Review of Deep Learning in Pediatric Oncology Radiology

Aart van Bochove, Utrecht University
July 8, 2022

***Abstract*— *Introduction.* Deep learning has seen many applications in oncology imaging, but these applications are mainly limited to the adult population. They can not be blindly applied to children, due to large differences in anatomy between children and adults. In the last few years, more and more research has been done into deep learning applications for pediatric oncology. This article provides a review of deep learning applications in the pediatric oncology radiology field. *Results.* 21 relevant papers were found. Most of them ($n = 10$) describe classification models, with other models perform segmentation ($n = 5$), image synthesis ($n = 4$) and image quality improvement ($n = 2$) tasks. For most papers, convolutional neural networks were used, except for five papers on tumor classification. MRI images were the most used input for the models. *Conclusion & Discussion.* Research on deep learning applications in pediatric oncology radiology is following the developments in deep learning for adult oncology, resulting in a rapid increase of research into this topic. In the future, models will most likely become more complex and deep, which can be supported by larger available datasets.**

I. Introduction

AUTOMATED medical image analysis has been around for several decades. Initially, this consisted of manually created algorithms, which used humanly understandable image features to do classification, segmentation, etc. Soon machine learning methods emerged, which were able to find optimal solutions based on training data [1]. An advanced machine learning method is deep learning, which has seen rapid developments recently. Deep learning makes use of artificial neural networks, which take images, image features, or other information as input. This data is fed through multiple layers of artificial neurons, towards an output layer. This output can for example be the segmentation of certain organs or tumors, the classification of tumor type or patient outcome, or even complete synthetic images. An important advantage of using deep learning over classical machine learning methods is that a neural network can construct features itself. These features can be very complicated, such that they cannot be manually constructed by humans. A downside of this is that the network behaves like a 'black box', it is very hard to determine how the model reached its conclusion [2]. Another challenge of neural networks is that they have a huge number of variables that need to be trained, which makes them very prone to overfitting to the training data [3]. However, several methods exist to counter this, such as data augmentation, allowing the use of deep learning with small datasets.

Deep learning has already seen many applications in oncology. However, deep learning has not been used as much in pediatric oncology radiology, possibly because of the rarity of large imaging datasets in pediatric oncology [4]. Methods that are developed for adults cannot necessarily be applied to children, because the anatomy of children is different compared to adults. For example, development in the brain differs between tissue types and is non-linear, resulting in a different brain anatomy in children compared to adults [5]. This paper creates an overview of deep learning applications in pediatric oncology radiology. Some review papers have made an overview of machine learning in pediatric oncology imaging, such as Daldrup-Link [3] and Ramesh *et al.* [6], as well as Huang *et al.* [7] with a review focused specifically on machine learning applications in pediatric brain tumor imaging. This paper is different compared to the other papers because it focuses on deep learning specifically. This allows for a more in-depth look into the different deep learning techniques, such as the used model architectures. Furthermore, research in this field is increasing rapidly, resulting in 12 recently ($\geq$2020) published relevant papers which are not yet discussed in the previously mentioned review papers, compared to 9 papers that are discussed there.

To find relevant papers, in June 2022, a SCOPUS search based on the following query was performed: "(cancer OR oncology) AND (children OR pediatric) AND (deep learning OR neural network OR artificial intelligence)". The ASReview toolbox [8] was used to select relevant papers from the query results. ASReview sorts the papers in order of relevance based on user input, using active learning. In this stage, only the title and abstract were used. Papers cited in the relevant articles were also considered for inclusion. Papers were included if their dataset included only pediatric or young adult oncology patients. Papers before 2012 were not included, due to the rapid development of deep learning technology. Only papers discussing deep learning methods were included, where deep learning is defined as being a method using a neural network with at least two hidden layers. Only papers discussing radiology are included, so papers on histology were excluded.

II. Results

21 relevant papers were found, as is shown in Fig. 1. Fig. 2 shows the distribution of the papers. They were about tumors in the nervous system ($n = 17$), including brain tumors ($n = 12$), about kidney tumors ($n = 2$), tumors in the lymphatic system ($n = 2$), and tumors in soft tissue in general ($n = 1$). The models had a classification ($n = 10$),

TABLE I: Overview of papers discussed in this review paper. Abbreviations: AUC = Area under ROC curve, BAR = Balanced accuracy rate, CNN = Convolutional neural network with any other architecture than AlexNet, U-Net or ResNet, D_{diff} = Dose difference, DiCon = diagnostic confidence, DSC = Dice score, FNN = Fully connected neural network, ICC = Intraclass correlation coefficients between predicted and manual segmentations, IoU = Intersection over union, LOOCV = Leave one out cross validation, MAE = Mean average error, MIBG = metaiodobenzylguanidine, NS = Not specified, RPSP = relative proton stopping power, T_{MAX}/B = tumor maximum activity/background

(a) Overview of papers describing a classification model.

