Network Modeling Analysis in Health Informatics and Bioinformatics (2025) 14:58
https://doi.org/10.1007/s13721-025-00556-8

REVIEW ARTICLE

Convolutional neural networks for tumor segmentation by cancer type and imaging modality: a systematic review

Oona Rainio[1] · **Riku Klén**[1]

Received: 20 January 2025 / Revised: 11 May 2025 / Accepted: 20 June 2025

Abstract

During the last few decades, a convolutional neural network (CNN) has become a very popular deep learning technique in automatic tumor segmentation of medical images of cancer patients. However, unlike for semantic segmentation of regular photographs, there are few publicly available datasets that can be used to train a CNN to perform tumor segmentation in medical images. Consequently, it is difficult to compare these methods or tell how well they should work for a specific combination of cancer type and imaging modality when trained with a certain amount of data. To answer this question, we analyzed 325 recent articles about CNNs trained for tumor segmentation in order to give a comprehensive overview of the current state of research. The articles study several different types of cancer, including brain tumors and breast, liver, lung, skin, head and neck, prostate, thyroid, cervical, colorectal, pancreatic, kidney, and bladder cancer, imaged with magnetic resonance imaging (MRI), computed tomography, positron emission tomography, ultrasound, and other similar modalities. According to our analysis, the most popular CNN for tumor segmentation is U-Net and its new modifications. Conversely, Mask region-based CNNs are rarely used outside of MRI images. Out of the other CNN designs, SegNet and DeepLabV3 are most common but still significantly less studied than U-Net. Furthermore, several methods have not yet been tested for rarer types of cancer.

Keywords Convolutional neural network · Medical imaging · Tumor segmentation · U-Net

1 Introduction

Cancer is a leading cause of death globally, responsible of nearly one in six deaths in the world (Debela 2021). It is a complicated disease whose successful treatment often requires multiple medical imaging scans to monitor the disease, create individualized treatment plans, and estimate the treatment response. To analyze these medical images in an efficient way, several image processing tasks are required.

In clinical oncology, tumor segmentation is one of the necessary routine tasks performed after imaging of a cancer patient. It means denoting tumor segments in the image by creating a tumor mask, which is typically a matrix in the same size as the original image with all the elements labelled as cancerous or non-cancerous according to the corresponding pixels or voxels. Since tumor segmentation is both time-consuming and relatively monotonic, several automatic, deep learning-based applications have been developed for this purpose (Jiang et al. 2022).

The most popular method for automatic tumor segmentation, a convolutional neural network (CNN), is a special type of machine learning technique commonly used for image processing because it is able to understand the spatial relationships between adjacent data points instead of treating them as separate variables without a meaningful order. The signature features of CNNs were introduced by several researchers on different decades during the 20th century, though the most notably of the persons involved is perhaps Yann LeCun, who coined the term "convolution" and developed CNNs capable of interpreting hand-written digits in 1989 (Li et al. 2021). Since then, CNNs have become increasingly popular with numerous applications related to, for instance, face recognition (Lawrence et al. 1997), self-driving vehicles (Farag 2018), and video surveillance (Dhiyanesh et al. 2021).

✉ Oona Rainio
ormrai@utu.fi

Riku Klén
riku.klen@utu.fi

[1] Turku PET Centre, University of Turku and Turku University Hospital, Turku, Finland

Published online: 07 July 2025

CNNs have been applied to medical purposes from very early on Zhang et al. (1994). However, due to the privacy concerns, there are a quite limited number of publicly available datasets (Anwar et al. 2018) and the research of medical CNNs is focused on the hospitals and research centers that are able to collect their own image data to evaluate the models used. While the CNNs trained to perform segmentation and object detection in typical photographs can be evaluated in large open-source datasets and therefore easily compared with each other, it is more difficult to say how well a CNN should work for medical images in a specific task, whether certain modifications significantly improve its performance, or which CNN architecture could truly be considered the state-of-the-art.

To address this issue, we analyze CNNs used for tumor segmentation in different medical images. Our review includes articles based on brain tumors, head and neck cancer (HNC), and breast, liver, lung, skin prostate, thyroid, cervical, colorectal, pancreatic, kidney, and bladder cancer. The main imaging modalities considered include magnetic resonance imaging (MRI), computed tomography (CT), positron emission tomography (PET), and ultrasound (US). We select 325 articles in total for our systematic review synthesis and, based on the reported results, choose 54 of them for a more careful review. In this way, we form a representative collection of the best CNNs for each cancer type and modality combination.

The structure of this review is as follows. First, we briefly explain our methods of article selection and analysis. Then we introduce necessary concepts related to medical imaging, CNNs, and the evaluation of segmentation results, together with the results obtained during our systematic review synthesis. After that, we introduce a few accurate CNN methods for each cancer type separately. In Discussion, we summarize our key findings and consider the future prospects of this area of study. At the end, we also present a brief conclusion.

2 Article selection

We performed literature analysis according to the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) reporting guidelines of year 2020 (Page et al. 2021). The articles for this review were chosen among the 810 search results on the three-year time period 1.4.2021−31.3.2024 for the search phrase *cancer tumor segmentation convolutional neural network* from PubMed. We screened the titles and abstracts of these search results and excluded 466 articles as shown in Fig. 1.

Our exclusion criteria were to leave out such articles that 1. only studied diagnosis, classification, outcome prediction, or some task other than segmentation, 2. performed only organs-at-risk type segmentation without denoting the cancer, 3. performed segmentation either for non-cancerous targets such as cysts or only for metastases outside the organ with the primary tumor, 4. tested the CNN on datasets on several fully different cancer types, 5. used histopathological images, 6. did not use CNNs for the actual segmentation, 7. did not use human patients, or 8. were reviews or editorials. However, we included articles about segmentation of breast tumors and pulmonary or thyroid nodules, even though a majority of them are benign.

For the 325 included articles in the systematic review synthesis, we collected the following details from abstract if mentioned: cancer type according to the organ with the primary tumor (i.e. breast cancer instead of the specific subtype of breast cancer), imaging modality, CNN architecture, name of the public dataset if used, and the numeric values of the evaluation metrics from the segmentation results of the test set. From the 325 articles, we selected 54 articles for more careful analysis. This choice was made to include articles that 1.) composed a diverse collection of imaging modalities and cancer types and 2.) reported high segmentation accuracy when compared to other studies about the same modality and cancer type combination. The details of the all 810 articles, including the reason for excluding them or the details collected from the systematic review synthesis, can be found from the excel file included in the supplementary material of this article.

3 Cancer types

According to our systematic review synthesis from 325 articles, CNN-based tumor segmentation has been studied for brain tumors, breast, liver, lung, skin, head and neck, prostate, thyroid, cervical, colorectal, pancreatic, kidney, and bladder cancer. There are notably discrepancies between the popularity of certain cancer types in the reviewed articles and their global prevalence and mortality rates, as illustrated in Fig. 2. In particular, brain tumor segmentation is very commonly studied despite the relatively low prevalence of brain tumors whereas the amount of research of colorectal cancer segmentation is very limited when considering its high prevalence.

