P, Metzger MC. Automated segmentation of head CT scans for computer-assisted craniomaxillofacial surgery applying a hierarchical patch-based stack of convolutional neural networks. 2022; 17(11): 2093-2101.
109. Hassanzadeh T, D Essam. Eevou-Net: An Ensemble of Evolutionary Deep Fully Convolutional Neural Networks for Medical Image Segmentation. 2023; 34: 1387-1404.
110. Dang H, M Li. LVSegNet: A novel deep learning-based framework for left ventricle automatic segmentation using magnetic resonance imaging. 2023; 208: 124-35.
111. Sun Y, H Kang, H Zhang, Y Wang, H Shen, et al. A Global-Local Cascade Network for Multi-bone Segmentation in Chest CT. 2022; 809-12.
112. Abdo Y, N Mohamed, M Alsawaf, M Elsaed. Teeth and Jaw Segmentation from CBCT images Using 3D Deep Learning Models. 2022; 25-30.
113. Hsu K, DY Yuh, SC Lin, PS Lyu, GX Pan, et al. Improving performance of deep learning models using 3.5D U-Net via majority voting for tooth segmentation on cone beam computed tomography. 2022; 12(1): 1-15
114. Peng Q, X Chen, C Zhang, W Li, J Liu, et al. Deep learning-based computed tomography image segmentation and volume measurement of intracerebral hemorrhage. 2022; 16: 965680.
115. Wang Y, H Feng. Method for Automatic mandibular canal detection on improved U-Net. 2022; 206-9.
116. Singh KR, A Sharma, GK Singh. Cardiac Magnetic Resonance Imaging Segmentation using Ensemble of 2D and 3D Deep Residual U-Net. 2022; 1-6.
117. Inkeaw P, S Angkurawaranon, P Khumrin, N Inmutto, P Traisathit, et al. Automatic hemorrhage segmentation on head CT scan for traumatic brain injury using 3D deep learning model. 2022; 146(2022): 105530.
118. Jiang S, J Li, Z Hua. DPCFN: Dual path cross fusion network for medical image segmentation. 2022; 116.
119. Peng Y, D Yu, Y Guo. MShNet: Multi-scale feature combined with h-network for medical image segmentation. 2022; 79.
120. Guan X. 3D AGSE-VNet: An automatic brain tumor MRI data segmentation framework. 2022; 22(1): 1-18.
121. Diniz JOB, JL Ferreira, OAC Cortes, AC Silva, AC de Paiva. An automatic approach for heart segmentation in CT scans through image processing techniques and Concat-U-Net. 2022; 196.
122. Zhou Q, Q Wang. LAEDNet: A Lightweight Attention Encoder-Decoder Network for ultrasound medical image segmentation. 2022.
123. Sun Q, M Dai, Z Lan, F Cai, L Wei, et al. UCR-Net: U-shaped context residual network for medical image segmentation. 2022; 151(Part A).
124. Xue Z, P Li, L Zhang, X Lu, G Zhu, et al. Multi-Modal Co-Learning for Liver Lesion Segmentation on PET-CT Images. 2021; 40(12): 3531-42.
125. Alam MS, D Wang, A Sowmya. Image data augmentation for improving performance of deep learning-based model in pathological lung segmentation. 2021; 1-5.
126. Zhang C, T Yang, Y Zhang. Learning Medical Image Segmentation via U-net Based Positive and Unlabeled Learning Method. 2021; 1-6.
127. Xie H, JF Zhang. Application of Deep Convolution Network to Automated Image Segmentation of Chest CT for Patients with Tumor. 2021; 11.

128. Zhang F, Y Wang. Efficient Context-Aware Network for Abdominal Multi-organ Segmentation. 2021.
129. Fu X, L Bi, A Kumar, M Fulham, J Kim. Multimodal Spatial Attention Module for Targeting Multimodal PET-CT Lung Tumor Segmentation. 2021; 25(9): 3507-16.
130. Enshaei N, P Afshar, S Heidarian, A Mohammadi, MJ Rafiee, et al. An Ensemble Learning Framework for Multi-Class Covid-19 Lesion Segmentation from Chest Ct Images. 2021; 1-6.
131. Almeida G. and J. M. R. S. Tavares. Versatile Convolutional Networks Applied to Computed Tomography and Magnetic Resonance Image Segmentation. 2021; 45(8): 1-10.
132. Luca AR, TF Ursuleanu, L Gheorghe, R Grigorovici, S Iancu, et al. Designing a High-Performance Deep Learning Theoretical Model for Biomedical Image Segmentation by Using Key Elements of the Latest U-Net-Based Architectures. 2021; 9(7): 8-20.
133. Zopes, J. 'Multi-Modal Segmentation of 3D Brain Scans Using Neural Networks. 2021; 12.
134. Mohammadi R, I Shokatian, M Salehi, H Arabi, I Shiri, et al. Deep learning-based auto-segmentation of organs at risk in high-dose rate brachytherapy of cervical cancer. 2021; 159: 231-40.
135. Asipong K, S Gabbualoy, P Phasukkit. Coronavirus infected lung CT scan image segmentation using Deep Learning. 2021; 773-76.
136. Xu G, H Cao, JK Udupa, Y Tong, DA Torigian. DiSegNet: A deep dilated convolutional encoder-decoder architecture for lymph node segmentation on PET/CT images. 2021; 88: 101851.
137. Khan A, H Kim, L Chua. PMED-Net: Pyramid Based Multi-Scale Encoder-Decoder Network for Medical Image Segmentation. 2021; 9: 55988-98.
