

Deep Learning CT Image Reconstruction in Clinical Practice

CT-Bildrekonstruktion mit Deep Learning in der klinischen Praxis

Authors

Clemens Arndt, Felix Güttler, Andreas Heinrich, Florian Bürckenmeyer, Ioannis Diamantis, Ulf Teichgräber

Affiliation

Department of Radiology, Jena University Hospital, Jena, Germany

Key words

computed tomography, image reconstruction, deep neural network, deep learning, artificial intelligence, image denoising

received 11.05.2020

accepted 20.08.2020

published online 10.12.2020

Bibliography

Fortschr Röntgenstr 2021; 193: 252–261

DOI 10.1055/a-1248-2556

ISSN 1438-9029

© 2020, Thieme. All rights reserved.

Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany

Correspondence

Dr. Clemens Arndt

Institut für Diagnostische und Interventionelle Radiologie, Universitätsklinikum Jena, Am Klinikum 1, 07751 Jena, Germany

Tel.: ++ 49/36 41/9 32 48 39

clemens.arndt@med.uni-jena.de

ZUSAMMENFASSUNG

Hintergrund Die Computertomografie (CT) ist eine zentrale Modalität der modernen Radiologie, die in nahezu allen medizinischen Fachdisziplinen, insbesondere aber in der Notfallmedizin, einen wichtigen Bestandteil zur Gesundheitsversorgung liefert. Die Berechnung bzw. Rekonstruktion von Schnittbildern aus den rohen Messwerten der CT-Untersuchungen stellt mathematisch ein inverses Problem dar. Bisher waren die gefilterte Rückprojektion und die iterative Rekonstruktion (IR) die Goldstandards, um schnell und zuverlässig Bilder zu berechnen. Aktuell gibt es im Rahmen neuerer Entwicklungen im Bereich der künstlichen Intelligenz mit dem Deep Learning (DL) einen weiteren Ansatzweg in der klinischen Routine.

Methode Dieser Übersichtsartikel erläutert die bisherigen Prinzipien der Bildrekonstruktion, das Konzept des DL und das Anwendungsprinzip zur Rekonstruktion. Anschließend werden kommerziell verfügbare Algorithmen und bisherige Studien diskutiert und Grenzen sowie Probleme dargestellt.

Ergebnis und Schlussfolgerung DL als Methode des Machine Learning nutzt im Allgemeinen ein trainiertes, künstliches, neuronales Netzwerk zur Lösung von Problemen. Aktuell sind DL-Rekonstruktionsalgorithmen von 2 Herstellern für die klinische Routine verfügbar. In den bisherigen Studien konnte eine Reduktion des Bildrauschens und eine Verbesserung der Gesamtqualität gezeigt werden. Eine Studie zeigte, bei höherer Qualität der mittels DL rekonstruierten Bilder, eine vergleichbare diagnostische Genauigkeit zur IR für die Detektion von Koronarstenosen. Weitere Studien sind notwendig und sollten vor allem auch auf eine klinische Überlegenheit abzielen, während dabei ein breiter Umfang an Pathologien umfasst werden muss.

Kernaussagen:

- Nach der aktuell verbreiteten iterativen Rekonstruktion können CT-Schnittbilder im klinischen Alltag nun auch mittels Deep Learning (DL) als Methode der künstlichen Intelligenz rekonstruiert werden.
- Durch DL-Rekonstruktionen können Bildrauschen vermindert, Bildqualität verbessert und möglicherweise Strahlung reduziert werden.
- Eine diagnostische Überlegenheit im klinischen Kontext sollte in kommenden Studien gezeigt werden.

ABSTRACT

Background Computed tomography (CT) is a central modality in modern radiology contributing to diagnostic medicine in almost every medical subspecialty, but particularly in emergency services. To solve the inverse problem of reconstructing anatomical slice images from the raw output the scanner measures, several methods have been developed, with filtered back projection (FBP) and iterative reconstruction (IR) subsequently providing criterion standards. Currently there are new approaches to reconstruction in the field of artificial intelligence utilizing the upcoming possibilities of machine learning (ML), or more specifically, deep learning (DL).

Method This review covers the principles of present CT image reconstruction as well as the basic concepts of DL and its implementation in reconstruction. Subsequently commercially available algorithms and current limitations are being discussed.

Results and Conclusion DL is an ML method that utilizes a trained artificial neural network to solve specific problems. Currently two vendors are providing DL image reconstruction algorithms for the clinical routine. For these algorithms, a decrease in image noise and an increase in overall image qual-

ity that could potentially facilitate the diagnostic confidence in lesion conspicuity or may translate to dose reduction for given clinical tasks have been shown. One study showed equal diagnostic accuracy in the detection of coronary artery stenosis for DL reconstructed images compared to IR at higher image quality levels. Consequently, a lot more research is necessary and should aim at diagnostic superiority in the clinical context covering a broadness of pathologies to demonstrate the reliability of such DL approaches.

Key Points:

- Following iterative reconstruction, there is a new approach to CT image reconstruction in the clinical routine using deep learning (DL) as a method of artificial intelligence.

- DL image reconstruction algorithms decrease image noise, improve image quality, and have potential to reduce radiation dose.
- Diagnostic superiority in the clinical context should be demonstrated in future trials.

Citation Format

- Arndt C, Güttler F, Heinrich A et al. Deep Learning CT Image Reconstruction in Clinical Practice. *Fortschr Röntgenstr* 2021; 193: 252–261

Background

In the clinical routine, the crucial aim of CT imaging is to provide clinically relevant information or more specifically, information as to whether a specific feature is found or not, with reasonable certainty. However, the certainty of that radiological decision depends heavily on the image quality, e. g. contrast, resolution, noise, and artifacts [1].

Alongside image quality, radiation dose is an important element in CT protocol optimization. Furthermore, acquisition and reconstruction time should be acceptable, especially in a high-level workflow setting such as emergency radiology. Both image quality and radiation dose are affected by acquisition parameters, patient constitution, and positioning [2, 3]. Another important element of the CT imaging process, influencing quality and reconstruction time, is the mathematical transformation of the raw data into a three-dimensional volume viewable as an anatomical slice image. To solve this transformation, several reconstruction algorithms have been developed [4]. Algorithms routinely used since the introduction of computer tomography are *filtered back projection* (FBP), which was used into the early 2010 s, and *iterative reconstruction* (IR), which was subsequently used [5]. To understand the ongoing developments in CT image reconstruction, the principles of FBP and IR are briefly reviewed, followed by an introduction to deep learning (DL).