4 Imaging modalities

For this review, we focus on 3D imaging modalities such as MRI, CT, PET as well as 2D US and photographs. Out of these modalities, MRI is a very common medical imaging method that uses strong magnetic fields and radio waves. Dynamic contrast enhanced MRI (DCE-MRI) measures changes in spin–lattice relaxation of tissue caused by a gadolinium contrast bolus and can be used to assess tumor

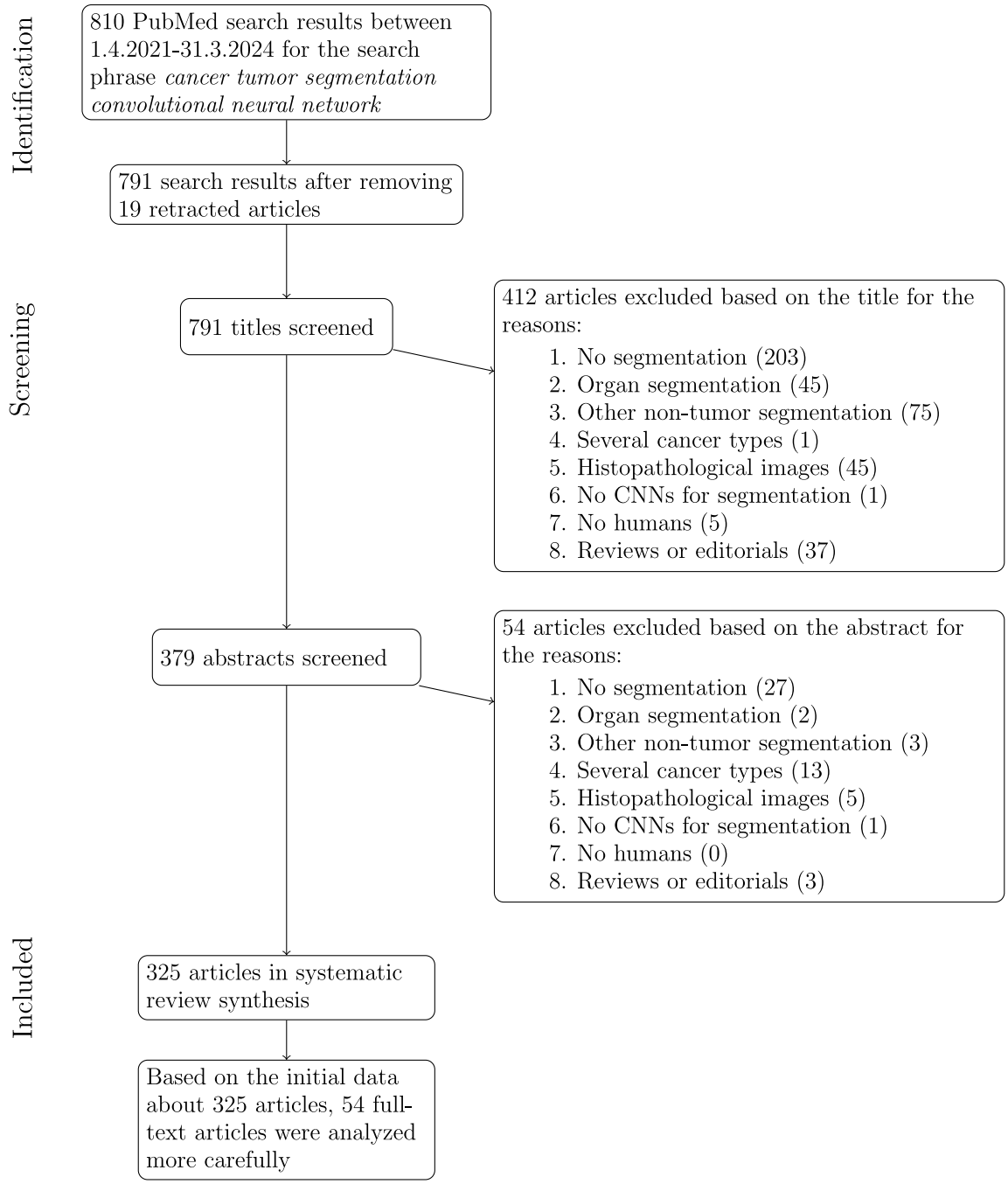

Fig. 1 The article selection process of this review visualized as a flow diagram according to the PRISMA 2020 guidelines

perfusion (Berman et al. 2016). CT is another common imaging technique, which uses rotating X-ray tube to gain anatomic information about different tissue types. MRI and CT both typically result in three-dimensional (3D) images, but it is also possible to record their dynamic sequence, giving us four-dimensional (4D) data.

PET is a nuclear medicine imaging method that uses short-lived radioactive isotopes to extract information about the function of the human body (Townsend et al. 2004). During imaging, a patient is first injected with a tracer substance and then a scanner device records how the substance travels in the body based on the radioactive decay (Townsend et al. 2004). The results are presented either as a 3D static image or a 4D dynamic image. In oncology, ^{18}F-fluorodeoxyglucose (^{18}F-FDG) is the most common tracer because it accumulates in the cancer tumors due to their higher metabolism. Single photon emission computed tomography (SPECT) is another nuclear medicine imaging

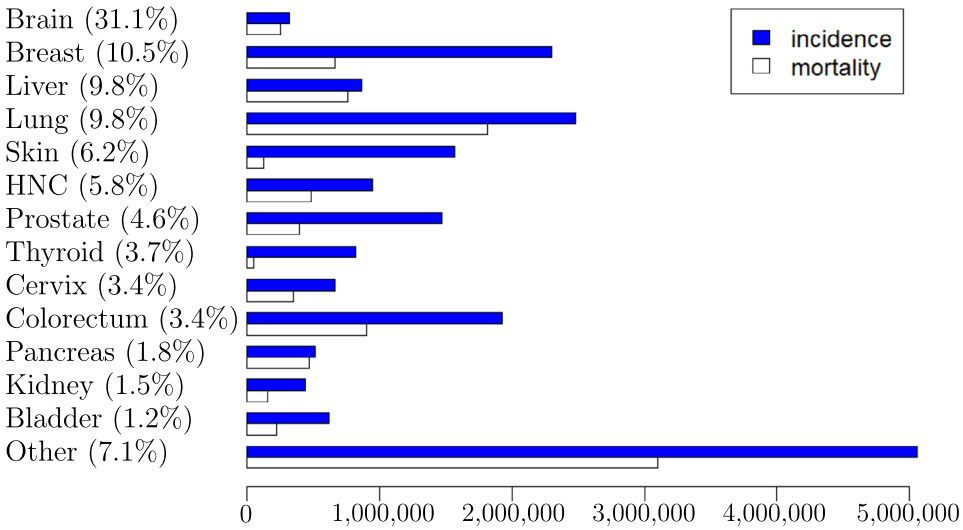

Fig. 2 The global incidence and mortality rates in 2022 according to Global Cancer Observatory (2025) for the most common cancer types in CNN-based tumor segmentation articles. The percentage is the relative prevalence of the cancer type in the 325 articles selected for this review. The incidence and the mortality of head and neck cancer (HNC) are obtained as the sums of cases in lips, oral cavity, larynx, nasopharynx, oropharynx, hypopharynx, and salivatory glands

method, though less common than PET and based on measuring radioactive decay through gamma rays instead of positrons.

US imaging is used especially for breast cancer patients, and, unlike MRI, CT, and PET images, US images are typically two-dimensional (2D). An exception to this is automated whole-breast US (ABUS) and automated breast volume scanner (ABVS), which both give 3D images. More traditional imaging methods for breast cancer patients include a special type of an X-ray image called a mammogram which can sometimes be 3D but is typically 2D.

Additionally, there are several methods based on 2D photographs. Depending on the location of the cancer, images taken during endoscopy, laparoscopy, colposcopy, or colonoscopy can be of interest. Furthermore, dermoscopic images are commonly analyzed for skin cancer and intraoral photographs can show HNC tumors located in the mouth.

The stacked bar chart of Fig. 3 shows the distribution of the selected articles between different imaging modalities within each cancer type of Fig. 2. Given imaging modalities are better suited for certain cancer types than other modalities, it is expected that certain combinations, such MRI for brain tumors and US for thyroid cancer, are more common. However, this distribution is also affected by the publicly available datasets.

