

Commentary

An Overview of Artificial Intelligence Applications in Liver and Pancreatic Imaging

Nicolò Cardobi ^{1,*}, Alessandro Dal Palù ², Federica Pedrini ³, Alessandro Belev ³, Riccardo Nocini ⁴,
Riccardo De Robertis ¹, Andrea Ruzzenente ⁵, Roberto Salvia ⁶, Stefania Montemezzi ¹ and Mirko D'Onofrio ³

- ¹ Radiology Unit, Department of Pathology and Diagnostics, University Hospital of Verona, Piazzale Aristide Stefani, 1, 37126 Verona, Italy; riccardo.derobertis@hotmail.it (R.D.R.); stefania.montemezzi@aovr.veneto.it (S.M.)
- ² Department of Mathematical, Physical and Computer Sciences, University of Parma, 43121 Parma, Italy; aledalpa@gmail.com
- ³ Department of Radiology, G.B. Rossi University Hospital, University of Verona, 37129 Verona, Italy; federica.pedrini@studenti.univr.it (F.P.); ale.belev@gmail.com (A.B.); mirko.donofrio@univr.it (M.D.)
- ⁴ Otolaryngology-Head and Neck Surgery Department, University Hospital of Verona, Piazzale Aristide Stefani, 1, 37126 Verona, Italy; riccardo.nocini@gmail.com
- ⁵ Department of Surgery, General and Hepatobiliary Surgery, University Hospital G.B. Rossi, University and Hospital Trust of Verona, 37126 Verona, Italy; andrea.ruzzenente@aovr.veneto.it
- ⁶ Unit of General and Pancreatic Surgery, Department of Surgery and Oncology, University of Verona Hospital Trust, 37126 Verona, Italy; roberto.salvia@univr.it
- * Correspondence: nicolo.cardobi@gmail.com



Citation: Cardobi, N.; Dal Palù, A.; Pedrini, F.; Belev, A.; Nocini, R.; De Robertis, R.; Ruzzenente, A.; Salvia, R.; Montemezzi, S.; D'Onofrio, M. An Overview of Artificial Intelligence Applications in Liver and Pancreatic Imaging. *Cancers* **2021**, *13*, 2162. <https://doi.org/10.3390/cancers13092162>

Academic Editor:
Christine Decaestecker

Received: 23 February 2021
Accepted: 28 April 2021
Published: 30 April 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Simple Summary: Artificial intelligence (AI) is gaining more and more attention in radiology. The efficiency of AI-based algorithms to solve specific problems is, in some cases, far superior compared to human-driven approaches. This is particularly evident in some repetitive tasks, such as segmentation, where AI usually outperforms manual approaches. AI may be also used in quantification where it can provide, for example, fast and efficient longitudinal follow up in liver tumour burden. AI, thanks to the association with radiomic and big data, may also suggest a diagnosis. Finally, AI algorithms can also reduce scan time, increase image quality and, in the case of computed tomography, reduce patient dose.

Abstract: Artificial intelligence (AI) is one of the most promising fields of research in medical imaging so far. By means of specific algorithms, it can be used to help radiologists in their routine workflow. There are several papers that describe AI approaches to solve different problems in liver and pancreatic imaging. These problems may be summarized in four different categories: segmentation, quantification, characterization and image quality improvement. Segmentation is usually the first step of successive elaborations. If done manually, it is a time-consuming process. Therefore, the semi-automatic and automatic creation of a liver or a pancreatic mask may save time for other evaluations, such as quantification of various parameters, from organs volume to their textural features. The alterations of normal liver and pancreas structure may give a clue to the presence of a diffuse or focal pathology. AI can be trained to recognize these alterations and propose a diagnosis, which may then be confirmed or not by radiologists. Finally, AI may be applied in medical image reconstruction in order to increase image quality, decrease dose administration (referring to computed tomography) and reduce scan times. In this article, we report the state of the art of AI applications in these four main categories.

Keywords: artificial intelligence; machine learning; deep learning; liver imaging; pancreatic imaging

1. Introduction

Artificial intelligence (AI) is one of the most promising fields of research to date. The applications of AI-based algorithms in medicine include drug development, health

monitoring, disease diagnosis, and personalized medical treatment [1–4]. Under the broad definition of AI, there are significantly different models, paradigms and implementations that differ in requirements and performance. Machine learning (ML) was the first computer science research field where a system was designed to discriminate or predict features based on algorithms that mimic human decisional processes and rely on statistical models.