Filtered back projection

FBP is the most commonly used analytical reconstruction method owing to its computational efficiency and stability creating images rapidly. The mathematical approach to FBP is primarily the idea that a projection, consisting of measurements in multiple angles, can be back projected into a model of the scanned object by an inverse radon transformation with a high-pass filter (► **Fig. 1**). Without a filter, smearing the projection values into the radiation path would result in a blurry object morphology. This filter or *kernel* can additionally be modified to facilitate evaluation of specific anatomic components, e. g. bones [4]. Later, FBP benefited from advanced methods compensating geometric problems [6]. The

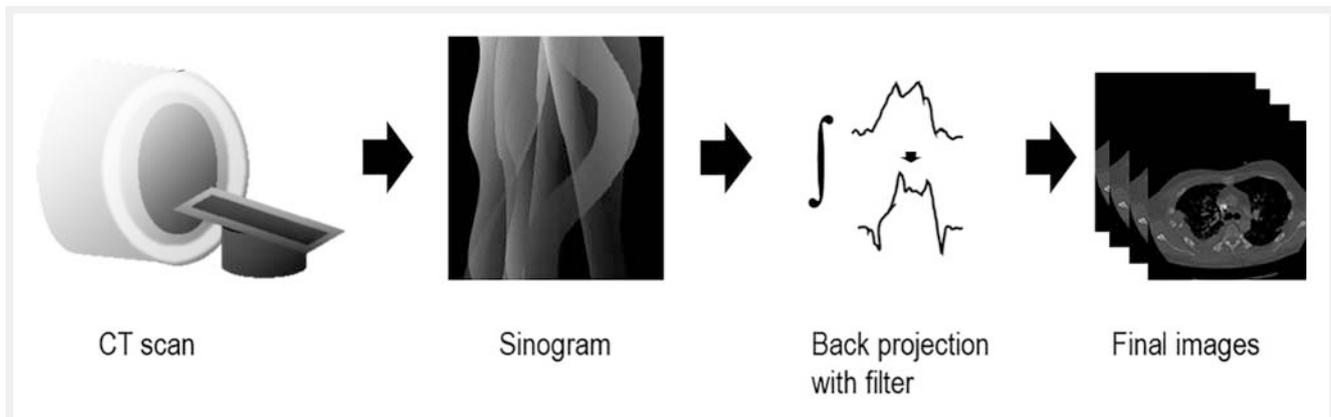
disadvantage of FBP is the strong relationship between radiation dose and noise, that is especially problematic in obese patients [7]. In the ongoing advancement in CT imaging from measured data to algorithmic optimized data, FBP is closest to measured data. As computational power for industrial purposes and graphical processing was advancing, FBP was gradually replaced by iterative reconstruction [8].

Iterative reconstruction

The IR method starts with the creation of an image estimate that is basically equal to FBP. Subsequently, the image estimate is projected forward into an artificial sinogram and iteratively corrected by comparison with the original raw data sinogram. When a pre-defined endpoint condition in the algorithm is fulfilled, the iterative cycle is stopped, and the images are readily reconstructed (► **Fig. 2**). However, the iterative cycle can get implemented in different steps, e. g. the raw data and the image domain. Additionally, further statistic adjustments and modelling can be used. Roughly iterative algorithms can be divided into methods without statistics (e. g. ART, the first iterative algorithm), methods with modeling of photon counting statistics (e. g. hybrid IR), and *model-based* methods beyond photon counting (MBIR) [9, 10]. With enhanced image quality, a reduction in patient dose was feasible. In a systematic review from 2015, the mean effective dose for contrast-enhanced chest CT with IR decreased by 50 % compared to FBP [11]. IR is now considered criterion standard for CT image reconstruction.

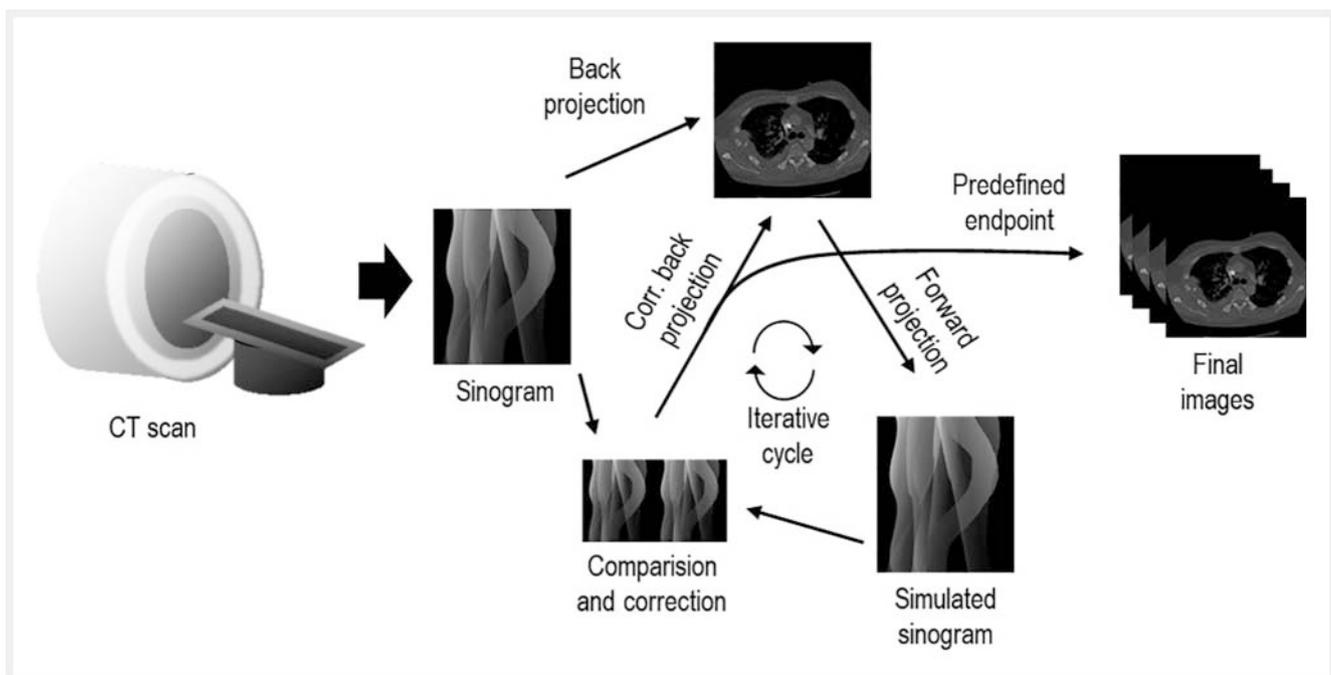
Artificial neural networks and deep learning

The recent rise in popularity of artificial intelligence is primarily due to the advances in the field of *artificial neural networks* (ANNs). ANNs are a subfield of *machine learning* (ML). In machine learning a model learns by training data and is subsequently able to perform specific tasks. This process is termed supervised if the training data contains both input and desired output. The system learns how to do tasks only by showing what the results should be and not how to perform those tasks [12]. Aside from the super-



► **Fig. 1** Schematic illustration of image reconstruction by filtered back projection. After the acquisition the raw data consisting of the measurement of attenuation profiles in multiple angles is transformed into an image domain with a filter or kernel to compensate for the blur emerging with reconstruction.