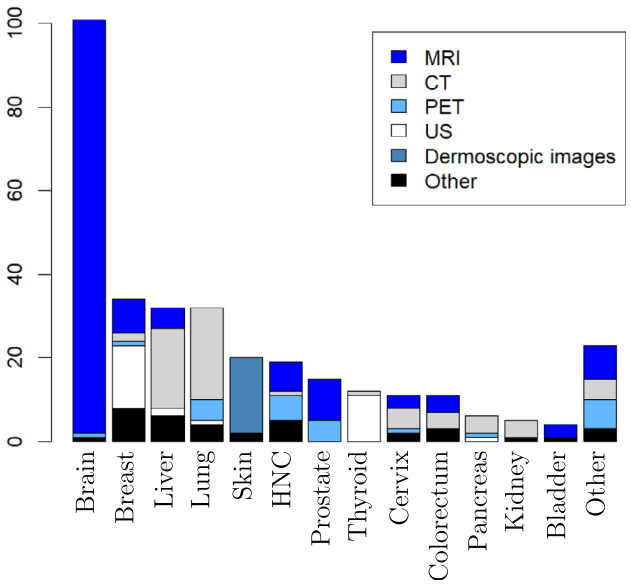

Fig. 3 The numbers of articles based on a specific imaging modality for different cancer types among the 325 selected tumor segmentation articles

5 Convolutional neural networks

A CNN is a special type of an artificial neural network (NN). In general, NNs are complex models consisting of several, even millions, of neurons divided on different layers and connected to each other in such a way that they can transmit instances of data from the input layer to the output layer. The connections between neurons have numerical values called weights and, while the layer-specific transformations are fixed by the chosen NN architecture, these weights are optimized during the training of the NN. Namely, a loss function is first used to estimate the difference between the ground-truth values and the outputs given by the NN with a specific choice of weights in the training data and an optimizer then determines how the weights should be updated in order to minimize this loss. The NN is typically trained with suitable data until the loss function converges. Given the training requires annotated data, the NNs are considered a form of supervised learning techniques, and they are typically also evaluated by using

a special test set of unseen data instances after the training (O'shea and Nash 2015).

CNNs are distinguished from other NNs due to their use of operations called convolutions to transform the data. On a convolutional layer, a filter matrix is placed over a receptive field of the same size in the input matrix, the dot product between the filter and the receptive field is computed, and the filter is then shifted by a stride along the input matrix onto another receptive field. A feature map is the series of these dot products in an order corresponding with the filter locations. The larger the stride is, the smaller the feature map is and, if several filters are used, the output will have an additional dimension to contain all the resulting feature maps. For dimension reduction, the CNNs have pooling layers on which a maximum or an average value of a receptive field is chosen for output instead of the dot product. Additionally, upsampling is possible through transposed convolutions, for instance. At the end of the CNN, there is typically also a fully-connected layer whose each neuron is directly connected with to every neuron of the previous layer (O'shea and Nash 2015).

The dimension of the data affects the CNN architecture and, in particular, the dimension of the filter. 2D CNNs can be re-designed for 3D data or even such data that has higher dimensions than this. However, due to the increased depth and complexity, 3D CNNs require more training data to avoid overfitting. If 3D medical image is given to a 2D CNN instead, they are typically divided into transaxial slices that are treated as separate images, which also means that one 3D image gives more data instances for training and test sets. Furthermore, it is possible to include information about the location of the slice in the 3D space or use some sort of a multi-view CNN that is given images of the same target from several directions. CNNs developed for processing color photographs use commonly three channels to process the RGB values of each pixel and, while medical images are typically grayscale, this channel structure can easily be used to input multimodality images instead.

U-Net introduced by Ronneberger et al. (2015) in 2015 is an example of a CNN commonly used for medical image segmentation. The idea behind the U-Net architecture is that a CNN needs to first see the whole image to understand the context before focusing on the details to perform accurate segmentation. The U-Net therefore consists of an encoder and a decoder. The encoder is a constricting path with some number of sequences, each of which contains two convolutional layers followed by a max-pooling layer. The decoder is correspondingly an expanding path with equivalently many sequences of two convolutional layers followed by an up-convolutional layer and there is a few additional convolutional layers at the end. Furthermore, before the ith max-pooling operation of the encoder for $i = 1, ..., n$, the data matrix is copied to be later combined with the output of the $(n - i)$th up-convolutional layer on the decoder. Between convolutions, U-Net uses the rectified linear unit defined as the function ReLU(x)=max$\{0, x\}$.

U-Net is an improvement of an earlier CNN called Fully convolutional network (FCN) by Long et al. (2015). FCN does not have fully-connected layers and, instead, it only performs convolution and downsampling or upsampling operations. U-Net has also been developed further: Its newer versions include UNet++ by Zhou et al. (2018), Attention U-Net by Oktay et al. (2018), ResU-Net by Diakogiannis et al. (2020), and nnU-Net by Isensee et al. (2021). In addition to U-Net and its modifications, other well-known CNNs for segmentation are DeepMedic by Kamnitsas et al. (2016), V-Net by Milletari et al. (2016), SegNet by Badrinarayanan et al. (2017), DeepLab by Chen et al. (2017), and RefineNet by Lin et al. (2017). Furthermore, AlexNet by Krizhevsky et al. (2012), VGG by Simonyan and Zisserman (2014), ResNet by He et al. (2016), DenseNet by Huang et al. (2017), and MobileNet by Howard et al. (2017) are also all very popular CNNs and, while they are perhaps more commonly used for classification, their architecture can be modified for segmentation.

While the common CNNs work very well for image segmentation and classification, they have limitations when it comes to object detection tasks. Girshick et al. (2015) proposed a region-based convolutional neural network (R-CNN) that first uses a selective search algorithm to choose such bounding boxes that might contain objects of interest, then gives these region proposals separately for a typical CNN architecture, and finally uses a support-vector machine (SVM) to classify the outputs. While R-CNNs are mainly for object detection, He et al. (2017) introduced a modified version called Mask R-CNN that uses the CNN architecture of an R-CNN to perform segmentation for each region proposal, enabling instance segmentation. Huang et al. (2019) developed this instance segmentation method further by introducing Mask-Scoring R-CNN that scores the segmentation mask rather than only evaluating it based whether the region proposal is correctly classified.

Given there are often limited amount of medical image data available to train the CNNs, different augmentation methods are common. While simple transformations such as reflections or rotations can be employed, one option is a generative adversarial network (GAN) which was introduced by Goodfellow et al. (2014) as a method to create synthetic data based on real images. The structure of GAN consists of both a generator producing new images and a discriminator evaluating how much they resemble the instances in the original dataset. Another possibility is a practice called transfer learning, which means that a CNN is first pre-trained to process similar data before it is trained to the actual task of interest.

In our systematic review synthesis, we were able to collect some details about the CNN from either the abstract or the full text-version for 207 articles. Out of these 207 articles, 22 articles (10.6%) used a standard 2D or 3D U-Net, 15 articles (7.2%) proposed either a well-known modified version of the U-Net such as ResU-Net or nnU-Net, and 77 articles (37.2%) introduced either a new modified version of the U-Net or a new CNN based on the U-Net architecture. In other words, at least 114 articles (55.1%) out of these 207 articles utilized the U-Net architecture in some way.

Conversely, the rest 93 articles (44.9%) used neither U-Net nor its modifications to perform the tumor segmentation task. Instead 22 articles (10.6%) used an unmodified existing CNN other than U-Net, 13 articles (6.3%) proposed a modified version for some existing CNN other than U-Net, and 58 articles (28.0%) introduced a new CNN without directly stating that it is a modification of some specific existing CNN. In the 22 articles proposing the use of an unmodified CNN other than U-Net, the most common CNNs were Mask R-CNN (in 6 articles), DeepLabV3 or DeepLabV3+ (in 4 articles), SegNet (in 2 articles), and DeepMedic (in 2 articles). DeepLabV3 and SegNet were also popular in the 13 articles proposing modified CNNs, though there were multiple times more U-Net based designs in the 207 articles.

6 Evaluation of segmentation results

The segmentation results of a CNN are evaluated by comparing the predicted tumor masks of the testing set and the related ground-truth tumor masks drawn by a medical doctor. In a binary mask, the pixels or the voxels are classified either positive if they contain cancer and otherwise negative, and these classifications are either true or false according to the ground-truth mask. The most common evaluation metric for segmentation, the Dice score, is defined as (Rainio et al. 2024)

$$D = \frac{2\mathrm{TP}}{2\mathrm{TP} + \mathrm{FN} + \mathrm{FP}},$$

where TP, FN, and FP refer to the numbers of true positive, false negative, and false positive classifications of the image points. The value of Dice score varies on the closed interval [0,1] so that the greater the value is the better the segmentation result is. Alternatively, some papers use the intersection over union (IoU), which is the number of points in the intersection of the positive segments in the predicted and the real mask divided by the number of points in the union of these positive segments. The Dice score can be computed from IoU with the formula $D = 2\mathrm{IoU}/(1 + \mathrm{IoU})$ (Rainio et al. 2024).