Author	Tumor location	Imaging modality	Metric(s)	Metric performance	Model type	Cross-validation	# Patients	Patient age
Orphanidou-Vlachou *et al.* (2013) [9]	Brain	MRI	Accuracy	0.933	FNN	LOOCV	40	NS
Zarinabad *et al.* (2016) [10]	Brain	MRI	BAR	0.82 & 0.92	FNN	10-fold CV	90	6.86 $\pm$ 4.22
Li *et al.* (2019) [11]	Brain	MRI	Accuracy	0.8058	FNN	Repeated hold-out validation (70-30)	58	0-14
Zhang *et al.* (2021) [12]	Brain	MRI	F1	0.9189	FNN	No	278	$<$19
Quon *et al.* (2020) [13]	Brain	MRI	Accuracy	0.92	ResNet	No	617	0-34 (median 8)
Prince *et al.* (2020) [14]	Brain	CT, MRI	Accuracy	0.878	ResNet	5-fold CV	86	NS
Ye *et al.* (2021) [15]	Brain	MRI	AUC	0.950 - 0.991	FNN	No	9	7-18
Banerjee *et al.* (2018) [16]	Soft tissue	MRI	Accuracy	0.85	AlexNet	LOOCV	21	1-20
Mayampurath *et al.* (2021) [17]	Nervous system	MIBG	AUC	0.63	CNN	4-fold CV	103	0-18
Liu *et al.* (2022) [18]	Nervous system	CT	AUC	0.63 - 0.83	FNN & CNN	5-fold outer, 3-fold internal CV	65	0-16

(b) Overview of papers describing a segmentation model.

Author	Tumor location	Imaging modality	Metric(s)	Metric performance	Model type	Cross-validation	# Patients	Patient age
Peng *et al.* (2020) [5]	Brain	MRI	ICC	0.912 & 0.960	U-Net	No	916	$<$19
Artzi *et al.* (2020) [19]	Nervous system	MRI	DSC	0.761	U-Net + ResNet	5-fold CV	29	5.7 $\pm$ 5.4
Nalepa *et al.* (2022) [20]	Nervous system	MRI	DSC	0.781	U-Net	4-fold CV	22 & 51	Mean: 7.5 & 9
Corbat *et al.* (2020) [21]	Kidney	CT	DSC	0.8806	CNN	$OV^2ASSION$	14	1-15
Yin *et al.* (2021) [22]	Lymphatic system	CT	DSC	0.9	CNN	NS	30	0-9

(c) Overview of papers describing an image synthesis model.

Author	Tumor location	Imaging modality	Metric(s)	Metric performance	Model type	Cross-validation	# Patients	Patient age
Florkow *et al.* (2020) [23]	Kidney/Nervous system	MRI	MAE, D_{diff}	57 HU, $<$0.5%	U-Net	3-fold CV	66	1-9
Maspero *et al.* (2020) [24]	Brain	MRI	MAE, D_{diff}	61 HU, -0.1% & 0.1%	GAN	No	60	10 $\pm$ 5
Wang *et al.* (2022) [25]	Brain	MRI	MAE	42 HU	GAN	No	195	1-20
Wang *et al.* (2021) [26]	Lymphatic system	PET, MRI	DiCon	0.942	ResNet	LOOCV	23	3-30

(d) Overview of papers describing models for image quality improvement.

Author	Tumor location	Imaging modality	Metric(s)	Metric performance	Model type	Cross-validation	# Patients	Patient age
Hales *et al.* (2020) [27]	Brain	MRI	SNR gain	62%	CNN	No	131	0-17
Ladefoged *et al.* (2019) [28]	Brain	PET, MRI	T_{MAX}/B	-0.1%	U-Net	4-fold CV	79	0-14

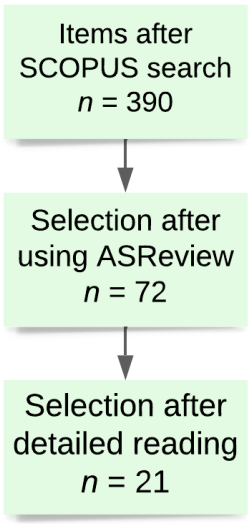


Fig. 1: Flow diagram showing the exclusion process.

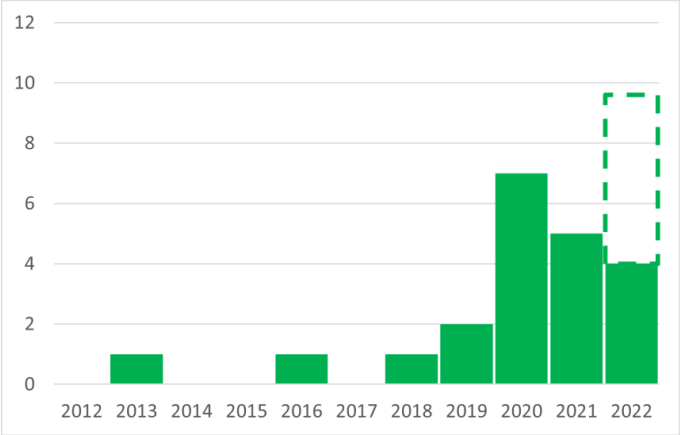
segmentation ($n = 5$), synthetic image generation ($n = 4$) or image improvement ($n = 2$) task. The mostly used modality was magnetic resonance imaging (MRI) ($n = 17$), followed with computed tomography (CT) ($n = 4$), positron emission tomography (PET) ($n = 2$), and metaiodobenzylguanidine imaging (MIBG) ($n = 1$). More information on the papers can be found in Table I.