► **Abb. 1** Schematische Darstellung einer Bildrekonstruktion durch gefilterte Rückprojektion. Nach Akquisition werden die Rohdaten, welche den Schwächungsprofilen in multiplen Winkel entsprechen, mittels Rückprojektion als anatomisches Schnittbild rekonstruiert. Ein Filter bzw. Kernel kompensiert dabei die Kantenverschmierung.



► **Fig. 2** Simplified exemplary illustration of iterative image reconstruction. From the raw data a reconstructed image estimate is generated, which is iteratively compared with the original sinogram in the forward projection and corrected until a predefined endpoint is reached.

► **Abb. 2** Vereinfachte, beispielhafte Darstellung einer iterativen Bildrekonstruktion. Aus den Rohdaten wird eine Rückprojektion erstellt, welche in der Vorwärtsprojektion iterativ mit dem originalen Sinogramm verglichen und verbessert wird, bis ein definierter Endpunkt erreicht wird.

vised approach, the learning process can be unsupervised, or reward-based as in reinforcement learning [13]. ANNs as a method of machine learning are inspired by the operation principle of neurons, although they do not simulate them in detail. ANNs consist of interconnected nodes, comparable to the cell body of a neuron with synapses for signal transduction. Those networks are usually built with an input layer, one or more hidden layers, and an

output layer. On information flow values are passed from one node to the next connected nodes. However, the connections themselves are weighted and will modify the incoming values from the previous nodes to a subsequent node. In the node itself, the sum of all weighted incoming values is passed through an activation function that determines the output of a single node. Depending on the sum and the *activation function*, similar to the

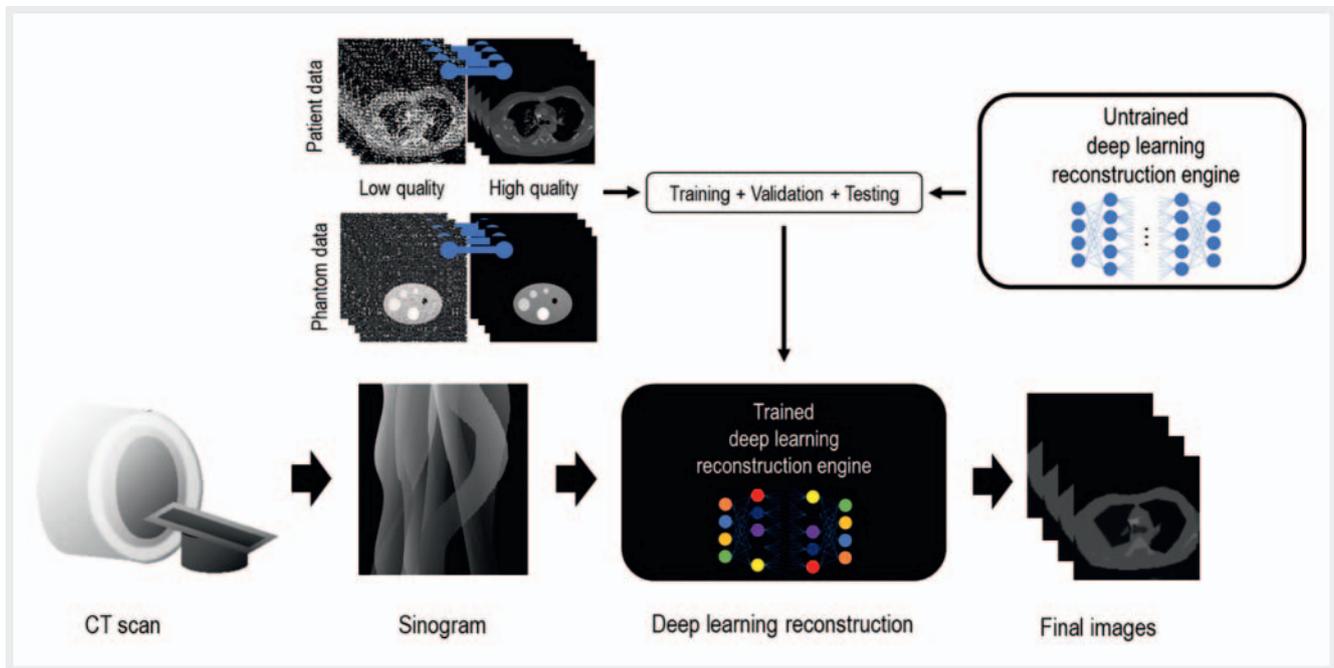
principle of a neuronal excitation threshold, there might be no activation at all. Leaving the output layer, the final calculated output is compared to the desired output by a *loss* or *error function*. Subsequently the previously mentioned weights are adjusted, and the network is recalculated. Using this process of iterative readjustment and reevaluation, the network is trained, in order to finally produce an output with an acceptable level of inaccuracy. The training data, consisting of multiple inputs with corresponding desired outputs (also known as *ground truth*), is usually divided into a training, a validation, and a test set. Different from the training set, the validation set will not be used directly for the training regarding the iterative adjustment of weights. However, the validation set is utilized to monitor network performance and to fine-tune the non-learnable hyperparameters that need to be manually adjusted beforehand. Aside from performance evaluation and hyperparameter tuning, the dataset split enables the recognition of overfitting. That means a network is too closely adapted to the training set, by identifying more features than the data provides. This inevitably leads to an inability to recognize unseen data correctly. When training has been finished, there is still a lot of testing ahead, usually with another unseen testing dataset. When the network has been trained and tested to satisfaction, it will eventually be applied to unknown data outside of the training environment. This is called *inferencing* [14]. With ANNs, two concepts are commonly mentioned: *deep learning* (DL) and *convolutional neural networks* (CNNs). Deep refers to the multiple layers of an ANN between the input and the output layer, increasing the complexity usually with millions of parameters being calculated. Those networks are sometimes labeled as *deep neural networks* (DNNs). Convolutional neural networks are a class of DNN. They are especially interesting for tasks in computer vision, as they operate on grid pattern data, for example images. Usually CNNs show a typical architecture with convolutional and pooling layers that can be passed iteratively and a fully connected layer. In the convolutional and pooling layers multiple filter kernels consisting of a grid of weights are applied to the image extracting different features. While the first layers usually represent low-level features like edges and corners, mid-level features might be parts of objects or organs and high-level features are whole structures or organs. The fully connected layers finally flatten the previous layer and create a one-dimensional vector so the output can be classified. In the training process additional to the adjustment of weights, the filter kernels are optimized, again, by loss functions comparing the output to the known ground truth [15]. Generative adversarial networks (GANs) are another type of ANN that show potential for tasks in medical imaging, e. g. denoising. They consist of two neural networks: a generator, creating artificial samples from input data, and a discriminator, learning to distinguish real from generated data [16]. The final principle of transforming an image with characteristics of another image is called image-to-image translation [17].