When summarizing our articles during systematic literature review, we specifically collected the reported Dice scores and, if only IoU was reported, we computed the Dice score from the IoU to be able to compare different studies with a single metric. Figure 4 shows the boxplots of the reported Dice scores for the most common cancer type and imaging modality combinations. As can be seen, the Dice scores are strongly affected by the cancer type and imaging modality combination: For brain tumors, the CNNs obtain Dice scores of around 90% whereas the segmentation accuracy of prostate cancer on MRI images is much lower, on average.

7 Literature review by cancer type

7.1 Brain tumors

The most common cancer type by far in the all 325 articles selected for systematic review synthesis was brain tumors, which were studied in 101 articles (31.1%). Despite the fact that this is relatively rare cancer type as shown in Fig. 2, nearly every third article introduced a method specially designed for brain tumors or at least trained and tested with brain tumor datasets. The reason for this is at least partially availability of brain tumor data: Since 2012, MICCAI has organised yearly Brain Tumor Segmentation (BraTS) challenges and provided MRI data collected from patients with brain tumors (Baid et al. 2021). Out of the 101 brain tumor

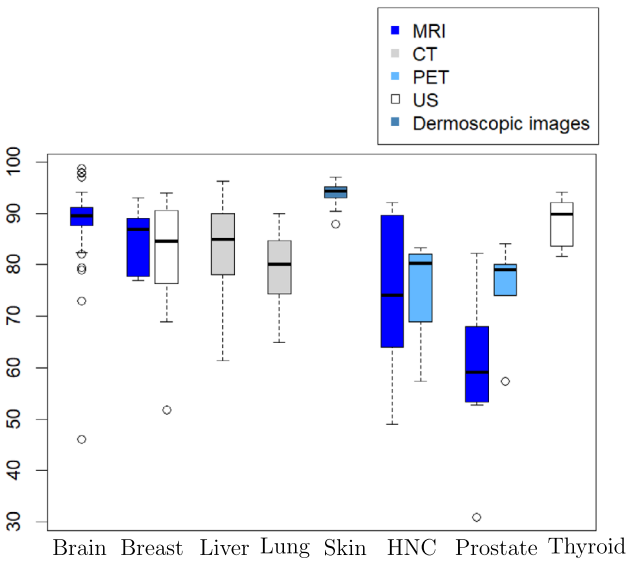

Fig. 4 Boxplots of the Dice scores (%) reported in the articles chosen for this review for the most common cancer and modality combinations. The Dice score for brain tumors is the Dice score of the whole tumor volume rather than the tumor core or the enhancing tumor

segmentation articles, 54 articles (53.5%) mentioned the use of BraTS data in the abstract.

According to Fig. 3, all the articles except two used MRI images, which would be likely the cause even without the BraTS datasets: MRI is the standard imaging modality brain tumors. From Fig. 4, we see that the reported Dice scores were very high in comparison with several other cancer types except skin cancer. This can be explained by both the large datasets provided by BraTS and the great soft-tissue separation ability of MRI. While the brain tumor segmentation is typically also performed for the tumor core and the enhancing tumor, we focus here on the whole tumor segmentation results so that the Dice scores are comparable with those from the other analyzed cancer types.

The highest Dice score for the brain tumor segmentation, 98%, was reported both by Gull et al. (2021) and Alqazzaz et al. (2022). Gull et al. proposed a simple 2D FCN after pre-processing images with median filtering. Alqazzaz et al. introduced a hybrid approach with a 2D SegNet to segment regions of interest followed by a decision tree to classify the pixels in these areas further. Other methods with a reported Dice score over 95% included a CNN combined with an optimization algorithm by Ranjbarzadeh et al. (2024) and a 3D U-Net by Lin et al. (2021). In the work of Lin et al., each input of the U-Net was a cube, obtained by transforming a ball-shaped selection of the brain area in the original images into a cube, and the voxel dimensions were decreased to lessen the computational complexity of the U-Net. The methods in all the four articles were evaluated on the BraTS data, though on datasets of different years, so the number of the patients in the used datasets varies from 477 to 660.

7.2 Breast cancer

Globally, breast cancer is the second most prevalent cancer after lung cancer and it was also very often studied in the selected tumor segmentation articles. Both DCE-MRI and US were common modalities among the selected articles as shown in Fig. 3 and there were also research about certain imaging techniques specially designed for breast cancer such as mammograms. The most commonly used public dataset in the selected articles was BUS-Set (Thomas et al. 2023) of US images.

The highest Dice score, 99.01%, was reported by Mohamed et al. (2022), but the full-text article revealed that they studied segmentation of the breast area only, not tumor segmentation. The best Dice score for actual breast tumor segmentation was 96.34% obtained by Alkhaleefah et al. (2022) for mammograms. Their method, called Connected-SegNets, was a combination of two 2D SegNets modified with skip connections. This method was evaluated by using both public and private datasets with a total of 1141 images of breast tumors. The highest Dice score for breast tumor segmentation in US images was reported by Yang et al. (2023), who used a 2D U-Net3+ with an attention mechanism originally introduced by Zhong (2022) for brain tumor segmentation in MRI images. The best Dice score for DCE-MRI-based breast cancer segmentation, 93%, was reported by Guo et al. (2022) who proposed a relatively simple 2D CNN combined with a SVM classifying the pixels in the output of the CNN as either positive or negative. Both Yang et al. and Guo et al. used private datasets collected from 1131 and 272 patients, respectively, at their own institutes.

One interesting approach was also introduced by Xing et al. (2020). Their aim was to combine the benefits of both R-CNNs that divide image into patches and FCNs that perform pixel-wise classification. Based the GAN architecture, the method used two generators to create both potential segmentation masks from original images and potential original images from generated masks and two discriminators evaluating their authenticity in a cyclic training process. The method reached a Dice score of 92% on a private dataset of 670 US images.

7.3 Liver cancer

Among the selected articles, the place of the third most common cancer type was shared by liver cancer and lung cancer, both with 32 articles, even though Fig. 2 shows that liver cancer has much lower incidence and mortality rate than lung cancer. Over a half (59.4%) of liver tumor segmentation articles was for CT images.

Three articles about CT-based liver tumor segmentation had a Dice score of at least 90% or higher: Wang et al. (2021); Gao and Almekkawy (2021), and Shao et al. (2024) reported Dice scores of 96.3%, 94.3%, and 90%, respectively. Additionally, all of them studied modified versions of U-Net: Wang et al. introduced a 3D U-Net modified by integrating octave convolutions to improve segmentation while Gao et al. combined atrous spatial pyramid pooling with U-Net++ architecture and Shao et al. introduced a 2D U-Net modified with axial attention and vision transformer modules. Wang et al. and Shao et al. both used the public Liver Tumor Segmentation (LiTS) (Bilic et al. 2023) dataset, though Shao et al. only chose 130 CT images of the LiTS training data rather than the whole data of 200 images and supplemented it with private data instead. Gao et al. used two private CT datasets and also one private US dataset.

Other modalities had lower segmentation accuracy: The highest Dice score for liver tumor segmentation on US images solely was 84% reported by Natali et al., who tested nnU-Net on a private dataset. For SPECT/CT images, Chaichana et al. (2021) proposed a new 3D CNN that performed segmentation by processing small 3D patches at a time. Their CNN architecture resembled a U-Net with additional dense blocks, and they obtained a Dice score of 85% on a

private dataset with 56 patients. Furthermore, for DCE-MRI images, Zheng et al. (2022) introduced a 3D U-Net with a convolutional long short-term memory for time information exploitation, resulting in a Dice score of 82.5% when trained and tested on a private dataset of 250 patients.

7.4 Lung cancer

Lung cancer has both the highest incidence and mortality of all cancers Global Cancer Observatory (2025). Out of the 32 selected lung cancer segmentation articles, 22 articles (68.8%) were based on CT images. Figure 4 suggests that the segmentation accuracy is slightly lower for lung cancer than for liver cancer on CT images.