A. Classification

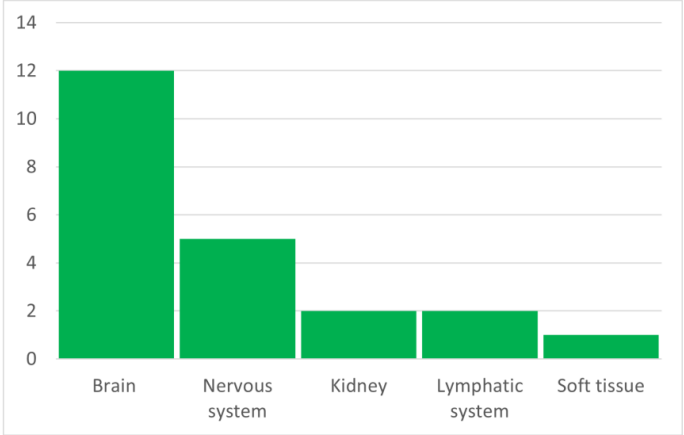
Classification is the task that currently has the most applications in deep learning for pediatric oncology ($n = 10$). This can involve classification into different tumor (sub)types, but also patient outcome prediction, both of which are discussed below.

1) Tumor classification: Seven out of eight papers on tumor classification deal with brain tumors, while the last paper describes a model classifying rhabdomyosarcoma tumors. All papers but one use solely MRI, because MRI is considered the standard technique for brain imaging due to its high soft-tissue contrast [29]. In one paper, CT images are also used in the model.

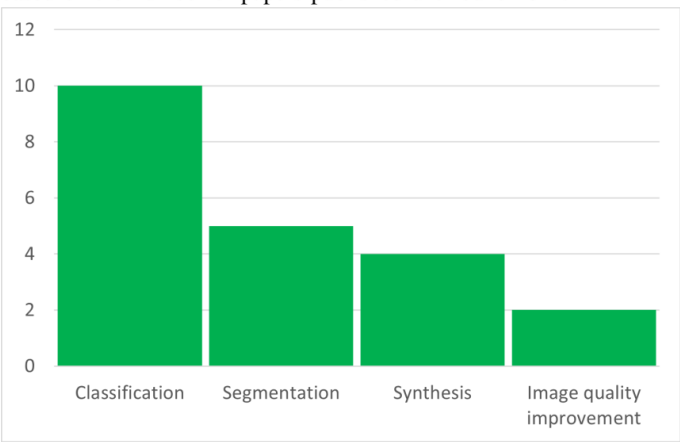
About 50-55% of the brain tumors in children are located in the posterior fossa (PF) [30]. Classification of these tumors is discussed in five papers. Four of them used features extracted from MR images as an input for the deep learning models, while one used the full input image in a convolutional neural network (CNN). An important advantage of using image features as input over simply using the images is that this makes the model less of a black box: the relevant features can be understood intuitively. All papers compare different machine learning methods. In 2013, Orphanidou-Vlachou *et al.* [9] applied a texture analysis algorithm to T1 weighted (T1w) and T2 weighted (T2w) MRI to extract 279 features. Principal component analysis yielded principal components which were processed using linear discriminant analysis (LDA) and a probabilistic neural network (PNN). A probabilistic neural network is a specific type of feedforward neural network. This implementation consisted of two hidden layers, three summation layers, and one output layer. The PNN outperformed the LDA. This was the first indication that using neural networks might be beneficial in pediatric brain tumor classification. A challenge with the classification of PF tumors is the different incidences of the different tumor types. Zarinabad *et al.* [10] describe a dataset where only 11% of the patients had ependymomas, with 89% of the cases having one of two other PF tumor types. This imbalanced dataset can result in low test accuracies for the classification of less frequent tumors. Zarinabad *et al.* tried to tackle this problem by overpopulating the minority class using the bSMOTE algorithm, which is an algorithm that creates artificial data. Different classification methods were tested, including a neural network with 3 hidden layers. The input for this model included features from magnetic resonance spectroscopy (MRS) data. The models that were trained using the expanded dataset performed better than the models using only the original dataset. Li *et al.* [11] was less successful in using deep learning for PF tumor classification. They used 10 different methods to extract features from MRI images, after which they used 11 different machine learning methods for classification, including a neural network. They assessed the accuracy, as well as the stability of the different classifiers. The neural network had an accuracy of 80.58%, while the best performing method (a support vector machine) had an accuracy of 85.38%. The neural network had the lowest stability of all methods. However, Zhang *et al.* [12] reached better results, by using 6 machine learning methods, including a neural network. First, a classifier was used to identify the appearance of pilocytic astrocytoma, and if this was a negative, a second classifier distinguished between ependymoma and medulloblastoma. As input for these classifiers, features from T2w- and contrast-enhanced T1w images were used. In the first step, the best performing classifier was a logistic regression, while the neural net performed the best in the second step. A paper by Quon *et al.* [13] was the only paper that used entire MRI images as input for their model, a ResNeXt convolutional neural network. A ResNeXt model is based on the ResNet architecture, but introduces a new model dimension apart from width and depth, called cardinality (C). For each ResNet block, the ResNeXt model has C parallel blocks, such that there are C paths from the start to the end of the block. They used a 2D-based model, because of a wide variety of slice thicknesses in the dataset. The model was pre-trained on the ImageNet dataset. It classified each slice as having a tumor or not a tumor, and further classified the slices with a tumor into 4 different tumor types. They used a confidence-weighted vote of models with 5 hyperparameter combinations to do final predictions. They compared using T2w, contrast-enhanced T1w, and ADC MRI as an input to using T2 only. Using T2w only performed better, probably because there was overfitting when using all images as input. The dataset used in this paper includes patient ages up to 34, but it is still included in this review because the mean age of the patients was 8. all scans were made in a pediatric hospital, suggesting that the