One of the main limitations of such an approach was the need to perform feature extraction, where the features need to be defined a priori. However, even for experts, it may be difficult to define features that are distinctive of an object [5]. To overcome this limitation, the more recent deep learning (DL) approaches can learn from data with no need to define such features a priori [6].

DL network architecture is based on modelling an artificial neural network (ANN) to perform one specific task. ANN simulates biologic neuronal systems, composed of multiple artificial neurons. Every single unit receives an input and, by means of an activation function, emulates an action potential that leads to an output. This output may serve as input to other neurons. In modern ANNs, multiple artificial neurons are organized in several layers, called hidden layers to form a DL network [7,8]. This structure creates a feedforward stream from the ANN input to the output.

A convolutional neural network (CNN) is a subgroup of ANN with input composed of images that are particularly suitable in image recognition tasks, such as LeNet, the first CNN back in 1998 [9] and the winner of the ImageNet challenge in 2012 [10]. For this reason, CNNs were broadly used in radiology in various tasks, which may be grouped into four main categories: segmentation, quantification, characterization and AI image reconstruction/image quality improvement. In the following sections, we provide details for each of the mentioned categories for applications to liver and pancreatic imaging.

Before diving into AI applications in pancreatic and liver imaging, let us briefly mention two flavours encountered in learning. Supervised learning [11] requires a training set as training input where the goal is to build a network that classifies unknown inputs. The training set contains examples of network input and corresponding expected output (typically a classification label, a regression value and/or a binary segmentation of an image). Such a scenario allows one to train the network; the control weights are modified in order to make the network's outputs converge on the training set outputs. Optimization algorithms, based on gradient descent [12], are used to revise the weights. Supervised learning provides high quality networks but requires an annotated input that may be time-consuming to set up, especially for radiology applications.

On the other side, the goal of unsupervised networks is to derive consistent features that are common in the dataset without the need to specify them as an input. The network learns the characteristics of clusters of similar data. Common applications are feature selection and reduction of the size of a problem, through dependent variables elimination.

We refer the interested reader to a survey on deep learning in medical images [13].

2. Segmentation

In segmentation, AI algorithms are used to perform semi-automatic or automatic segmentation of a given organ. The gold standard is usually a database of manually segmented organs, called ground truth, to compare the AI segmented ones. The comparisons are performed by using overlap-based, size-based, boundary-distance-based or boundary-overlap based methods [14]. The most used is the similarity coefficient (Dice–Sørensen coefficient, DSC) [15,16], which numerically describes the similarity between two groups (ground truth vs. segmentation obtained by an algorithm). In other words, DSC numerically represents the spatial overlap or the percentage of voxel in common between two segmentations (DSC = 1, complete overlap; DSC < 1 and > 0, partial overlap; DSC = 0, no overlap). DCS may be used to compare manual vs. manual, manual vs. automatic and automatic vs. automatic segmentation. Zou et al. [17] defined DSC as a special case of kappa statistics, a widely used and reliable agreement index. As recommended by Zijdenbos et al. [18], a good overlap occurs with a DSC > 0.7.

The performance of AI algorithms in liver segmentation is reported to be very good, exceeding a DSC of 0.9 with a few dozen cases used as training [19,20]. Using fewer cases for training and testing the algorithm leads to a lower DSC score. Kavur et al. tested various AI liver segmentation models on 20 CTs (8 training and 12 test CTs), and found a DSC in the range of 0.79–0.74 in the top four methods. When increasing the number of CTs in the dataset, up to over 800 exams, the DSC exceeded 0.97 [21]. Interestingly, the DSC was not significantly different across liver conditions (normal liver, fatty liver disease, non-cirrhotic liver disease, liver cirrhosis and post-hepatectomy), so the algorithm developed was robust across possible liver morphology variations. The liver volume variation between deep learning and ground truth was $-0.2 \pm 3.1\%$. Moreover, Ahn et al. found the segmentation performance of the developed algorithm was also very good on CT exams performed elsewhere, compared to CTs acquired in their institution (DSC 0.98 vs. 0.98, $p = 0.28$) [21].