The highest Dice score, 90%, was reported by Momin et al. (2021), who studied lung tumor segmentation with a 4D motion region CNN for 4D CT data obtained by recording typical static 3D CT images of 10 breathing phases. Otherwise, the CT-based lung cancer segmentation articles used typical static 3D CT images and the highest Dice scores for them were reported with 2D CNNs: Wang et al. (2022) introduced a new dual-path 2D CNN based on boundary enhancement and a hybrid transformer, Riaz et al. (2023) modified a 2D U-Net by using MobileNetV2 as its encoder, and Zhang et al. (2021) also improved a 2D U-Net but by using batch normalization and an α-hull algorithm for correcting the lung contour. They all obtained Dice scores between 86.2−89.7%, though they focused on the segmentation of pulmonary nodules, which are not necessarily cancerous. Wang et al. and Zhang et al. both tested their methods on 1010 images from the Lung Image Database Consortium and Image Database Resource Initiative (Armato III et al. 2011) whereas Riaz et al. used 96 images from the Cancer Imaging Archive (Clark et al. 2013) instead. For PET/CT images, the highest Dice score, 84%, was reported by Lei et al. (2023), who proposed a novel Mask R-CNN based approach and tested it on both a Cancer Imaging Archive dataset and a private one.

7.5 Skin cancer

Melanoma and non-melanoma skin cancer has high incidence rate but low mortality. All the 20 selected articles about skin cancer segmentation perform the segmentation task on either dermoscopic images or regular photographs. 15 (75.0%) of the articles used publicly available data from International Skin Imaging Collaboration (ISIC) (Codella et al. 2019).

From Fig. 4, we see that the reported Dice scores were high in all the skin cancer segmentation articles with very little variation between articles. A Dice score of at least 95% was obtained in four articles, which all proposed different novel CNN architectures: Bukhari et al. (2023) introduced a new CNN with multiple parallel depth-wise separable and dilated convolutions, Yang et al. (2021) modified to U-Net++ by replacing its encoder with Xception and EfficientNet architectures to create a new EfficientUNet++, Zheng et al. (2022) introduced a new L-shaped CNN, and Kaur et al. (2022) introduced a new CNN based on atrous convolutions. All of them tested their proposed CNNs on ISIC data and a few additional public datasets were used.

7.6 Head and neck cancer

HNC is a common name for a heterogeneous group of different cancer diseases originating from the human head and neck area, including for instance lips, oral cavity, larynx, and pharynx, but excluding brain tumors and eye cancer. Most of the selected HNC segmentation articles study segmentation of PET/CT or MRI images. There are public HNC PET/CT datasets available: Medical Image Computing and Computer-Assisted Intervention (MICCAI) organized Head and neck tumor (HECKTOR) challenge in 2020 and 2021, offering PET/CT image datasets with ground-truth HNC tumor masks drawn by physicians (Andrearczyk et al. 2021).

The highest Dice score for HNC segmentation, 92.06%, was reported by Wang et al. (2023), who tested DeepLabV3 for segmentation of nasopharynx gross tumor volume on a private dataset with MRI images of 200 patients. Wang et al. (2021) also studied segmentation of nasopharyngeal cancer on MRI images but they proposed a modified 2D Residual U-Net that obtained a Dice score of 89.6% on a private dataset of 45 patients. The highest Dice score for PET/CT data, 80.2%, was obtained by Zhao et al. (2023), who tested a 3D U-Net for the images of the 201 patients in the HECKTOR data of 2020.

7.7 Prostate cancer

Prostate cancer is very common cancer but often undiagnosed. As can be seen from Fig. 3, all our selection prostate cancer segmentation articles are based on either PET or MRI imaging. However, typical ^{18}F-FDG PET is rarely used for imaging of prostate cancer due to the low ^{18}F-FDG uptake of this cancer type. Instead, specially designed prostate specific membrane antigen (PSMA) is used imaging of prostate cancer patients.

The highest Dice score, 84%, was obtained by Kostyszyn et al. (2021), who used a standard 3D U-Net for segmentation of ^{68}Ga-PSMA and ^{18}F-PSMA PET images in several private datasets with a total of 209 patients. This method was also externally validated by Ghezzo et al. (2023), who reported a Dice score of 74% with the same method on a private dataset with ^{68}Ga-PSMA PET images of 85 patients. Additionally, Matkovic et al. (2021) studied segmentation of dominant intraprostatic lesions in PET/CT images based

on an unspecified PET tracer, obtaining a Dice score of 80.1% on a dataset of 49 patients. They proposed a cascaded method that first crops images based on the prostate segmentation by another CNN and then uses a Mask-Scoring R-CNN.

From Fig. 4, we can see that the MRI-based segmentation methods did not work as well as the PET-based in terms of the Dice scores on average. Still, the highest Dice score for MRI images was 82.2%: It was obtained with a standard 2D U-Net on a private dataset of 171 patients by Hong et al. (2023). Also for MRI images, Song et al. (2023) tested a modified 3D V-Net with a deep multi-scale attention mechanism, obtaining a Dice score of 70.14% on a dataset of 97 patients.

7.8 Thyroid cancer

While thyroid cancer is not especially common cancer and has low mortality, 12 of selected articles are related to it, most of them studying thyroid nodule segmentation on 2D US images. As shown in Fig. 4, the overall performance of the used methods is good with little variation between articles. The best Dice score, 94.1%, was obtained by Zheng et al. (2023), who introduced a 2D U-Net modified by adding atrous spatial pyramid pooling, a residual split-attention block, and a deformable convolution and tested it by using two private datasets with 5822 US images in total. U-Net-based methods were proposed by also other researchers: Gökmen Inan et al. (2024) modified a 2D ResUNet++ architecture by integrating feature extraction and upsampling with dropout operations while Lin et al. (2022) proposed a new 2D CNN with several atrous spatial pyramid pooling blocks, a multi-scale input layer, a U-Net type architecture but with attention blocks, and a parallel atrous convolution module. Inan et al. tested their method on US images of 880 patients and Lin et al. on a dataset of 1381 US images, obtaining Dice scores of 92.4% and 91.9%, respectively. A high Dice score, 90.26, was obtained also by Tao et al. (2022), who designed a fully new 2D CNN containing a transformers-based context-attention module, typical convolutions as the backbone, and a nodule adaptive convolution and tested it on 10781 US images.

7.9 Cervical cancer

Among the selected tumor segmentation articles, there were several articles studying cervical cancer, but all of them used private datasets. The highest Dice score, 87%, was reported by Wang et al. (2023), who used a 2D U-Net type CNN for CT slices of 60 patients to delineate their clinical target volume. Xiao et al. (2022) also studied cervical cancer segmentation in CT images but by using a 3D version of the RefineNet architecture, obtaining a Dice score of 82% on a dataset of 313 patients. For segmentation of MRI images, Zabihollahy et al. (2022) proposed a 3D U-Net after extracting suitable 3D blocks of interest with another CNN, Breto et al. (2022) tested a 2D Mask R-CNN, and Wang et al. (2021) introduced a new 3D CNN consisting of a few convolutions and residual units. They obtained Dice scores of 85%, 78%, and 67% on datasets of 125, 80, and 15 patients, respectively. Additionally, Iantsen et al. (2021) studied PET images and introduced a modified U-Net with new residual blocks and non-linear downsampling and upsampling blocks, which gave a Dice score of 80% on a dataset of 232 patients. The CNNs by both Wang et al. and Breto et al. performed organs-at-risk segmentation in addition to the cancer segmentation.

7.10 Colorectal cancer

Colorectal cancer is the third most common cancer disease in the world after lung and breast cancers according to Global Cancer Observatory (2025), but there were considerably fewer articles found about colorectal cancer segmentation. Most of the articles were also focused on the segmentation of rectal tumors specifically. The highest Dice score, 97%, was obtained by Wang et al. (2022), who used a modified U-Net with residual blocks instead of convolution blocks on a rectal cancer CT image data set provided by The Teddy Cup Data Mining Challenge. Dice scores over 90% for rectal tumor segmentation in CT images were also obtained by Li et al. (2023), who tested DeepLabV3+ on CT images of 47 rectal cancer patients, and Huiting (Zhang et al. 2023), who introduced a dual parallel network combining CNN and transformer architectures. For MRI images, Kenan Zhang et al. (2023) introduced a modified 2D U-Net with densely connected convolution blocks and additional atrous convolution blocks and reported a Dice score of 94.71% from the data of 572 patients.