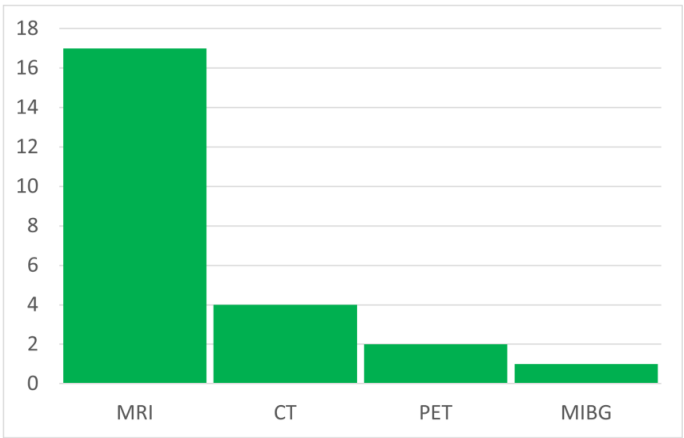
(a) The number of papers per year. The dashed box is an extrapolation based on the number of papers published until June 2022.



(b) The number of papers per tumor location.



(c) The number of papers per task.



(d) The number of papers per image modality.

Fig. 2: Breakdown of the papers included in this review.

dataset consisted of mostly children. Furthermore, the paper has a large focus on pediatric patients.

Contrary to what Quon *et al.* found, models investigated by Prince *et al.* [14] showed improved performance when including different imaging types, in this case CT and T1w MRI. They used many different pre-trained (ImageNet) deep learning models for the classification task. Network architectures included Inception models, ResNet models, and two models found using a (progressive) neural architecture search ((P)NAS). NAS, and PNAS, its successor, are recurrent neural networks (RNNs) that can be used to design a new CNN from scratch. Combinations of network architectures and hyperparameters were tried to find the best-performing model. They also applied different data augmentation methods, using either a stochastic method or transformation adversarial networks for data augmentation (TANDA), which is a method that uses Generative Adversarial Networks (GANs) and RNNs to do data augmentation. Independent of the input format (CT, MRI, or both), the best-performing models were ResNet models and had results comparable to clinical experts. The stochastic augmentation outperformed the TANDA method, but they mention that they did not optimize their TANDA method. The work is an initial exploration of the methodologies, but they suggest several improvements for further research. Another paper showing a proof of concept was by Ye *et al.* [15], who made diffusion-weighted MRI (DWI) scans of postmortem brain specimens of 9 children. A neural network was used to classify histopathological features, based on 10 diffusion metrics, provided by analysis of the DWI scans. The network consisted of 10 fully connected hidden layers. The results indicate that deep learning in combination with DWI could be used to classify histology, but in vivo testing of this method is still required.

Banerjee *et al.* [16] created a semi-automatic model to differentiate between the embryonal and alveolar subtypes of rhabdomyosarcoma (RMS), a soft tissue cancer, by using only T1w-MRI and DWI images [16]. They used transfer learning to train their CNN (AlexNet), which was pre-trained on images from ImageNet. They reached an 85% accuracy using this method. To use this model, a manual delineation of the tumor in a single slice is needed on both modalities, which decreases the level of automation.