An AI algorithm may also be trained to detect and segment liver lesions. In this case, compared to the ground truth, the DSC was generally lower than the whole liver segmentation. Automated detection of hepatocellular carcinoma (HCC) on a dataset of 174 MR and 231 lesions (respectively 70% train, 15% test and 15% validation) leads to a validation/test DSC of 0.91/0.91 for liver and 0.64/0.68 for lesions [22]. The increased algorithm performance may be obtained by increasing the number of exams used for training or by introducing some optimisation in the algorithm itself. Qayyum et al. proposed a hybrid 3D residual network with a squeeze-and-excitation. This approach, applied on Liver Tumour Segmentation Challenge (LiTS, 200 CT in total), leads to a DSC of 0.95 and 0.81 respectively for liver and tumour [19].

AI segmentation of the pancreas is a bit more challenging compared to the liver, which is probably due to the complex morphology of the pancreas and the higher individual variability of this organ. Bagheri et al. investigated the technical and clinical factors that may affect the success rate in pancreas DL segmentation. They identified five parameters, all in relationship with fat (such as body mass index and visceral fat), which better delineate organ profiles [23]. Generally, the DSC of pancreas segmentation is lower than the liver one, as reported in papers who perform both liver and pancreas AI assisted segmentation (Figure 1).

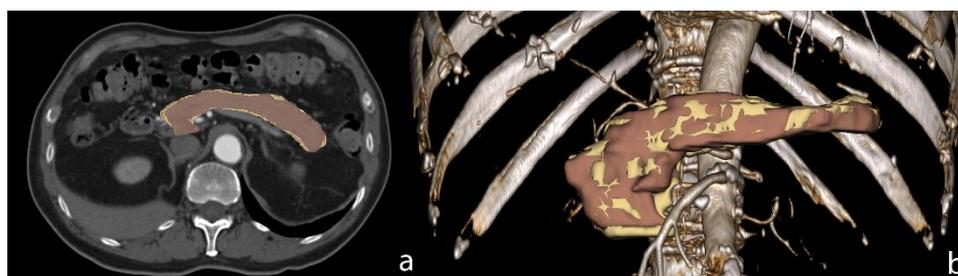


Figure 1. Pancreas AI segmentation (yellow) versus manual segmentation (brown) on axial CT image (a) and Volume Rendering image (b). The DSC of this case was 0.86.

A fully automated multiorgan segmentation on 102 MR databases (66 train, 16 validation and 20 test) led to a DSC of 0.96 for liver and 0.88 for pancreas [24]. A similar approach was also used in a CT dataset (66 for validation and 16 for test), with a DSC of 0.95 and 0.79 for liver and pancreas respectively [25].

Another aspect to consider in AI segmentation approaches, is the time required for segmentation. Accurate manual segmentation requires a lot of time, especially for big organs like the liver or complex morphology organs like the pancreas. Once trained, an AI algorithm can outperform manual segmentation in term of time required. A whole liver may be extracted in a few seconds, up to 0.04 s per CT slice, with a DSC of 0.95 [26]. In a polycystic liver and kidney disease CT series, an AI algorithm segmented the liver at 8333 slice/hour, compared to manual segmentation of an expert, which did not exceed 16 slice/hour, with a DSC of 0.96 [27]. In our small series of 39 pre- and post-chemotherapy liver CT segmenta-

tions, mean manual segmentation time was 660 s. AI-assisted segmentation, by means of NVIDIA AI assisted annotation client (NVIDIA Corporation, Santa Clara, CA, USA), was 6.70 times faster, with a mean segmentation time of 98 s ($p < 0.0001$, Figure 2). The DSC of this model was 0.956 [28].

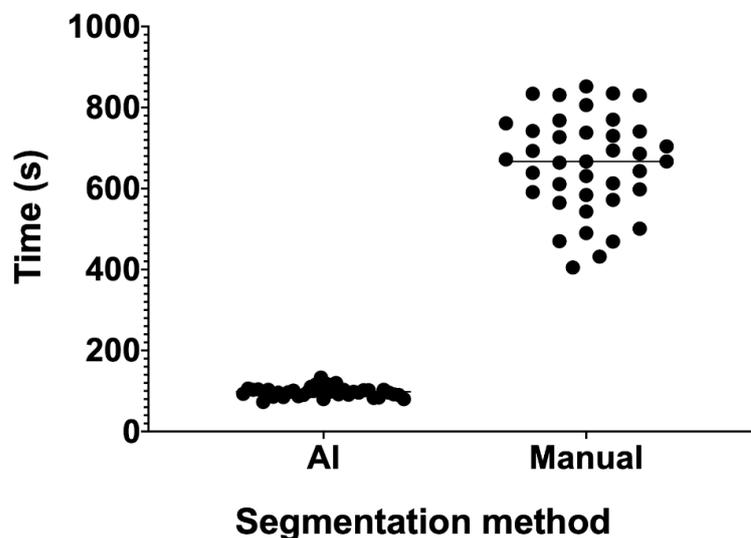


Figure 2. Liver segmentation time for an AI assisted approach and manual approach. With the AI assisted approach, there is also a lower time variance compared to the manual assisted one. This is probably due to the lower sensitivity of the AI algorithm to liver anatomical variations between subjects.