There are multiple possible objectives for ANNs in radiology. However, the most common can be assigned to the following applications: classification, detection, segmentation, and image optimization, e. g. denoising [18]. Classification tasks in radiology include concluding whether lesions, e. g. pulmonary nodules, are benign or malignant. While lesions are characterized during clas-

sification, the existence of lesions and their location are the focus of detection tasks in the CNN. A promising example is the study by Lakhani et al. who trained two DNNs on the detection of tuberculosis. The network ensemble showed a sensitivity of 97.3% and a specificity of 94.7%. This approach could facilitate mass screenings and have a positive impact on healthcare in regions with TBC epidemics and low radiological infrastructure or expertise [19]. Segmentation defines borders between anatomical compartments or structures, for example of the liver. Segmentation of organs will, for example, accelerate volumetric tasks and radiation therapy planning [14]. Image denoising as an application is especially beneficial for low-dose CT. Since 2017, multiple networks and methods have been proposed. In 2017 Wolterink et al. combined a generator CNN with an adversarial CNN and showed the improved ability to generate images similar to routine dose from low-dose CT [20].

Deep learning image reconstruction

The intention of developing DL image reconstruction algorithms is mainly the improvement of image quality in comparison to the performance of current IR algorithms, while reduction of radiation dose might be a secondary benefit. Although there are differences with respect to the structure and the development of the DLR algorithms, the basic concept is comparable (► Fig. 3). The first step is designing the network architecture. This includes setting the hyperparameters, i. e., parameters not learned during the training process, for instance the size and topology of the network, the type of activation function, as well as the learning rate. The definition of the hyperparameters is crucial to the performance of a network. The next step is the training process. As previously explained, this requires a dataset consisting of multiple low-quality input images and corresponding high-quality ground truth images. The dataset originates from phantom scans as well as patient examinations in the clinical setting. Starting the supervised training process, the DLR algorithm will create an output image out of the low-quality input data that will be immediately compared to the corresponding ground truth image, calculating the error function to the output. In a so-called backpropagation process, the impact of each weight to the error will be calculated and subsequently weights in the network will be adjusted accordingly. This process is repeated iteratively, and the network will successively learn to eliminate most of the noise, while keeping anatomical detail [21, 22]. With the complexity of the image data, the training process will likely involve millions of calculations on each iteration and require an extreme amount of computational power, even when facilitated by use of GPU. As previously mentioned, the complete dataset will usually be split into a training set, a validation set, and a test set to ensure no overfitting is present. Apart from stability, a thorough evaluation of anatomical details is necessary. There is no guarantee that important information will not be lost during reconstruction. Conversely, a reconstruction algorithm might also be able to create new artificial structures, known as hallucinations, that may simulate a pathology [23].



► **Fig. 3** Schematic steps of deep learning image reconstruction. In the training phase low- and high-quality images are fed into a neural network. After completed training and validation, the algorithm can generate high-quality images out of unseen low-quality inputs.

► **Abb. 3** Schematische Darstellung des Trainings und der Bildrekonstruktion durch ein Deep-Learning-Netzwerk. Während der Trainingsphase werden dem Netzwerk Bilder mit hoher und niedriger Qualität zugeführt. Nach Abschluss dieser Phasen und der Validierung ist das Netzwerk in der Lage, hochqualitative Bilder aus bisher ungesehenen Daten mit geringer Qualität zu generieren.

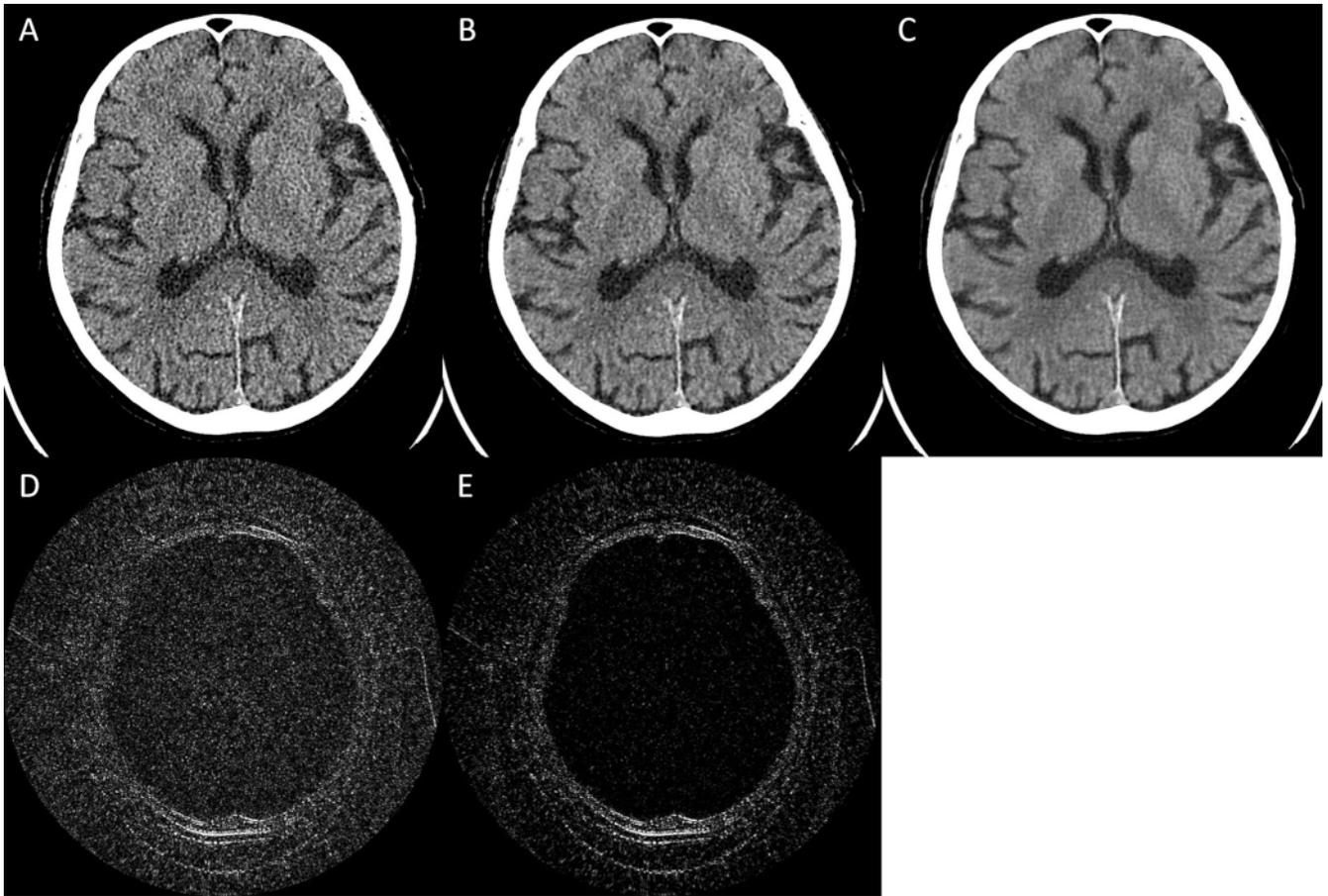
Advancements from deep learning image reconstructions