7.11 Pancreatic cancer

The found articles about pancreatic cancer were mostly for CT images. Both Mukherjee et al. (2023) and Mahmoudi et al. (2022) studied segmentation of CT images by using a 3D nnU-Net and 3D Attention U-Nets, respectively. While Mukherjee et al. obtained a considerably higher Dice score than Mahmoudi et al. (84% vs 57.3%), Mukherjee et al. had a dataset of 1151 patients while Mahmoudi et al. had only 157 patients. For segmentation of PET/CT images, Wang et al. (2023) introduced a modified 3D U-Net with an added multi-modal fusion downsampling block and a multi-modal mutual calibration block, obtaining a Dice score of 76.20% from 93 PET/CT images.

7.12 Kidney cancer

For kidney tumor segmentation, the best Dice score, 93.9%, was obtained by Yang et al. (2022), who introduced a new 3D multi-scale residual FCN and tested it on the CT images of the public Kidney Tumor Segmentation (KiTS) Challenge Heller et al. (2020) and a private dataset. Ji et al. (2021) and Gottlich et al. (2023) also studied a similar combination of KiTS data and private data, though the private data used by Gottlich et al. contained both CT and MRI images and they focused on researching the impact of the amount of training data on segmentation accuracy. Ji et al. proposed a cascaded 3D U-Net and DeepLabV3+ using atrous convolution feature pyramid to adaptively control receptive field, reporting a Dice score of 81.72%, whereas Gottlich et al. obtained a Dice score of 84% with a standard 3D nnU-Net.

7.13 Bladder cancer

Varnyú and Szirmay-Kalos (2022) tested several well-known CNN architectures for bladder cancer segmentation on a private dataset of 2578 cystoscopy images and obtained the best Dice score, 91%, with a pyramid attention network originally introduced by Li et al. (2018) in 2018. Both Yu et al. (2022) and Moribata et al. (2023) proposed modified U-Net designs for MRI images: The model by Yu et al. had an additional cascade strategy to eliminate the redundant information in the background, and Moribata et al. had a slightly deeper version of U-Net with five sequences instead of the original four. Yu et al. obtained a Dice score of 87.40% from 1545 MRI images while Moribata et al. reported a Dice score of 81% by using MRI images of 170 patients.

8 Discussion

According to our analysis, the U-Net architecture was the most popular CNN design by far. Among different cancer types, the standard U-Net was especially popular within the articles about prostate cancer and HNC segmentation in PET images. Since PET imaging often results in blurry-looking images where tumors are visible as spots slightly darker or lighter than the background, it is quite expected that pixel-wise segmentation CNNs, such as U-Net, works well for them. As PET data is commonly used in the research of prostate and HNC tumor segmentation, partially due to the HECKTOR challenge, it is likely that the standard U-Net was often used for prostate cancer and HNC segmentation here because of its suitability for PET images rather some characteristic of these specific cancer type.

Additionally, over a third of the analyzed articles proposed a new method that either used the standard 2D or 3D U-Net with some additional modification or was introduced as a new CNN inspired by the U-Net architecture. In comparison, the use of well-known alternative versions of U-Net such as nnU-Net or ResU-Net was considerably less common in the tumor segmentation articles. Understandably, there is a pressure for the researchers to create new methods instead of testing the already existing ones so that their results have higher significance and easier to publish, though.

One of the most common modifications of U-Net was adding some sort of attention mechanism. The point of an attention mechanism is to allow the model focus on certain areas of the input rather than viewing all parts as equally meaningful. In our analyzed articles, the CNNs introduced by for instance Yang et al. (2023); Shao et al. (2024); Song et al. (2023), and Zheng et al. (2023) included attention mechanisms integrated in the U-Net architecture for segmentation of breast, liver, prostate, and thyroid tumors, respectively. The attention mechanism could potentially allow the CNN focus on the organ of interest to find the tumor within. Another common addition to U-Net was atrous spatial pyramid pooling used by Gao and Almekkawy (2021); Zheng et al. (2023), and Lin et al. (2022). This structure allows the CNN resample feature layers at multiple rates prior to convolution. Other U-Net modifications in our analyzed articles included for instance adding residual connections as in Zheng et al. (2023) and transformer layers as in Shao et al. (2024).

The Mask R-CNN based methods were rare for tumor segmentation, appearing as unmodified or modified in only 8 (3.9%) of the 207 articles analysed in terms of the CNN architecture. The cancer type varied much in these articles but, out of the possible imaging modalities, the Mask R-CNNs were mostly used for MRI images and, in case of skin cancer, dermoscopic images. Given the R-CNN based methods were originally developed for object detection in typical photographs, it is natural that they are better suited for MRI and dermoscopic images than for other modalities: MRI and dermoscopic images have sharp boundaries separating different tissue types and extracting specific targets from these images is therefore easy, especially in comparison with PET images.

Other well-known pixel-wise segmentation architectures besides U-Net and Mask R-CNN appeared a few times at most. Most commonly, DeepLabV3 or DeepLabV3+ and SegNet were used. DeepLabV3+ is a modification of the DeepLabV3 architecture based on using spatial pyramid pooling module and a simple but effective decoder to refine the segmentation accuracy on the object boundaries. DeepLabV3 produced higher Dice scores than U-Net in HNC segmentation in research by Yandan Wang et al. (2023) and DeepLabV3+ outperformed ResU-Net in rectal tumor segmentation in research by Li et al. (2023). SegNet is quite similar to U-Net but, unlike U-Net, SegNet only transfer

pooling indices from the encoder to the decoder, therefore requiring less computer memory. Despite the simple architecture, SegNet performed very well for breast cancer segmentation in Alkhaleefah et al. (2022).

However, one important aspect to note is that the results clearly depend on the used data. For instance, 2D SegNet outperformed 2D U-Net for breast tumor segmentation in 2D slices of 3D DCE-MRI images in the research by Carvalho et al. (2021) but 2D U-Net worked better than 2D SegNet for this same purpose in the research by Adoui et al. (2019). Because of this, it is difficult to say which method is actually better and the reported Dice scores are not directly comparable.

In fact, it is a common limitation present in the analyzed studies that the methods are only analyzed based on a single dataset. Still, it is understandable that there are strict restrictions about sharing imaging data about patients, especially if the data includes 3D tomography images of the head area, due to the threat to the patient anonymity (Parks and Monson 2017). Consequently, segmentation challenges such as HECKTOR or BraTS are crucial for fair comparison between methods because they allow the researchers to evaluate their CNNs by using same training and test datasets as their colleagues.

Another common issue with the analyzed articles was a lack of available code. We recommend authors sharing the code of their new methods. Namely, if it is not possible for different researchers to compare their methods on the same data, it would be important that the researchers could at least compare both their new method and the existing ones to their own data. Still, it was noticed during the review that the authors typically only compare their new methods with unmodified versions of the well-known CNNs such as U-Net and SegNet but not with new modifications introduced by other researchers studying the same tumor segmentation task. The comparison with other CNNs specifically modified for tumor segmentation is naturally only possible if these CNNs are accessible and it was noted during the analysis that relatively few articles included publicly available code for the methods.

During the future research, several topics could be considered. There is a quite little research on certain common cancer types, including gastric cancer, ovarian cancer, and testicular cancer. Furthermore, for several cancer types, automatic tumor segmentation applications have only been developed for a single imaging modality, even when the cancer is commonly imaged also with other modalities. In particular, only one article based on SPECT imaging was found so to understand better the potential of automatic tumor segmentation from SPECT images, more research would be needed.

In this review, there are also certain limitations that could be addressed with more study. First of all, we only concentrated on the CNN-based automatic tumor segmentation techniques so new methods, such as visual transformers, could be also reviewed to obtain a more comprehensive picture of the current state of automatic tumor segmentation. Due to our identification of the original articles based on a PudMed search only, there is a risk of bias caused by potentially missing more technical articles not indexed in PubMed. Furthermore, there might be well-performing techniques introduced in studies published prior to our time period of 1.4.2021−31.3.2024, and more research would be needed to see any trends in the accuracy of the automatic methods with respect of time.