3. Quantification

AI algorithms may be useful to perform quantification of a particular parameter. For example, a segmentation may be useful to obtain organ volumetry to evaluate the response of a therapy. In our small cohort of 27 patients with autoimmune pancreatitis, we used an AI assisted segmentation approach to measure pancreas volume on 3D T1-weighted MR images, respectively at diagnosis (T0, mean volume 88.25 mL), post steroid therapy (T1, mean volume 62.85 mL) and at first follow-up (T2, mean volume 55.30 mL) (Figure 3a). Like that previously observed for liver segmentation, AI segmentation was 2.38 times faster than the manual approach (48.40 vs. 115.10 s, $p < 0.0001$; Figure 3b).

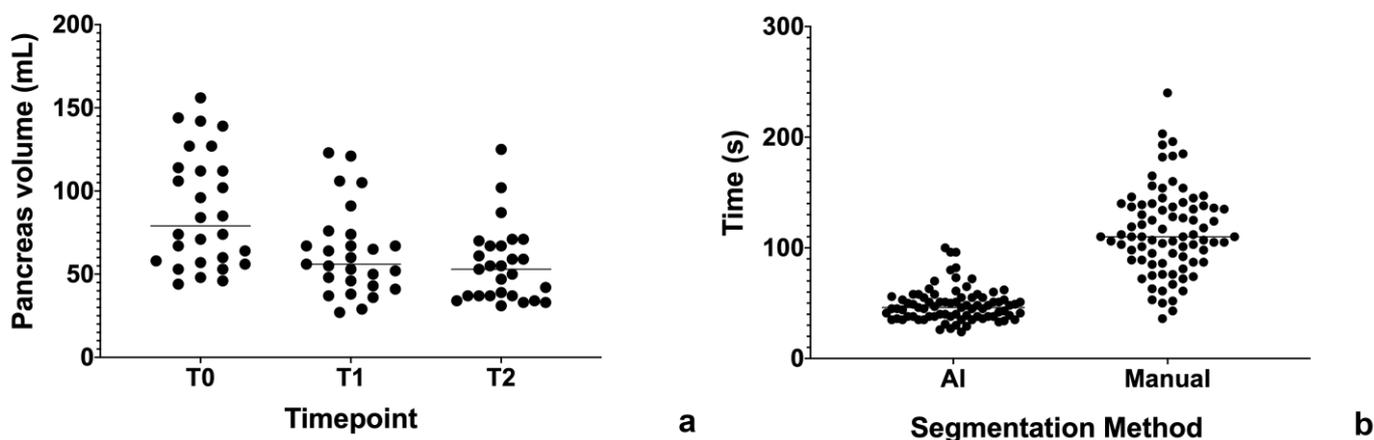


Figure 3. (a) Time required for pancreas segmentation. (b) Pancreas volume across timepoints (T0 vs. T1, $p = 0.0034$; T1 vs. T2, $p = 0.2650$; T0 vs. T2, $p = 0.0001$).

In regards to the liver, there are several papers regarding quantification of liver steatosis and fibrosis in various imaging modalities. Treacher et al. used a CNN approach to correlate grey scale elastography image texture with shear wave velocity (SWV), and they