There are several advancements from DL image reconstruction that are intricately linked to each other. The most essential advancement is improved noise reduction. Noise is the variation of attenuation coefficients in homogeneously dense material. In reconstruction this leads to a grainy image appearance [24]. While DL algorithms are trained on low-dose data, they are capable of noise reduction while containing a true signal. This aspect can be quantified by the signal-to-noise ratio (SNR) [25]. In 2017, Jin et al. proposed a CNN based on the U-net, which showed improved SNR in experimental datasets compared to IR [26]. The loss of anatomical information can be visualized by subtraction of CT slices reconstructed in DL and iterative methods (► **Fig. 4**). Denoising has two immediate clinically relevant aspects. Firstly, lesions of any kind might be better detectable. This is expected to facilitate the daily work of a radiologist, e. g. in oncological imaging (► **Fig. 5**). Secondly, noise is directly associated with radiation dose. A lower tube current will reduce the dose, but noise will increase concurrently, thereby deteriorating image quality and assessability. With DL image reconstruction, the increase in noise can be compensated. This aspect will contribute to imaging in low-dose scenarios, e. g. CT of pediatric patients (► **Fig. 6**). Another feasible opportunity of DL image reconstruction is the virtual improvement of spatial resolution by creating thin-slice images out of thick slices. This was demonstrated by Umehara et al. in chest CT imaging as well by Park et al. who further success-

fully showed deblurring of bone edges due to partial volume effect [27, 28]. An additional advancement due to DL image reconstruction is improved artifact reduction. Beam hardening artifacts remain a highly relevant problem especially in head and neck imaging due to dental fillings as well in the case of imaging after osteosynthesis implants, impeding detection of implant loosening or periprosthetic fractures. In 2018, Zhang et al. developed a CNN trained on metal-free, metal-inserted, and precorrected images capable of superior metal artifact suppression while providing anatomical detail [29]. In addition, DL image reconstruction will be able to reduce beam hardening artifacts from bony structures. Especially in head CT, such artifacts can mask subtle intracranial hemorrhage (► **Fig. 7**).

Deep learning reconstruction algorithms in the clinical routine

To our knowledge, there are currently two commercially available CT image reconstruction algorithms using DL methods cleared by the FDA, TrueFidelity by GE Healthcare and AiCE by Canon Medical Systems. For this review we did a literature search, performed via PubMed with the search terms: “deep learning CT reconstruction” “DLR”, “DLIR”, “AiCE” and “TrueFidelity”. Studies performed with TrueFidelity or AiCE as reconstruction algorithms were of interest. Altogether seven studies were relevant. The studies, their characteristics, and important results are depicted in ► **Table 1**. Both TrueFidelity and AiCE have been investigated in phantom and

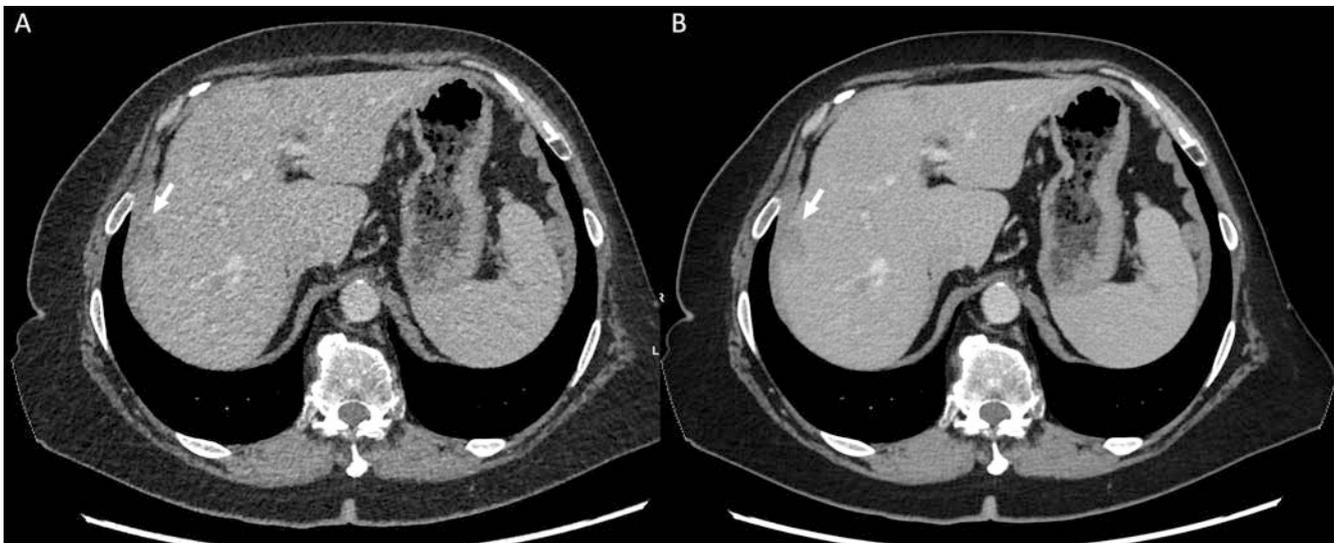


► **Fig. 4** Comparison of a head CT of an 87-year-old woman, reconstructed by **A** FBP, **B** ASiR-V-50% and **C** DLIR-H, show reduced image noise for DL image reconstruction. Subtraction images of **D** FBP-DLIR-H and **E** ASiR-V-50%-DLIR-H demonstrate no anatomical structure besides the calvaria, indicating that anatomical detail is preserved. Slice thickness: 0.625 mm. W: 100, C: 40.

► **Abb. 4** Der Vergleich einer CCT einer 87-jährigen Frau, rekonstruiert durch **A** FBP, **B** ASiR-V-50% und **C** DLIR-H, ergibt einen deutlich enträuschten Bildeindruck. In den Differenzbildern **D** FBP-DLIR-H und **E** ASiR-V-50%-DLIR-H zeigt sich neben dem dichten Schädelknochen keine anatomische Struktur, als Indikator für die Erhaltung anatomischer Details. Schichtdicke: 0,625 mm. W: 100, C: 40.