9 Conclusion

We analyzed 325 articles about CNN-based tumor segmentation for different cancer types. The most commonly studied cancer type was brain tumors, followed by breast, liver, lung, skin, HNC, prostate, thyroid, cervical, colorectal, pancreatic, kidney, and bladder cancer. Excluding skin cancer segmentation nearly exclusively studied for dermoscopic images or regular photographs, we found mostly articles using MRI, CT, PET, or US images. Based on our analysis, the most common CNN for tumor segmentation is U-Net and its modifications. The R-CNN based methods are rarely used, especially outside MRI images. Much of the current research is solely focused on MRI images of brain tumors, despite the fact that this cancer is not so common. In comparison, there is relatively little research about tumor segmentation for colorectal cancer, which has much higher global incidence and mortality rate than brain tumors. The CNNs for segmentation of brain tumor, skin cancer, and thyroid nodules also outperformed the ones for other cancer types in terms of segmentation accuracy measured by the Dice score. Several methods could yet be tested for certain cancer types in the future research and also for rarer imaging modalities such as SPECT.

Funding Open Access funding provided by University of Turku (including Turku University Central Hospital). The first author was financially supported by the Finnish Cultural Foundation and Sakari Alhopuro Foundation.

Data availability An excel file summarizing the 810 PubMed search results is included as supplementary material.

Code availability Not applicable.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no Conflict of interest.

References

Alkhaleefah M et al (2022) Connected-SegNets: a deep learning model for breast tumor segmentation from X-ray Images. Cancers 14(16):4030

Alqazzaz S et al (2022) Combined features in region of interest for brain tumor segmentation. J Digit Imaging 35(4):938–946

Andrearczyk V et al (2021)"Overview of the HECKTOR challenge at MICCAI 2021: automatic head and neck tumor segmentation and outcome prediction in PET/CT images". In: 3D head and neck tumor segmentation in PET/CT challenge. Springer,, pp. 1–37

Anwar SM et al (2018) Medical image analysis using convolutional neural networks: a review. J Med Syst 42:1–13

Armato SG III et al (2011) The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans. Med Phys 38(2):915–931

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

Baid U et al (2021)"The rsna-asnr-miccai brats 2021 benchmark on brain tumor segmentation and radiogenomic classification". In: arXiv preprint arXiv:2107.02314

Berman Rose M et al (2016) DCE MRI of prostate cancer. Abdom Radiol 41:844–853

Bilic P et al (2023) The liver tumor segmentation benchmark (lits). Med Image Anal 84:102680

Breto AL et al (2022) Deep learning for per-fraction automatic segmentation of gross tumor volume (GTV) and organs at risk (OARs) in adaptive radiotherapy of cervical cancer. Front Oncol 12:854349

Bukhari M et al (2023) (2023) A novel framework for melanoma lesion segmentation using multiparallel depthwise separable and dilated convolutions with swish activations'. J Healthc Eng 1:1847115

Carvalho ED et al (2021) "Tumor segmentation in breast DCE-MRI slice using deep learning methods". 2021 IEEE symposium on computers and communications (ISCC). IEEE. pp. 1–6

Chaichana A et al (2021) Automated segmentation of lung, liver, and liver tumors from Tc-99m MAA SPECT/CT images for Y-90 radioembolization using convolutional neural networks. Med Phys 48(12):7877–7890

Chen L-C et al (2017) Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. IEEE Trans Pattern Anal Mach Intell 40.4:834–848

Clark K et al (2013) The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. J Digit Imaging 26:1045–1057

Codella N et al (2019) Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (isic). In: arXiv preprint arXiv:1902.03368

Debela DT et al (2021) New approaches and procedures for cancer treatment: Current perspectives. SAGE Open Med 9:20503121211034370

Dhiyanesh B, Rajkumar S, Radha R et al (2021)"Improved object detection in video surveillance using deep convolutional neural network learning". In: 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC). IEEE. , pp. 1–8

Diakogiannis FI et al (2020) ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data. ISPRS J Photogramm Remote Sens 162:94–114

El Adoui M et al (2019) MRI breast tumor segmentation using different encoder and decoder CNN architectures. Computers 8(3):52

Farag W (2018) Recognition of traffic signs by convolutional neural nets for self-driving vehicles. Int Knowl-based Intell Eng Syst 22(3):205–214

Gao Q, Almekkawy M (2021) ASU-Net++: A nested U-Net with adaptive feature extractions for liver tumor segmentation. Comput Biol Med 136:104688

Ghezzo S et al (2023) External validation of a convolutional neural network for the automatic segmentation of intraprostatic tumor lesions on 68Ga-PSMA PET images. Front Med 10:1133269

Girshick R et al (2015) Region-based convolutional networks for accurate object detection and segmentation. IEEE Trans Pattern Anal Mach Intell 38.1:142–158

Global Cancer Observatory. Cancer Today. https://gco.iarc.fr (2025)

Goodfellow I et al (2014)"Generative adversarial nets". Adv Neural Inform Process Syst 27

Gottlich HC et al (2023) Effect of dataset size and medical image modality on convolutional neural network model performance for automated segmentation: A CT and MR renal tumor imaging study. J Digit Imaging 36(4):1770–1781

Gull S, Akbar S, Khan HU et al (2021) Automated detection of brain tumor through magnetic resonance images using convolutional neural network". BioMed Res Int 2021(1):3365043

Guo Y-Y et al (2022) Breast MRI tumor automatic segmentation and triple-negative breast cancer discrimination algorithm based on deep learning. Comput Math Methods Med 202(1):2541358

He K et al (2016)"Deep residual learning for image recognition". In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778

He K et al (2017) Mask r-cnn. Proceed IEEE Int Conf Comput Vis 2017:2961–2969

Heller N et al (2020) An international challenge to use artificial intelligence to define the state-of-the-art in kidney and kidney tumor segmentation in CT imaging

Hong S et al (2023) Deep learning algorithm for tumor segmentation and discrimination of clinically significant cancer in patients with prostate cancer. Curr Oncol 30(8):7275–7285

Howard AG et al "Mobilenets: Efficient convolutional neural networks for mobile vision applications". In: arXiv preprint arXiv:1704.04861 (2017)

Huang G et al (2017)"Densely connected convolutional networks". In: Proceedings of the IEEE conference on computer vision and pattern recognition 4700–4708

Huang Z et al (2019)"Mask scoring r-cnn". In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 6409–6418

Iantsen A et al (2021) Convolutional neural networks for PET functional volume fully automatic segmentation: development and validation in a multi-center setting. Eur J Nucl Med Mol Imaging 48:3444–3456

Isensee F et al (2021) nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nat Methods 18.2:203–211

Ji H et al (2021) Automatic segmentation of kidney tumor based on cascaded multiscale convolutional neural networks. J Biomed Eng 38(4):722–731

Jiang H, Diao Z, Yao Y-D (2022) Deep learning techniques for tumor segmentation: a review. J Supercomput 78.2:1807–1851

Kamnitsas K et al (2016) "DeepMedic for brain tumor segmentation". In: *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: Second International Workshop, BrainLes 2016, with the Challenges on BRATS, ISLES and mTOP 2016, Held in Conjunction with MICCAI 2016, Athens, Greece, October 17, 2016, Revised Selected Papers 2*. Springer. pp. 138–149

Kaur R et al (2022) Automatic lesion segmentation using atrous convolutional deep neural networks in dermoscopic skin cancer images. BMC Med Imaging 22(1):103

Kostyszyn D et al (2021) Intraprostatic tumor segmentation on PSMA PET images in patients with primary prostate cancer with a convolutional neural network. J Nucl Med 62(6):823–828

Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. Adv Neural Inform Process Syst 25

Lawrence S et al (1997) Face recognition: a convolutional neural-network approach. IEEE Trans Neural Netw 8.1:98–113

Lei Y et al (2023) Automated lung tumor delineation on positron emission tomography/computed tomography via a hybrid regional network. Med Phys 50(1):274–283

Li H et al (2018)"Pyramid attention network for semantic segmentation". In: arXiv preprint arXiv:1805.10180

Li Z et al (2021) "A survey of convolutional neural networks: analysis, applications, and prospects". In: IEEE transactions on neural networks and learning systems

Li J et al (2023) Clinical evaluation on automatic segmentation results of convolutional neural networks in rectal cancer radiotherapy. Front Oncol 13:1158315

Lin G et al (2017)"Refinenet: Multi-path refinement networks for high-resolution semantic segmentation". In: Proceedings of the IEEE conference on computer vision and pattern recognition pp. 1925–1934