did not find a statistically significant association [29]. On the other hand, Schawkat et al. used a ML approach to classify patients in two categories, respectively low- and high-stage fibrosis based on T1-weighted imaging derived texture parameters. Using histopathology as a gold standard, the percentage of correct assessment of the ML approach was 85.7% with an area under the curve (AUC) of 0.82, compared to 0.92 for magnetic resonance elastography (MRE) [30]. Another study showed similar results, but using a gadoteric acid-enhanced MRI, with an increasing AUC in more severe liver fibrosis [31]. Similar to fibrosis, AI algorithms were applied to quantify liver steatosis. Cao et al. used DL quantitative analysis in non-alcoholic fatty liver disease (NAFLD) applied to 2D ultrasound imaging. They found this approach fairly good in identifying NAFLD (AUC > 0.7) and very promising in distinguishing moderate and severe NAFLD (AUC = 0.958) [32]. Liver steatosis was also estimated on liver-enhanced CT, where DL was used as a segmentation tool. The segmented ROIs were then used as an input to calculate liver density. The AUC was found to increase with the fat fraction threshold, reaching an AUC of 0.962 with a fat fraction threshold >15% [33]. All of these instruments may be useful in order to detect and monitor chemotherapy-induced liver steatosis and long-term fibrosis [34–37]. In our small cohort of patients with resectable pancreatic cancer in a phase II study of liposomal irinotecan with 5-fluorouracil, leucovorin and oxaliplatin (nITRO) [38], we found chemotherapy-induced steatosis after 6 months of chemotherapy (Figure 4a). Interestingly, liver volume was not statistically significantly different across timepoints (Figure 4b). This may be a sign of drug-induced hepatotoxicity, together with steatosis.

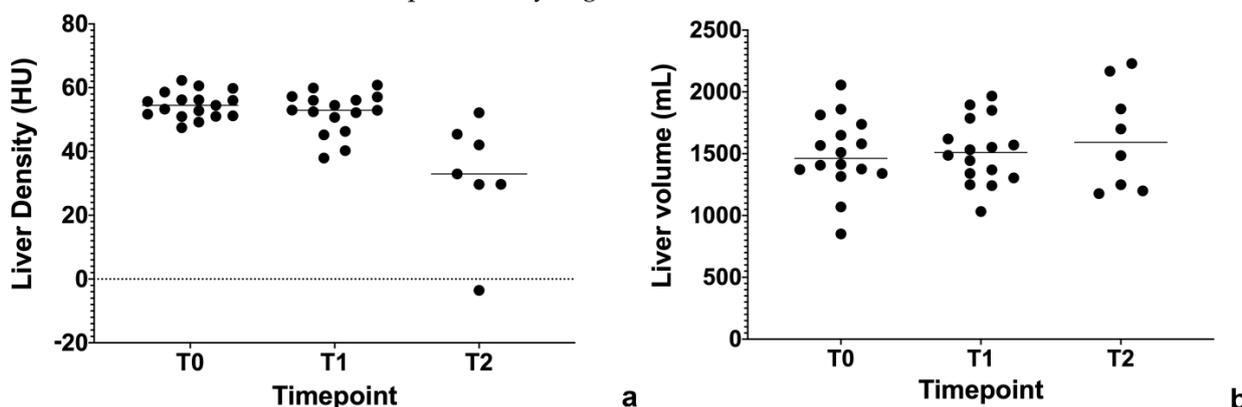


Figure 4. (a) Liver density (Hounsfield Unit, HU) across timepoints. Baseline (T0, 54.54 HU), 3 months follow-up (T1, 52.02 HU), 6 months follow-up (T2, 32.62 HU). T0 vs. T1, $p = 0.1958$; T1 vs. T2, $p = 0.0009$; T0 vs. T2, $p < 0.0001$. (b) Liver volume across timepoints.

Another interesting application of AI algorithms in the oncology setting is the longitudinal monitoring of the therapy response. These algorithms may be applied both to CT and MR. Vivanti et al. proposed an automatic approach to longitudinally detect new liver tumours, using the baseline scan as a reference. Their approach reaches a true positive new tumour detection rate of 86%, compared to 72% for stand-alone detection, with a tumour burden volume overlap error of 16% [39]. A similar approach was used to estimate liver tumour burden in neuroendocrine neoplasia on consecutive MRs. The DL approach was concordant with the radiologists' manual assessment in 91%, with a sensitivity of 0.85 and specificity of 0.92. The DSC lesion coefficient was 0.73–0.81 [40].

4. Characterization and Diagnosis

Automatic lesion characterization is another promising AI field of research. A CNN developed with a database of 494 liver lesions (434 train and 60 test, respectively) studied with MR leads to a correct classification in 92% of cases, with a sensitivity and specificity of 92% and 98% respectively. Focusing on HCC, the AUC was 0.992. Interestingly, the computational time per lesion was only 5.6 ms, so the integration in the clinical workflow may be very time efficient [41]. Referring to HCC in a clinical setting, another paper

evaluated the performance of DL applied to the Liver Imaging Reporting and Data System (LI-RADS) [42]. This model, based on MR images, reaches a correct classification rate of 90% and an AUC of 0.95 compared to radiologists in the differentiation between LI-RADS 3 and LI-RADS 4–5 lesions [43].