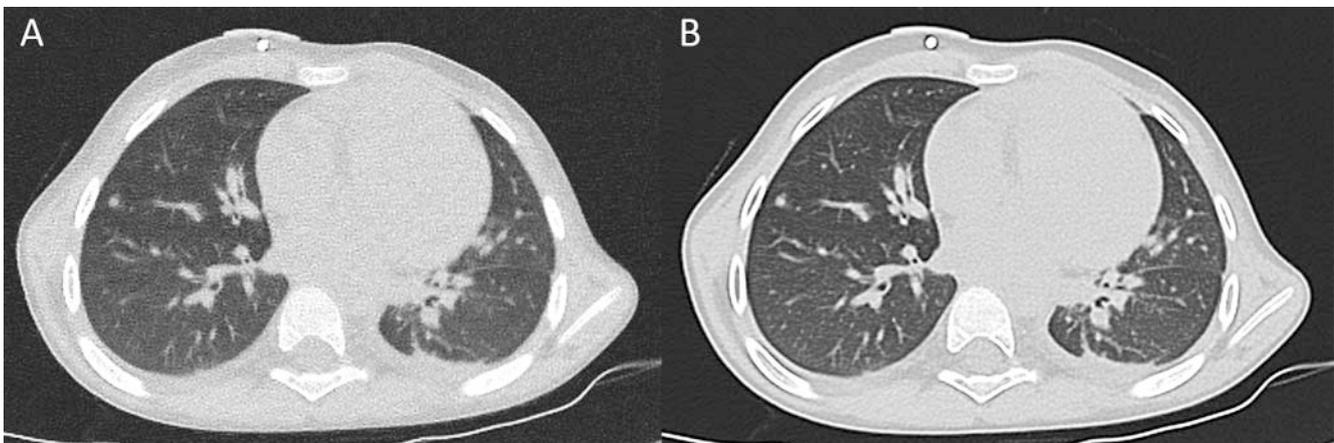
patient studies. The evaluated criteria usually consisted of quantitative and qualitative measurements. Utilizing DLIR in coronary computed tomography angiography, Benz et al. showed a significant reduction in noise and higher image quality, while DL image reconstruction was equal to IR with regard to diagnostic accuracy, sensitivity, and specificity in the detection of significant coronary artery stenosis with invasive coronary angiography as criterion standard [30]. The phantom study by Greffier et al. published in 2020 showed a potential for dose reduction of up to 56% for TrueFidelity, by achieving comparable detectability with iterative reconstruction and DL while lowering the dose [31]. This was especially successful for small and subtle features. Detectability was calculated utilizing a non-prewhitening matched filter with eye filter (NPWE) model as a surrogate for human perception that includes noise and resolution [32]. A potential for dose reduction was furthermore shown in the clinical context in a study from our institution in 2020. The DL image reconstruction allowed improved SNR and CNR at the same dose levels compared to IR [33]. AiCE was examined by Akagi et al. in 2019 by measuring noise, CNR, and image quality in abdominal CT scans, as stated

by two radiologists. They were able to show improved CNR and image quality compared to images generated by hybrid iterative reconstruction and MBIR [34]. Another study in 2019 by Nakamura et al. evaluated the detectability of hypovascular hepatic metastasis in images reconstructed with AiCE in addition to measuring noise and CNR. They demonstrated less image noise and superior conspicuity for DLR compared to the iterative algorithm [35]. A phantom study by Higaki et al. from 2020 additionally examined spatial resolution with a task-based modulation transfer function at 10% [36]. MBIR outperformed the other algorithms, while the DL algorithm was still superior to FBP and hybrid IR. Furthermore, this study calculated the detectability index for the aforementioned reconstruction methods at different doses. Detectability as the quality criterion is dependent on noise as well as spatial resolution. Consequently, MBIR showed the highest detectability in most dose settings. However, for low dosage the DL method outperformed the other algorithms. This underlines the potential for clinical scenarios with the aim of low dosage, e. g. pediatric CT. The imaging quality of the common bile duct in maximum intensity projection was improved when reconstructed with



► **Fig. 5** Comparison of an abdominal CT of a 66-year-old woman with hepatic metastases generated by **A** ASiR-V-50 % and **B** DLIR-H shows a more homogeneous liver texture due to denoising with DLIR-H. Slice thickness: 0.625 mm. W: 400, C: 40.

► **Abb. 5** Abdomen-CT einer 66-jährigen Patientin mit Rekonstruktionen durch **A** ASiR-V-50 % und **B** DLIR-H. Letztere zeigt durch das verminderte Bildrauschen eine deutlich homogenere Lebertextur. Schichtdicke: 0,625 mm. W: 400, C: 40.



► **Fig. 6** Comparison of a low-dose chest CT of a 7-year old boy generated by **A** ASiR-V-50 % and **B** DLIR-H (without lung kernel) shows a sharper image impression and reduced noise with DLIR-H. Slice thickness: 1.25 mm. W: 1600, C: -500.

► **Abb. 6** Der Vergleich eines Low-Dose-Thorax-CT eines 7-jährigen, rekonstruiert durch **A** ASiR-V-50 % und **B** DLIR-H (ohne Lungenkernel), zeigt einen deutlich schärferen Bildeindruck und ein deutlich vermindertes Bildrauschen mit DLIR-H. Schichtdicke: 1,25 mm. W: 1600, C: -500.

DL compared to iterative methods in a study by Narita et al. in 2020 [37]. This could facilitate scenarios with bile duct pathologies when MR is not available or possible and preoperative planning is needed.

To conclude, for both DL image reconstruction methods, a decrease of noise and improved CNR were consistently shown, while one study showed diagnostic comparability regarding the presence of significant coronary artery stenosis. Furthermore, both TrueFidelity and AiCE were rated superior for subjective image quality by radiologists. However, this must be interpreted with caution. As already stated by Hoeschen, a good image impression is not necessarily associated with a higher diagnostic value [38].

Limitations

Although DLR algorithms seem to be highly effective for improving image quality, there are several limitations or issues to be discussed.

Firstly, further external validation of DL image reconstruction is necessary. Although comparison of non-diagnostic quantitative parameters e.g. CNR and noise, is important, advancements should be driven by clinical superiority, for instance showing improved detection of specific lesions or even a higher rate of more certain report statements regarding ambiguous lesions. When comparing different DL image reconstruction methods, it

► **Table 1** Summary of published studies investigating DL image reconstruction algorithms by GE Healthcare and Canon Medical Systems.

► **Tab. 1** Übersicht über bisherige Studien, welche die DL-Rekonstruktion von GE HealthCare und Canon Medical Systems untersucht haben.