Lin W-W et al (2021) 3D brain tumor segmentation using a two-stage optimal mass transport algorithm. Sci Rep 11(1):14686

Lin X et al (2022) A super-resolution guided network for improving automated thyroid nodule segmentation. Comput Methods Programs Biomed 227:107186

Long J, Shelhamer E, Darrell T (2015) "Fully convolutional networks for semantic segmentation". In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 3431–3440

Mahmoudi T et al (2022) Segmentation of pancreatic ductal adenocarcinoma (PDAC) and surrounding vessels in CT images using deep convolutional neural networks and texture descriptors. Sci Rep 12(1):3092

Matkovic LA et al (2021) Prostate and dominant intraprostatic lesion segmentation on PET/CT using cascaded regional-net. Phys Med Biol 66.24:245006

Milletari F, Navab N, Ahmadi S-A (2016) "V-net: Fully convolutional neural networks for volumetric medical image segmentation". In: 2016 fourth international conference on 3D vision (3DV). IEEE. pp. 565–571

Mohamed EA et al (2022) A Novel CNN pooling layer for breast cancer segmentation and classification from thermograms. PLoS ONE 17(10):e0276523

Momin Shadab et al (2021) Lung tumor segmentation in 4D CT images using motion convolutional neural networks. Med Phys 48(11):7141–7153

Moribata Y et al (2023) Automatic segmentation of bladder cancer on MRI using a convolutional neural network and reproducibility of radiomics features: A two-center study. Sci Rep 13(1):628

Mukherjee S et al (2023) Bounding box-based 3D AI model for user-guided volumetric segmentation of pancreatic ductal adenocarcinoma on standard-of-care CTs. Pancreatology 23(5):522–529

ökmen Inan NG et al (2024) Multi-class classification of thyroid nodules from automatic segmented ultrasound images: Hybrid ResNet based UNet convolutional neural network approach. Comput Methods Programs Biomed 243:107921

Oktay O et al (2018)"Attention u-net: Learning where to look for the pancreas". In: arXiv preprint arXiv:1804.03999

O'shea K, Nash R (2015) "An introduction to convolutional neural networks". In: arXiv preprint arXiv:1511.08458

Page MJ et al (2021) "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews".BMJ 372

Parks CL, Monson KL (2017) Automated facial recognition of computed tomography-derived facial images: patient privacy implications. J Digit Imaging 30(2):204–214

Rainio O, Teuho J (2024) Evaluation metrics and statistical tests for machine learning. Sci Rep 14.1:6086

Ranjbarzadeh R et al (2024) Brain tumor segmentation based on optimized convolutional neural network and improved chimp optimization algorithm. Comput Biol Med 168:107723

Riaz Z et al (2023) Lung tumor image segmentation from computer tomography images using MobileNetV2 and transfer learning. Bioengineering 10(8):981

Ronneberger O, Fischer P, Brox T (2015) "U-net: Convolutional networks for biomedical image segmentation". In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, Proceedings, Part III 18. Springer. pp. 234–241

Shao J et al (2024) Attention connect network for liver tumor segmentation from CT and MRI images. Tech Cancer Res Treat 23:15330338231219366

Simonyan K, Zisserman A (2014)"Very deep convolutional networks for large-scale image recognition". In: arXiv preprint arXiv:1409.1556

Song E et al (2023) Prostate lesion segmentation based on a 3D end-to-end convolution neural network with deep multi-scale attention. Magn Reson Imaging 99:98–109

Tao Z et al (2022) Local and context-attention adaptive LCA-Net for thyroid nodule segmentation in ultrasound images. Sensors 22(16):5984

Thomas C et al (2023) BUS-Set: a benchmark for quantitative evaluation of breast ultrasound segmentation networks with public datasets. Med Phys 50(5):3223–3243

Townsend DW et al (2004) Physical principles and technology of clinical PET imaging. Annal-Acad Med Singa 33(2):133–145

Varnyú D, Szirmay-Kalos L (2022) A comparative study of deep neural networks for real-time semantic segmentation during the transurethral resection of bladder tumors. Diagnostics 12(11):2849

Wang B et al (2021) Accurate tumor segmentation via octave convolution neural network. Front Med 8:653913

Wang D et al (2021) Convolutional neural network intelligent segmentation algorithm-based magnetic resonance imaging in diagnosis of nasopharyngeal carcinoma foci. Contrast Media Mol Imaging 1:2033806

Wang B et al (2021) Multimodal MRI analysis of cervical cancer on the basis of artificial intelligence algorithm. Contrast Med & Mol Imaging 1:1673490

Wang Sihui et al (2022) DPBET: A dual-path lung nodules segmentation model based on boundary enhancement and hybrid transformer. Comput Biol Med 151:106330

Wang H et al (2022) A computed tomography image segmentation algorithm for improving the diagnostic accuracy of rectal cancer based on U-net and residual block. J Biomed Eng 39.1:166–174

Wang Y et al (2023) Automatic detection and recognition of nasopharynx gross tumour volume (GTVnx) by deep learning for

nasopharyngeal cancer radiotherapy through magnetic resonance imaging. Radiat Oncol 18(1):76

Wang Jiahao et al (2023) Evaluation of auto-segmentation for brachytherapy of postoperative cervical cancer using deep learning-based workflow. Phys Med Biol 68(5):055012

Wang F et al (2023) MFCNet: A multi-modal fusion and calibration networks for 3D pancreas tumor segmentation on PET-CT images. Comput Biol Med 155:106657

Xiao C et al (2022) RefineNet-based 2D and 3D automatic segmentations for clinical target volume and organs at risks for patients with cervical cancer in postoperative radiotherapy. J Appl Clin Med Phys 23(7):e13631

Xing J et al (2020) Lesion segmentation in ultrasound using semi-pixel-wise cycle generative adversarial nets. IEEE/ACM Trans Comput Biol Bioinf 18(6):2555–2565

Yang C-H et al (2021) Deep hybrid convolutional neural network for segmentation of melanoma skin lesion. Comput Intell Neurosci 2021:1–15

Yang E et al (2022) 3D multi-scale residual fully convolutional neural network for segmentation of extremely large-sized kidney tumor. Comput Methods Programs Biomed 215:106616

Yang Lei et al (2023) Rapid segmentation and diagnosis of breast tumor ultrasound images at the sonographer level using deep learning. Bioengineering 10(10):1220

Yu J et al (2022) Cascade Path Augmentation Unet for bladder cancer segmentation in MRI. Med Phys 49(7):4622–4631

Zabihollahy F et al (2022) Fully automated segmentation of clinical target volume in cervical cancer from magnetic resonance imaging with convolutional neural network. J Appl Clin Med Phys 23(9):e13725

Zhang W et al (1994) Computerized detection of clustered microcalcifications in digital mammograms using a shift-invariant artificial neural network. Med Phys 21(4):517–524

Zhang X et al (2021) Accurate segmentation for different types of lung nodules on CT images using improved U-Net convolutional network. Medicine 100:40

Zhang H et al (2023) Dual parallel net: A novel deep learning model for rectal tumor segmentation via CNN and transformer with Gaussian Mixture prior. J Biomed Inform 139:104304

Zhang K et al (2023) Automatic segmentation of rectal tumors from MRI using multiscale densely connected convolutional neural network based on attention mechanism. Phys Med Biol 68(16):165001

Zhao L-m et al (2023) Head and neck tumor segmentation convolutional neural network robust to missing PET/CT modalities using channel dropout. Phys Med Biol 68(9):095011

Zheng R et al (2022) Automatic liver tumor segmentation on dynamic contrast enhanced mri using 4D information: Deep learning model based on 3D convolution and convolutional lstm. IEEE Trans Med Imaging 41(10):2965–2976

Zheng T et al (2023) Segmentation of thyroid glands and nodules in ultrasound images using the improved U-Net architecture. BMC Med Imaging 23(1):56

Zhong Y (2022) "Attention mechanisms with UNet3+ on brain tumor MRI segmentation". In: AIP Conference Proceedings. 2589. 1. AIP Publishing

Zhou Z et al (2018)"Unet++: A nested u-net architecture for medical image segmentation". In: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4. Springer. 3–11