2) Outcome prediction: In two papers, several machine learning methods were compared in predicting patient outcome in patients with neuroblastoma. Mayampurath *et al.* [17] used multiple machine learning models to predict response

to chemotherapy in these patients. Their neural network was a 2-layer CNN, which classified based on 2D metaiodobenzylguanidine (MIBG) scans. In an MIBG scan, a small amount of radioactive MIBG is administered, after which a gamma camera is used to determine the distribution of MIBG in the body. They also used a logistic regression model based on clinical parameters, and a Naive Bayes classifier (NB) that used the outcomes of both methods to do a final prediction, as well as a geometric mean of both methods. Their results indicated that the CNN performed worse than the logistic regression model, although not significantly. The NB model performed similarly, but the geometric mean method performed better, with an AUC of 0.73. Liu *et al.* [18] also used several machine learning methods. This included a 3-layer fully-connected neural network (FNN) and a 2D CNN. Patient outcome was predicted in six categories. The machine learning methods and the FNN used 105 radiomics features extracted from 3D CT images, while the CNN used 2D slices of these images. The convolutional neural network, which was pre-trained on the ImageNet dataset, performed poorer than the best performing other models. The FNN performed best in five out of six classification categories. Both works can be seen as a proof of concept for combining radiomics with machine learning to predict patient outcome, rather than only using a CNN.

B. Segmentation

Segmentation is the second most appearing use case for deep learning in pediatric oncology radiology. Manual tumor segmentation is a very time-consuming and difficult task, so using deep learning to speed up this process can save valuable time [29]. Five papers describe deep learning models used for segmentation. All of the papers use convolutional neural networks for this purpose. A widely used architecture for medical image segmentation is the U-Net [31], which is reflected in the fact that three out of the five papers have a U-Net-like architecture.

Peng *et al.* [5] used a 3D U-Net to segment brain tumors (high-grade gliomas, medulloblastomas, and leptomeningeal seeding tumors) in preoperative and postoperative children. The U-Net consisted of 5 levels. They furthermore created an automatic method to estimate the tumor volume and the RAPNO score, which is used to determine treatment response. This is the first fully automated method for segmentation and volume estimation of pediatric brain tumors using deep learning. The performance of the methods was similar to human experts.

Two papers used a U-Net-like architecture to segment optic pathway gliomas (OPG) on MRI images. OPG accounts for about 3-5% of the pediatric central nervous system tumors [19]. Because of this, only small pediatric OPG datasets are available. Both papers tackle this problem using transfer learning. The model of Artzi *et al.* [19] was initiated using the weights of a model trained on the ImageNet database. The convolution layers in the U-Net were replaced by residual blocks from the ResNet architecture. After segmentation, the tumors were classified using fuzzy c-means clustering. Their research showed promising results, despite their small and heterogeneous dataset. Nalepa *et al.* [20] used a nnU-Net (no new U-Net), a U-Net which configures itself, making it applicable to many different segmentation tasks. They pre-trained their model on segmentation of glioblastoma (GBM), for which a large dataset was available. Afterward, they fine-tuned their model to apply to OPG, by either updating all weights or only updating the weights of the final layer. They made their network architecture open source, allowing other research groups to use it.

Two other papers describe segmentation in patients with tumors in the kidney and the lymphatic system respectively, based on CT images. Corbat *et al.* [21] created a semi-automated method to segment the kidney and tumor in children with nephroblastoma based on CT images. A fully convolutional network (FCS-8s) was used to create a segmentation of the kidney and of the tumor separately. For this method to work, the segmentation of about 20% of the kidney and tumor must be known, on which the model trains, to be able to segment the other parts of the kidney and tumor automatically. They call this the OV^{2}ASSION method. The segmentations of the kidney and the tumor are combined using a deep learning network with seven convolutional layers and two pooling layers. The total time of the segmentation process is eight hours. Even though this is a long time, and partial segmentations are needed as input, it could be seen as an improvement compared to other networks (who needed more training data than 20%) or a completely manual method. Yin *et al.* [22] created a 3-layer CNN to segment lymphangioma lesions in CT images before and after interventional therapy. The model consisted of three convolutional layers, two pooling layers, and one fully connected layer. The model is said to outperform the Canny segmentation algorithm based on the dice score, but the cost function and the validation and test set size were not specified in the paper.

C. Image synthesis

Some imaging modalities, like PET and CT, use electromagnetic radiation that can be harmful at high doses. Therefore it can be beneficial to only acquire MRI scans, and use those as a basis to create synthetic CT images [24]. Furthermore, this removes the need for inter-modality image registration, resulting in fewer uncertainties [23, 24]. Four papers discuss models to generate synthetic images. Florkow *et al.* [23] adapted a U-Net to create synthetic CT images of children with Wilm's tumor or neuroblastoma, based on T1w and T2w MRI images. They used this to do dose calculations for radiotherapy, which they compared to dose calculations on a real CT scan. The MRI images first went through two convolution layers, before being concatenated to enter the U-Net structure. Most dosimetric differences were within clinically acceptable limits, showing that an MR-only radiotherapy workflow can be feasible using deep learning. Maspero *et al.* [24] also created synthetic CT images based on MRI images, but these were brain images. To do this, they used a conditional generative adversarial network (cGAN). The used dataset was very heterogeneous, both in patient size, shape, and age,