138. Li L, X Zhao, W Lu, S Tan. Deep Learning for Variational Multimodality Tumor Segmentation in PET/CT. 2020; 392: 277-95.
139. Jemaa S, J Fredrickson, RAD Carano. Tumor Segmentation and Feature Extraction from Whole-Body FDG-PET/CT Using Cascaded 2D and 3D Convolutional Neural Networks. 2020; 33: 888-94.
140. Wang EK, CM Chen, MM Hassan, A Almogren. A deep learning based medical image segmentation technique in Internet-of-Medical-Things domain. 2020; 108: 135-44.
141. Liu L, Y Liu, J Zhou, C Guo, H Duan. A novel MCF-Net: Multi-level context fusion network for 2D medical image segmentation. 2020; 226.
142. Wang R, S Chen, C Ji, J Fan, Y Li. Boundary-aware context neural network for medical image segmentation. 2020; 78.
143. Zhang Z, C Wu, S Coleman, D Kerr. DENSE-INception U-net for medical image segmentation. 2020; 192.
144. Jha D, MA Riegler, D Johansen, P Halvorsen, HD Johansen. DoubleU-Net: A Deep Convolutional Neural Network for Medical Image Segmentation, (IEEE Computer Society). 2020; 558-64.
145. Kim H, J Jung, J Kim, B Cho, J Kwak, et al. Abdominal multi-organ auto-segmentation using 3D-patch-based deep convolutional neural network. 2020; 10: 1-9.
146. Moradi S, MG Oghli, A Alizadehasl, I Shiri, N Oveisi, et al. MFP-UNet: A novel deep learning based approach for left ventricle segmentation in echocardiography. 2019; 67: 58-69.
147. Duan J, G Bello, J Schlemper, W Bai, TJ W Dawes, et al. Automatic 3D Bi-Ventricular Segmentation of Cardiac Images by a Shape-Refined Multi- Task Deep Learning Approach. 2019; 38(9): 2151-64.
148. Chan JW, V Kearney, S Haaf, S Wu, M Bogdanov, et al. A convolutional neural network algorithm for automatic segmentation of head and neck organs at risk using deep lifelong learning. 2019; 46(5): 2204-13.
149. Weng Y, T Zhou, Y Li, X Qiu. NAS-Unet: Neural Architecture Search for Medical Image Segmentation. 2019; 7: 44247-57.
150. Harten LD, van JMH, Noothout J, Verhoeff J, Wolterink, et al. Automatic Segmentation of Organs at Risk in Thoracic CT scans by Combining 2D and 3D Convolutional Neural Networks. 2019.
151. Zhang J, A Saha. Hierarchical Convolutional Neural Networks for Segmentation of Breast Tumors in MRI With Application to Radiogenomics. 2019; 38: 435-47.
152. Liang S, KH Thung, D Nie, Y Zhang, D Shen. Multi-View Spatial Aggregation Framework for Joint Localization and Segmentation of Organs at Risk in Head and Neck CT Images. 2020; 39(9): 2794-2805.
153. Raja M, S Vijayachitra. A hybrid approach to segment and detect brain abnormalities from MRI scan. 2023; 216.
154. Tyagi S, SN Talbar. CSE-GAN: A 3D conditional generative adversarial network with concurrent squeeze-and-excitation blocks for lung nodule segmentation. 2022; 147.
155. Wei X, X Chen, C Lai, Y Zhu, H Yang, et al. Automatic Liver Segmentation in CT Images with Enhanced GAN and Mask Region-Based CNN Architectures. 2021; 1-11.
156. Xu L. GAN-based Automatic Segmentation of Thoracic Aorta from Non-Contrast-Enhanced CT Images. 2021.
157. Murugesan B, K Sarveswaran, V Raghavan S, SM Shankaranarayana, K Ram, et al. A Context Based Deep Learning Approach for Unbalanced Medical Image Segmentation. 2020; 1949-53.
158. Costa P, A Galdran, MI Meyer, M Niemeijer, M Abràmoff, et al. End-to-End Adversarial Retinal Image Synthesis. 2018; 37(3): 781-91.
159. Tu D, P Lin, H Chou, M Shen, S Hsieh. Slice-Fusion: Reducing False Positives in Liver Tumor Detection for Mask R-CNN. 2023; 99: 1-11.
160. Cai Z, N Vasconcelos. Cascade R-CNN: High Quality Object Detection and Instance Segmentation. 2021; 43(5): 1483-98
161. Krizhevsky A, I Sutskever, G Hinton. ImageNet Classification with Deep Convolutional Neural Networks. 2012; 25.
162. Shorten C, TM Khoshgoftaar. A survey on Image Data Augmentation for Deep Learning. 2019; 6(1): 1-48.
163. Weiss K, TM Khoshgoftaar, D Wang. A survey of transfer learning. 2016; 1(3): 1-40.
164. Apostolopoulos ID, TA Mpesiana. Covid-19: Automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. 2020; 43(2): 635-40.
165. Das D, KC Santosh, U Pal. Truncated inception net: COVID-19 outbreak screening using chest X-rays. 2020; 43(3): 915-25.
166. Mann P, S Jain, S Mittal, A Bhat. Generation of COVID-19 Chest CT Scan Images using Generative Adversarial Networks. 2021; 1-5.
167. Goyal P. Shallow SegNet with bilinear interpolation and weighted cross-entropy loss for Semantic segmentation of brain tissue. 2022; 1: 361-65.
168. Atug. Class Distance Weighted Cross-Entropy Loss for Ulcerative Colitis Severity Estimation, (Springer Science and Business Media Deutschland GmbH. 2022; 13413.