AI characterization algorithms were also developed using CT imaging as an input. Cao et al. proposed a multiphase CNN network that considers four-phase dynamic contrast enhanced CT (DCE-CT). The AUC in differentiating 517 (410 train, 107 test) focal liver lesions from each other (HCC, metastases, benign non-inflammatory lesion and abscess) ranged from 0.88 to 0.99 [44]. DL was found to be useful for comparing the performance differences between three-phase compared to four-phase DCE-CT in differentiation of HCC from other focal liver lesions. Shi et al. found that the three-phase DCE-CT protocol (arterial, portal-venous and delayed), combined with a convolutional dense network, had an AUC of 0.920 in differentiating HCC from other focal liver lesions compared to an AUC of 0.925 of the model with the four-phase protocol ($p = 0.765$), thus potentially reducing radiation dose [45].

In regard to the pancreas, the importance of detecting a possible pancreatic adenocarcinoma [46] as soon as possible is well known, and consequently, developing a computer-aided detection tool may be very useful for radiologists and clinicians. A recently published paper proposed a CNN-based analysis to classify cancer and non-cancer patients, based on CT imaging. This approach, revealed to be very promising, with high sensitivity and specificity. Despite the CNN model missing three pancreatic cancers (1.7%, 1.1–1.2 cm), it was more efficient than a radiologist (7% missing rate) [47]. Similar results were reported for the diagnosis of malignancies in intraductal papillary mucinous neoplasm (IPMN) based on EUS imaging. The DL approach demonstrated a sensitivity, specificity and correct classification rate of 95.7%, 92.6% and 94.0%, respectively [48]. In conclusion, if correctly developed and trained, AI algorithms may outperform humans in a specific task.

AI models may also be useful in aiding the radiologist in the diagnosis of some relevant features of a given exam, like the semiautomatic evaluation of arterial infiltration by pancreatic adenocarcinoma. Our research group are working on AI assisted abdominal artery automatic classification [49], including anatomical variants, and automatic artery shape modification induced by pancreatic adenocarcinoma (Figure 5).

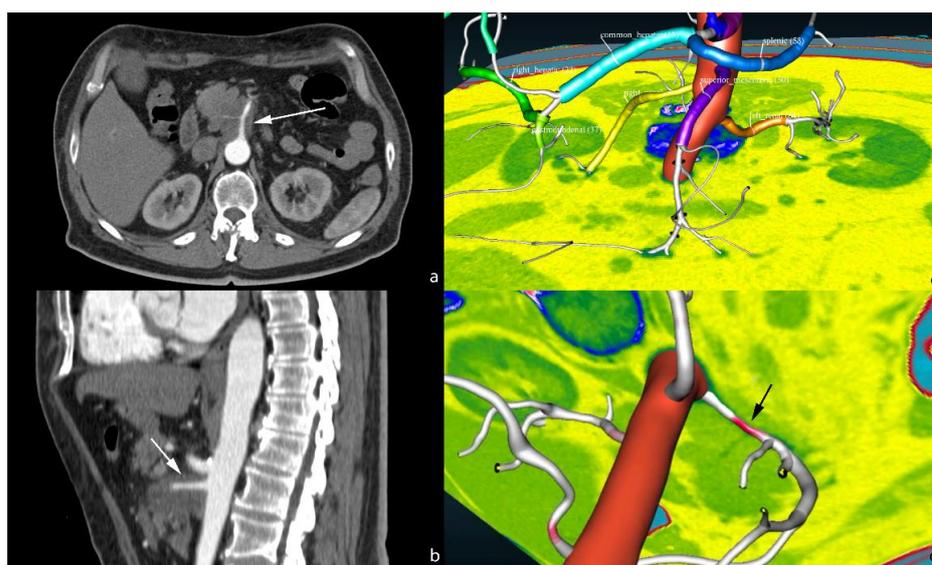


Figure 5. Axial (a) and sagittal (b) CT scan showing the infiltration of the superior mesenteric artery by pancreatic adenocarcinoma (white arrow). 3D view of automatic vessel labelling (c). Highlighting of the stenosis on the superior mesenteric artery in red (d, black arrow). The software may be configured to automatically output the length and entity of stenosis due to the tumour infiltration.