author of study and year of publication	study subject	number of patients included	reconstruction algorithms and vendor	evaluated criteria	important results
Benz DC et al. (2020) [30]	patients	43	<ul style="list-style-type: none"> ASiR-V-70 % SD ASiR-V-70 % HD DLIR-M DLIR-H by GE Healthcare	<ul style="list-style-type: none"> noise and SNR of aorta CNR of coronary arteries image quality assessed by three radiologists presence of significant luminal stenosis (criterion standard: ICA) 	<ul style="list-style-type: none"> less noise in DLIR compared to ASiR higher quality in DLIR compared to ASiR no differences in sensitivity, specificity, and diagnostic accuracy between ASiR and DLIR
Greffier J et al. (2019) [31]	phantoms	n/a	<ul style="list-style-type: none"> FBP ASiR-V-50 % ASiR-V-100 % DLIR-L DLIR-M DLIR-H by GE Healthcare	<ul style="list-style-type: none"> noise by NPS spatial resolution by TTF detectability index at different dose levels 	<ul style="list-style-type: none"> less noise in DLIR higher spatial resolution in DLIR higher detectability for small low-contrast lesions in DLIR, comparable detectability for other lesions
Heinrich A et al. (2020), submitted manuscript [33]	patients	100	<ul style="list-style-type: none"> ASiR-V-50 % DLIR-H 	<ul style="list-style-type: none"> SNR and CNR in abdominal aorta, liver, spleen, kidney, pelvic bone, and abdominal fat 	<ul style="list-style-type: none"> higher SNR and CNR in DLIR at equal doses
Higaki T et al. (2020) [36]	phantoms	N/A	<ul style="list-style-type: none"> FBP AiDR 3 D MBIR (FIRST) DLR (AiCE) by Canon Medical Systems	<ul style="list-style-type: none"> noise (NPS) spatial resolution (MTF) detectability 	<ul style="list-style-type: none"> less noise in DLR highest spatial resolution in MBIR highest detectability in DLR with low dose
Akagi M et al. (2019) [34]	patients	276	<ul style="list-style-type: none"> AiDR 3 D MBIR (FIRST) DLR (AiCE) by Canon Medical Systems	<ul style="list-style-type: none"> noise (SD of attenuation of paraspinal muscle) CNR of aorta, portal vein, and liver image quality on 5-point scale rated by two radiologists 	<ul style="list-style-type: none"> lowest noise in DLR highest CNR in DLR highest score for overall image quality in DLR
Narita K et al. (2020) [37]	patients	30	<ul style="list-style-type: none"> AiDR 3 D MBIR DLR (AiCE) by Canon Medical Systems	<ul style="list-style-type: none"> noise (SD of attenuation of paraspinal muscle) CNR in common bile duct overall visual image quality of the bile duct on thick-slab maximum intensity projections rated by two radiologists on 5-point confidence rating scale 	<ul style="list-style-type: none"> less noise in DLR highest CNR in DLR best overall visual image quality in DLR
Nakamura Y et al. (2019) [35]	patients	58	<ul style="list-style-type: none"> AiDR 3 D DLR (AiCE) by Canon Medical Systems	<ul style="list-style-type: none"> noise (SD of attenuation of paraspinal muscle) CNR from liver and hepatic metastasis conspicuity of smallest metastasis by two radiologists on 5-point scale 	<ul style="list-style-type: none"> lowest noise in DLR highest CNR in DLR higher conspicuity score for metastasis in DLR

AiCE = Advanced intelligent Clear-IQ Engine; ASiR = Adaptive statistical iterative reconstruction; AiDR = Adaptive iterative dose reduction; CNR = Contrast-to-noise-ratio; DLIR = Deep learning image reconstruction; DLR = Deep learning reconstruction; ICA = Invasive coronary angiography; MBIR = Model-based iterative reconstruction; MTF = Modulation transfer function; NPS = Noise power spectrum; SD = Standard deviation; SNR = Signal-to-noise-ratio; TTF = Task-based transfer function.

is important not to generalize results. Due to unique training and testing data, every algorithm might show different findings for particular imaging scenarios.

Secondly, although dose reduction potential has been shown in phantoms and patients, actual dose reduction in the clinical

routine as a result of altering acquisition parameters while performing diagnostic investigations has yet to be confirmed.

Thirdly, the decision-making process of trained algorithms is a black box to human perception. The complexity of a neural network especially in this field of image reconstruction is immense



► **Fig. 7** Comparison of a head CT detail of an 85-year-old woman reconstructed by **A** FBP, **B** ASiR-V-50% and **C** DLIR-H demonstrates improved reduction of the beam hardening artifact caused by a dental filling in the left upper quadrant. Slice thickness: 2.5 mm. W: 100, C: 40.

► **Abb. 7** Der Vergleich eines CCT-Ausschnitts einer 85-jährigen, rekonstruiert durch **A** FBP, **B** ASiR-V-50% und **C** DLIR-H, zeigt die verbesserte Reduktion von Hartstrahlartefakten, die hier durch eine Zahnfüllung im linken oberen Quadranten verursacht sind. Schichtdicke: 2,5 mm. W: 100, C: 40.

and conceptually different from human decision-making based on reasoning and memory. While we are already using DL image reconstruction in the clinical routine, computer scientists are examining methods and techniques to construct algorithms more reasonable to human understanding. Even though a DL image reconstruction algorithm might produce a correct image, it might be based on wrong reasoning. A specific lesion might be removed or blurred because it was underrepresented in the training data and thereby evaluated as noise. Conversely a non-existing lesion might be hallucinated into the reconstructed images by the reconstruction algorithm. This problem of unreliability will be difficult to solve. However, there are approaches to illuminate complex algorithms in the concept of *explainable AI* [39].

Conclusion and future outlook

In our view machine learning will inevitably affect many aspects of radiology positively. DL image reconstruction demonstrates improved image quality in terms of denoising and artifact reduction and shows potential for dose optimization. However, the implementation is not straightforward and currently in early clinical stages while the explainability and reliability of machine learning are a focus of research. While early research findings are promising, DL image reconstruction should be further introduced into clinical practice and extensively investigated to provide evidence of clinical superiority. Everything considered, after FBP and IR, reconstruction algorithms of CT images in the clinical routine are expected to move towards an ML-based approach using CNNs and possibly other upcoming methods, for example GAN. Aside from algorithm optimization, image reconstruction will additionally benefit from advancements in imaging hardware, for example photon counting CT (PCCT) [40]. As the performance of DL image reconstruction algorithms depends on the quality of the training data, there might also be an additional advancement in the future when the ground truth training data is provided by PCCT. Vice versa DL algorithms might facilitate reconstruction of those PCCT images, as currently used algorithms are presumably underachieving in terms of more complex data [40].

Conflict of Interest

F. Güttler received lecture fees from GE Healthcare. The Department of Radiology, University Hospital Jena received a research grant from GE Healthcare. The other authors declare that they have no conflict of interest.