as well as imaging protocols. Despite this heterogeneous dataset, the model still performed well in creating synthetic CT images. The created images could be successfully used to do dose calculations for radiotherapy. Wang *et al.* [25] created synthetic CT-derived relative proton stopping power (RPSP) images using a consistent cycle generative adversarial network (ccGAN) with MRI images as input. They created RPSP images instead of CT images because dose calculations can be directly done on RPSP images. ccGAN is an adaptation of the cGAN, by adding a consistent loss, which compares the output image with the RPSP corresponding to the input MRI. cGAN would not have used this RPSP image, resulting in unsupervised learning. They found that using T1w or T2w MRI images yielded the best results, while FLAIR MRI, or combinations of these inputs had less good results.

Wang *et al.* [26] developed a CNN to generate diagnostic whole-body ^{18}F-FDG PET images for lymphoma patients from ultra-low-dose ^{18}F-FDG PET and T1w-MRI input. They used an enhanced deep super-resolution network (EDSR), which is a residual neural network (ResNet). They adapted this EDSR by including the T1w-MRI input using middle fusion, adding a skip connection from the PET input to the final prediction layer, and using a slice-wise method. Lymph nodes are better delineated on the AI-augmented scan compared to the low-dose scan, but the augmented scan does not discriminate each lesion as well as in the full-dose scan. This paper includes patients aged up to 30 years old, but the focus of this paper is on children and young adults, so it is left in for this review.

D. Image quality improvement

Two papers discuss deep learning methods to increase image quality. MRI images created using arterial spin labeling (ASL) are known to have a low signal-to-noise ratio (SNR). Hales *et al.* [27] used a convolutional neural network to denoise ASL images. Their network used an encoding component with convolutional and max pool layers, and a decoding component with convolutional an up sampling layers. Two skip connections were applied between the encoding and decoding components. Model training was done using 131 pediatric neuro-oncology patients, but evaluation was performed using 11 healthy adult volunteers. They found a significant increase in SNR when applying the CNN to the MRI data. Using denoising methods like this can save scan time in ASL acquisitions.

Ladefoged *et al.* [28] made a deep learning method to do PET attenuation correction for simultaneous PET/MRI acquisitions. They compared this with the MR-derived RESOLUTE method. The network was a modified version of a U-Net, with convolutions (stride 2) instead of max pool operations, and batch normalisation, a ReLU activation function and a dropout layer after each convolution. The deep learning method performed similarly to the RESOLUTE method, but the images corrected using the deep learning method were more similar to the ground truth according to both visual inspection and quantitative metrics.

III. Discussion

A. Overview

21 papers on deep learning applications in pediatric oncology radiology were discussed. Research into this topic is increasing rapidly, which is visualized in Fig. 2(a). The papers were categorized based on model task. Most models were classification models ($n = 10$). Five of these papers did not involve convolutional neural networks, while all of the other papers did. As expected, the models using large datasets did not use cross-validation or hold-out validation, while most models using smaller datasets did. Several papers applied widely used methods for specific tasks, such as U-Nets for segmentation and GANs for image synthesis, but other papers had different approaches, such as manually constructed CNNs for segmentation and a U-Net or a ResNet for image synthesis. In several papers, the use of multiple types of input images was compared to the use of only a single type. In some cases, using multiple input images was beneficial due to more available information for the model, while in other cases, using a single input image was preferable. This suggest that the number and type of input images should always be reconsidered when applying a method in a new situation. Papers focusing on a single model or a small number of models usually reached higher performance than papers comparing many different models, because the multiple models were often out-of-the-box solutions, and not optimized for the task. With deep learning, well-thought-out network architectures and hyperparameter choices are essential for reaching high performances [32].

B. Outlook

Only two papers published in 2020 or later do not involve convolutional neural networks. This suggests that network architectures are getting more complex over time. This trend is likely to be continued in the future, resulting in models that are even deeper than currently used models.

There are many more adults than children with cancer, resulting in more research and available data for adult oncology compared to pediatric oncology [4]. New methods are usually tested using data from the adult population. However, when these methods are successful, they can be adapted to apply to children as well. This could result in many new techniques in the pediatric field in the near future. Furthermore, data sharing initiatives, such as the Childhood Cancer Data Initiative (CCDI) can help with gathering larger datasets, which can make the use of deep learning easier.

C. Limitations

It is possible that not all relevant papers were discussed here. Not all abstracts of the found articles were read, since ASReview sorted the papers based on relevance. After ASReview suggested over 80 irrelevant papers, only the titles of the remaining papers were read. Furthermore, some papers might have not appeared when using the query, if relevant search terms were missing in the title, abstract, or keywords. We have tried to be as transparent as possible by explaining the search strategy.