5. Reconstruction and Image Quality Improvement

AI has been shown to be useful in biomedical image reconstruction as confirmed by already available commercial tools by major vendors [50–53]. CTs reconstructed by means of DL are less noisy, with higher spatial resolution and with improved detectability, compared to the standard iterative reconstruction (IR) algorithm on phantom and real abdominal CT scans [54,55]. This was confirmed by Park et al., where two radiologists evaluated this kind of reconstruction on different vessels and abdominal organs, liver included. The contrast-to-noise ratio (CNR), signal-to-noise ratio (SNR) and sharpness, were significantly higher for the DL reconstruction [56]. Similar results were found by Akagi et al. on abdominal ultra-high-resolution CT. However, it is interesting to report that they found significant differences in attenuation values of each organ among different reconstruction algorithms, albeit small [57]. Another application of DL reconstruction on ultra-high-resolution CT was investigated in drip infusion cholangiography (DIC-CT). DL DIC-CT outperformed the other image reconstruction algorithms [58].

DL reconstruction on CT also offers the opportunity to reduce exposure dose [59], and this is particularly relevant in oncologic and paediatric patients. Lee et al. analysed the combination of dual-energy CT with DL reconstruction. They found a reduction of 19.6% of CT dose index and 14.3% of iodine administration [60]. A marked reduction of radiation dose, up to 76%, was also found by Cao et al., with the application of the DL reconstruction method to low dose contrast enhanced CT in patients with hepatic lesions [61].

DL reconstruction was also applied in MR imaging. In a study proposed by Hermann et al., DL reconstruction was used to accelerate the acquisition of T2-weighted (T2W) imaging of the upper abdomen. The DL acquisition required a single breath hold (16 s), compared to multiple breath hold of traditional single shot triggered T2W (1:30 min) and non-cartesian respiratory triggered T2W (4:00 min). The DL T2W was rated superior to conventional T2W but inferior to non-cartesian T2W. The noise, sharpness and artifacts were not statistically different between DL T2W and non-cartesian T2W [62]. DL was also used to reduce respiratory artifacts, and this led to an increase of liver lesion conspicuousness without removing anatomical details [63]. DL applied to MR may reduce acquisition time and improve image quality. However, further investigations are required, as demonstrated by Antun et al. [64]. This work investigated the drawback of DL reconstruction, represented by severe image artefacts and structural change.

6. Limitations

There are several drawbacks in the use of AI in radiology, and liver and pancreatic imaging are no exceptions. First of all, the development of AI models requires a lot of data in the form of patient imaging studies. This raises some concerns about the ownership of the data and how it is used for research. For this reason, informed consent must be obtained from the patients and data privacy needs to be guaranteed, according to local law (for example, EU General Data Protection and Regulation; GDPR) [65].

After meeting all the informed consent, privacy and data protection requirements, it is possible to proceed to data collection and preparation. In supervised learning, a large dataset of accurately labeled data is required. This is a very time intensive task and so, in order to reduce development time, a smaller dataset may be used. Despite the good performance of an AI model developed with a small dataset, there are some risks of data bias. Using a single center small dataset to develop an AI model may produce unreliable results if the model developed is applied to a different population [66,67]. Moreover, every AI model ground truth should be generated by more than one radiologist in order to increase the accuracy of the model itself [68,69]. A large multicenter dataset, together with an accurate ground truth, is necessary to mitigate the problem of overfitting in AI models, where performance of the models decreases if applied to new datasets, compared to the training datasets [70].

Finally, before clinical application, AI models need to be externally validated. Unfortunately, the majority of papers in the literature are about the feasibility of AI approaches, without external validation [71,72].

7. Future Perspectives

To date, the vast majority of AI models act as a “Black Box”. For every given input, the model produces an output, whose correctness relies on accuracy of the model itself. Usually, there is no way to understand why this output is generated. This is a problem, especially in medical imaging, where the output may be a part of the radiology report and it is necessary to understand how the model made its decision.

To overcome this limitation, a new branch of AI is under development, the Explainable AI (XAI) [73,74]. XAI analyses report human understandable features that are responsible for the AI model output [75]. Couteaux et al., using the LiTS CT liver tumour database, analysed not only the performance of DL liver tumour segmentations but also what influenced the output of the algorithm. They found the DL model was sensitive to focal liver density change and shape of the lesions [76].