References

- [1] Goldman LW. Principles of CT: radiation dose and image quality. *J Nucl Med Technol* 2007; 35: 213–225; quiz 226–228
- [2] Primak AN, McCollough CH, Bruesewitz MR et al. Relationship between Noise, Dose, and Pitch in Cardiac Multi-Detector Row CT. *RadioGraphics* 2006; 26: 1785–1794
- [3] Alkadhi H, Leschka S, Stolzmann P et al. *Wie funktioniert CT?* Berlin, Heidelberg: Springer Berlin Heidelberg; 2011. Im Internet: <http://link.springer.com/10.1007/978-3-642-17803-0>
- [4] Geyer LL, Schoepf UJ, Meinel FG et al. State of the Art: Iterative CT Reconstruction Techniques. *Radiology* 2015; 276: 339–357
- [5] Fleischmann D, Boas FE. Computed tomography – old ideas and new technology. *Eur Radiol* 2011; 21: 510–517
- [6] Feldkamp LA, Davis LC, Kress JW. Practical cone-beam algorithm. *J Opt Soc Am A* 1984; 1: 612
- [7] Desai GS, Uppot RN, Yu EW et al. Impact of iterative reconstruction on image quality and radiation dose in multidetector CT of large body size adults. *Eur Radiol* 2012; 22: 1631–1640
- [8] Noël PB, Walczak AM, Xu J et al. GPU-based cone beam computed tomography. *Computer Methods and Programs in Biomedicine* 2010; 98: 271–277
- [9] Gordon R, Bender R, Herman GT. Algebraic Reconstruction Techniques (ART) for three-dimensional electron microscopy and X-ray photography. *Journal of Theoretical Biology* 1970; 29: 471–481
- [10] Beister M, Kolditz D, Kalender WA. Iterative reconstruction methods in X-ray CT. *Physica Medica* 2012; 28: 94–108
- [11] den Harder AM, Willemink MJ, de Ruyter QMB et al. Achievable dose reduction using iterative reconstruction for chest computed tomography: A systematic review. *European Journal of Radiology* 2015; 84: 2307–2313
- [12] Kohli M, Prevedello LM, Filice RW et al. Implementing Machine Learning in Radiology Practice and Research. *American Journal of Roentgenology* 2017; 208: 754–760
- [13] Erickson BJ, Korfiatis P, Akkus Z et al. Machine Learning for Medical Imaging. *RadioGraphics* 2017; 37: 505–515

- [14] Chartrand G, Cheng PM, Vorontsov E et al. Deep Learning: A Primer for Radiologists. *RadioGraphics* 2017; 37: 2113–2131
- [15] Yamashita R, Nishio M, Do RKG et al. Convolutional neural networks: an overview and application in radiology. *Insights Imaging* 2018; 9: 611–629
- [16] Shan H, Zhang Y, Yang Q et al. 3-D Convolutional Encoder-Decoder Network for Low-Dose CT via Transfer Learning From a 2-D Trained Network. *IEEE Trans Med Imaging* 2018; 37: 1522–1534
- [17] Isola P, Zhu JY, Zhou T et al. Image-to-Image Translation with Conditional Adversarial Networks. arXiv:161107004 [cs] 2018. Im Internet: <http://arxiv.org/abs/1611.07004>
- [18] McBee MP, Awan OA, Colucci AT et al. Deep Learning in Radiology. *Academic Radiology* 2018; 25: 1472–1480
- [19] Lakhani P, Sundaram B. Deep Learning at Chest Radiography: Automated Classification of Pulmonary Tuberculosis by Using Convolutional Neural Networks. *Radiology* 2017; 284: 574–582
- [20] Wolterink JM, Leiner T, Viergever MA et al. Generative Adversarial Networks for Noise Reduction in Low-Dose CT. *IEEE Trans Med Imaging* 2017; 36: 2536–2545
- [21] Boedecker K. AiCE Deep Learning Reconstruction: Bringing the power of Ultra-High Resolution CT to routine imaging. 2019
- [22] Hsieh J, Liu E, Nett B et al. A new era of image reconstruction: TrueFidelityTM. Technical white paper on deep learning image reconstruction. 2019
- [23] Machine Learning for Medical Image Reconstruction: Second International Workshop, MLMIR 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 17, 2019, Proceedings. Cham: Springer International Publishing 2019. Im Internet: <http://link.springer.com/10.1007/978-3-030-33843-5>
- [24] Sprawls P. AAPM tutorial. CT image detail and noise. *RadioGraphics* 1992; 12: 1041–1046
- [25] Verdun FR, Racine D, Ott JG et al. Image quality in CT: From physical measurements to model observers. *Physica Medica* 2015; 31: 823–843
- [26] Jin KH, McCann MT, Froustey E et al. Deep Convolutional Neural Network for Inverse Problems in Imaging. *IEEE Trans on Image Process* 2017; 26: 4509–4522
- [27] Umehara K, Ota J, Ishida T. Application of Super-Resolution Convolutional Neural Network for Enhancing Image Resolution in Chest CT. *J Digit Imaging* 2018; 31: 441–450
- [28] Park J, Hwang D, Kim KY et al. Computed tomography super-resolution using deep convolutional neural network. *Phys Med Biol* 2018; 63: 145011
- [29] Zhang Y, Yu H. Convolutional Neural Network Based Metal Artifact Reduction in X-ray Computed Tomography. *IEEE Trans Med Imaging* 2018; 37: 1370–1381
- [30] Benz DC, Benetos G, Rampidis G et al. Validation of deep-learning image reconstruction for coronary computed tomography angiography: Impact on noise, image quality and diagnostic accuracy. *Journal of Cardiovascular Computed Tomography* 2020. doi:S1934592519304642
- [31] Greffier J, Hamard A, Pereira F et al. Image quality and dose reduction opportunity of deep learning image reconstruction algorithm for CT: a phantom study. *Eur Radiol* 2020. Im Internet: <http://link.springer.com/10.1007/s00330-020-06724-w>
- [32] Greffier J, Frandon J, Larbi A et al. CT iterative reconstruction algorithms: a task-based image quality assessment. *Eur Radiol* 2020; 30: 487–500
- [33] Heinrich A, Engler M, Dachoua D et al. Deep Learning Image Reconstruction in CT: Quantification of Image Quality in Clinical Practice. Manuscript submitted for publication. 2020
- [34] Akagi M, Nakamura Y, Higaki T et al. Deep learning reconstruction improves image quality of abdominal ultra-high-resolution CT. *Eur Radiol* 2019; 29: 6163–6171
- [35] Nakamura Y, Higaki T, Tatsugami F et al. Deep Learning-based CT Image Reconstruction: Initial Evaluation Targeting Hypovascular Hepatic Metastases. *Radiology: Artificial Intelligence* 2019; 1: e180011
- [36] Higaki T, Nakamura Y, Zhou J et al. Deep Learning Reconstruction at CT: Phantom Study of the Image Characteristics. *Academic Radiology* 2020; 27: 82–87
- [37] Narita K, Nakamura Y, Higaki T et al. Deep learning reconstruction of drip-infusion cholangiography acquired with ultra-high-resolution computed tomography. *Abdom Radiol* 2020. Im Internet: <http://link.springer.com/10.1007/s00261-020-02508-4>
- [38] Hoeschen C. Einsatz künstlicher Intelligenz für die Bildrekonstruktion. *Radiologe* 2020; 60: 15–23
- [39] Zhang H, Dong B. A Review on Deep Learning in Medical Image Reconstruction. arXiv:190610643 [physics] 2019. Im Internet: <http://arxiv.org/abs/1906.10643>
- [40] Willeminck MJ, Persson M, Pourmorteza A et al. Photon-counting CT: Technical Principles and Clinical Prospects. *Radiology* 2018; 289: 293–312