IV. Conclusion

In this paper, 21 papers on deep learning applications in pediatric oncology radiation were reviewed, and categorized according to model task.

References

[1] G. Litjens et al. "A survey on deep learning in medical image analysis". In: *Medical Image Analysis* 42 (2017), pp. 60–88.

[2] K. Raza and K.N. Singh. "A Tour of Unsupervised Deep Learning for Medical Image Analysis". In: *Current Medical Imaging* 17.9 (2021), pp. 1059–1077.

[3] H. Daldrup-Link. "Artificial intelligence applications for pediatric oncology imaging". In: *Pediatr Radiol* 49.11 (2019), pp. 1384–1390.

[4] N. Davendralingam et al. "Artificial intelligence in paediatric radiology: Future opportunities". In: *The British Journal of Radiology* 94.1117 (2021), p. 20200975.

[5] J. Peng et al. "Deep learning-based automatic tumor burden assessment of pediatric high-grade gliomas, medulloblastomas, and other leptomeningeal seeding tumors". In: *Neuro Oncol* 24.2 (2022), pp. 289–299.

[6] S. Ramesh et al. "Applications of Artificial Intelligence in Pediatric Oncology: A Systematic Review". In: *JCO Clin Cancer Inform* 5 (2021), pp. 1208–1219.

[7] J. Huang et al. "Artificial Intelligence Applications in Pediatric Brain Tumor Imaging: A Systematic Review". In: *World Neurosurgery* 157 (2022), pp. 99–105.

[8] R. van de Schoot et al. "An open source machine learning framework for efficient and transparent systematic reviews". In: *Nature Machine Intelligence* 3.2 (2021), pp. 125–133.

[9] E. Orphanidou-Vlachou et al. "Texture analysis of T1 - and T2 -weighted MR images and use of probabilistic neural network to discriminate posterior fossa tumours in children". In: *NMR Biomed* 27.6 (2014), pp. 632–639.

[10] N. Zarinabad et al. "Multiclass imbalance learning: Improving classification of pediatric brain tumors from magnetic resonance spectroscopy". In: *Magn Reson Med* 77.6 (2016), pp. 2114–2124.

[11] M. Li et al. "Machine learning methods for MRI biomarkers analysis of pediatric posterior fossa tumors". In: *Biocybernetics and Biomedical Engineering* 39.3 (2019), pp. 765–774.

[12] M. Zhang et al. "Machine Assist for Pediatric Posterior Fossa Tumor Diagnosis: A Multinational Study". In: *Neurosurgery* 89.5 (2021), pp. 892–900.

[13] J.L. Quon et al. "Deep Learning for Pediatric Posterior Fossa Tumor Detection and Classification: A Multi-Institutional Study". In: *American Journal of Neuroradiology* 41.9 (2020), pp. 1718–1725.

[14] E.W. Prince et al. "Robust deep learning classification of adamantinomatous craniopharyngioma from limited preoperative radiographic images". In: *Scientific Reports* 10.1 (2020), p. 16885.

[15] Z. Ye et al. "Diffusion histology imaging differentiates distinct pediatric brain tumor histology". In: *Sci Rep* 11.1 (Feb. 2021), p. 4749.

[16] I. Banerjee et al. "Transfer learning on fused multiparametric MR images for classifying histopathological subtypes of rhabdomyosarcoma". In: *Computerized Medical Imaging and Graphics* 65 (2018). Advances in Biomedical Image Processing, pp. 167–175.

[17] A. Mayampurath et al. "Predicting Response to Chemotherapy in Patients With Newly Diagnosed High-Risk Neuroblastoma: A Report From the International Neuroblastoma Risk Group". In: *JCO Clinical Cancer Informatics* 5 (2021), pp. 1181–1188.

[18] G. Liu et al. "Incorporating Radiomics into Machine Learning Models to Predict Outcomes of Neuroblastoma". In: *Journal of Digital Imaging* 35.3 (2022), pp. 605–612.

[19] M. Artzi et al. "Automatic segmentation, classification, and follow-up of optic pathway gliomas using deep learning and fuzzy c-means clustering based on MRI". In: *Medical Physics* 47.11 (2020), pp. 5693–5701.

[20] J. Nalepa et al. "Segmenting pediatric optic pathway gliomas from MRI using deep learning". In: *Computers in Biology and Medicine* 142 (2022), p. 105237.

[21] L. Corbat et al. "Fusion of multiple segmentations of medical images using OV^{2}ASSION and Deep Learning methods: Application to CT-Scans for tumoral kidney". In: *Comput Biol Med* 124 (2020), p. 103928.

[22] C. Yin et al. "Computed Tomography Image Characteristics before and after Interventional Treatment of Children's Lymphangioma under Artificial Intelligence Algorithm". In: *Computational and Mathematical Methods in Medicine* 2021 (2021), p. 2673013.

[23] M.C. Florkow et al. "Deep learning-enabled MRI-only photon and proton therapy treatment planning for paediatric abdominal tumours". In: *Radiotherapy and Oncology* 153 (2020), pp. 220–227.

[24] M. Maspero et al. "Deep learning-based synthetic CT generation for paediatric brain MR-only photon and proton radiotherapy". In: *Radiotherapy and Oncology* 153 (2020), pp. 197–204.

[25] C. Wang et al. "Toward MR-only proton therapy planning for pediatric brain tumors: Synthesis of relative proton stopping power images with multiple sequence MRI and development of an online quality assurance tool". In: *Medical Physics* 49.3 (2022), pp. 1559–1570.

[26] Y. Wang et al. "Artificial intelligence enables whole-body positron emission tomography scans with minimal radiation exposure". In: *European Journal of Nuclear Medicine and Molecular Imaging* 48.9 (2021), pp. 2771–2781.

[27] P.W. Hales, Josef Pfeuffer, and Chris A. Clark. "Combined Denoising and Suppression of Transient Artifacts in Arterial Spin Labeling MRI Using Deep Learning". In: *Journal of Magnetic Resonance Imaging* 52.5 (2020), pp. 1413–1426.

[28] C.N. Ladefoged et al. "Deep Learning Based Attenuation Correction of PET/MRI in Pediatric Brain Tumor

Patients: Evaluation in a Clinical Setting". In: *Frontiers in Neuroscience* 12 (2019).

[29] A. Işin, C. Direkoglu, and M. Şah. "Review of MRI-based Brain Tumor Image Segmentation Using Deep Learning Methods". In: *Procedia Computer Science* 102 (2016). 12th International Conference on Application of Fuzzy Systems and Soft Computing, ICAFS 2016, 29-30 August 2016, Vienna, Austria, pp. 317–324.

[30] V.N. Kornienko and I.N. Pronin. *Diagnostic Neuroradiology*. Springer, 2009.

[31] G. Du et al. "Medical image segmentation based on u-net: A review". In: *Journal of Imaging Science and Technology* 64 (2020), pp. 1–12.

[32] F. Isensee et al. "nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation". In: 18.2 (2021), pp. 203–211.

APPENDIX A
(DUTCH) PLAIN LANGUAGE SUMMARY

Kunstmatige intelligentie wordt steeds meer gebruikt bij medische beeldvorming. Een geavanceerde methode die recentelijk veel ontwikkelingen heeft doorgemaakt is het kunstmatige neurale netwerk. In een kunstmatig neuraal netwerk wordt de structuur van hersenen gesimuleerd in zijn eenvoudigste vorm, het bestaat uit kunstmatige neuronen die met elkaar in verbinding staan. Er kan informatie (zoals een medische afbeelding) worden ingevoerd in het netwerk, waarna het model een voorspelling doet. Deze vorm van kunstmatige intelligentie wordt ook wel *deep learning* genoemd. Voor dit artikel heb ik onderzocht wat voor toepassingen van deep learning er zijn voor medische beeldvorming van kinderen met kanker. Ik heb 21 relevante artikelen gevonden.

De meest voorkomende vorm (tien keer) van deep learning was classificatie. Bij een classificatie is de voorspelling van het netwerk een categorie waar de patiënt of tumor bij hoort. Het kan bijvoorbeeld zijn dat het type tumor wordt gedetecteerd door het netwerk, of dat er een voorspelling wordt gedaan over of de patiënt te genezen is. Na classificatie kwam segmentatie het meeste voor (vijf keer). Bij segmentatie wordt bepaald welke onderdelen van de medische afbeelding een tumor of orgaan bevat, en welk onderdelen niet. Dit kan bijvoorbeeld gebruikt worden om het volume van een tumor te bepalen. Vier artikelen beschreven methodes om nieuwe beelden te creëren. Vaak zijn er bijvoorbeeld zowel CT als MRI beelden nodig. Om te voorkomen dat er verschillende soorten scans moeten worden gemaakt, kan een neuraal netwerk gebruikt worden om een kunstmatige CT scan te creëren op basis van een MRI scan. In de laatste twee artikelen werden neurale netwerken gebruikt om de kwaliteit van medische beelden te vergroten, bijvoorbeeld door ruis te verwijderen.

Er zijn verschillende soorten neurale netwerken, die ieder gespecialiseerd zijn in andere taken. Voorbeelden zijn het U-Net, dat ontwikkeld is voor segmentaties, en de GAN, die gebruikt wordt voor het creëren van kunstmatige scans. Echter, sommige artikelen gebruikten ook minder gangbare methodes. Zo was er zelfs een paper die een U-Net gebruikte om een kunstmatige scan te creëren.

De verwachting voor de toekomst is dat de modellen steeds complexer zullen worden. Verder zijn er verschillende initiatieven die erop gericht zijn om meer data beschikbaar te maken voor onderzoek, wat ervoor kan zorgen dat er meer mogelijk wordt met neurale netwerken op het vlak van kinderkanker.