8. Conclusions

In accordance with the literature, AI algorithms in liver and pancreas medical imaging are already a reality and may be useful to radiologists to speed up repetitive tasks such as segmentation, acquire new quantitative parameters such as lesion volume and tumor burden, improve image quality, reduce scanning time, and optimize imaging acquisition.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Amisha Malik, P.; Pathania, M.; Rathaur, V.K. Overview of artificial intelligence in medicine. *J. Fam. Med. Prim. Care* **2019**, *8*, 2328–2331. [[CrossRef](#)] [[PubMed](#)]
2. Rieder, T.N.; Hutler, B.; Mathews, D.J.H. Artificial Intelligence in Service of Human Needs: Pragmatic First Steps Toward an Ethics for Semi-Autonomous Agents. *AJOB Neurosci.* **2020**, *11*, 120–127. [[CrossRef](#)] [[PubMed](#)]
3. Cavasotto, C.N.; Di Filippo, J.I. Artificial intelligence in the early stages of drug discovery. *Arch. Biochem. Biophys.* **2020**, *698*, 108730. [[CrossRef](#)] [[PubMed](#)]
4. Weiss, J.; Hoffmann, U.; Aerts, H.J. Artificial intelligence-derived imaging biomarkers to improve population health. *Lancet Digit. Health* **2020**, *2*, e154–e155. [[CrossRef](#)]
5. Erickson, B.J.; Korfiatis, P.; Akkus, Z.; Kline, T.L. Machine learning for medical imaging. *Radiographics* **2017**. [[CrossRef](#)] [[PubMed](#)]
6. Chartrand, G.; Cheng, P.M.; Vorontsov, E.; Drozdal, M.; Pal, C.J.; Kadoury, S.; Tang, A. Deep learning: A primer for radiologists. *Radiographics* **2017**. [[CrossRef](#)] [[PubMed](#)]
7. Liu, W.; Wang, Z.; Liu, X.; Zeng, N.; Liu, Y.; Alsaadi, F.E. A survey of deep neural network architectures and their applications. *Neurocomputing* **2017**. [[CrossRef](#)]
8. Lecun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)]
9. LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**. [[CrossRef](#)]
10. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. *Commun. ACM* **2017**. [[CrossRef](#)]
11. Schmidhuber, J. Deep Learning in neural networks: An overview. *Neural Netw.* **2015**. [[CrossRef](#)]
12. Du, S.S.; Lee, J.D.; Li, H.; Wang, L.; Zhai, X. Gradient descent finds global minima of deep neural networks. In Proceedings of the 36th International Conference on Machine Learning, ICML, Long Beach, CA, USA, 10–15 June 2019.
13. Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; van der Laak, J.A.W.M.; van Ginneken, B.; Sánchez, C.I. A survey on deep learning in medical image analysis. *Med. Image Anal.* **2017**. [[CrossRef](#)] [[PubMed](#)]
14. Yeghiazaryan, V.; Voiculescu, I. Family of boundary overlap metrics for the evaluation of medical image segmentation. *J. Med. Imaging* **2018**, *5*, 1. [[CrossRef](#)]
15. Dice, L.R. Measures of the Amount of Ecologic Association Between Species. *Ecology* **1945**. [[CrossRef](#)]
16. Taha, A.A.; Hanbury, A. Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool. *BMC Med. Imaging* **2015**, *15*, 1–28. [[CrossRef](#)]
17. Zou, K.H.; Warfield, S.K.; Bharatha, A.; Tempany, C.M.C.; Kaus, M.R.; Haker, S.J.; Wells, W.M., III; Jolesz, F.A. Statistical Validation of Image Segmentation Quality Based on a Spatial Overlap Index. *Acad. Radiol.* **2004**. [[CrossRef](#)]

18. Zijdenbos, A.P.; Dawant, B.M.; Margolin, R.A.; Palmer, A.C. Morphometric Analysis of White Matter Lesions in MR Images: Method and Validation. *IEEE Trans. Med. Imaging* **1994**. [[CrossRef](#)] [[PubMed](#)]
19. Qayyum, A.; Lalonde, A.; Meriaudeau, F. Automatic segmentation of tumors and affected organs in the abdomen using a 3D hybrid model for computed tomography imaging. *Comput. Biol. Med.* **2020**, *127*, 104097. [[CrossRef](#)] [[PubMed](#)]
20. Hu, P.; Wu, F.; Peng, J.; Liang, P.; Kong, D. Automatic 3D liver segmentation based on deep learning and globally optimized surface evolution. *Phys. Med. Biol.* **2016**, *61*, 8676–8698. [[CrossRef](#)]
21. Ahn, Y.; Yoon, J.S.; Lee, S.S.; Suk, H.I.; Son, J.S.; Sung, Y.S.; Lee, Y.; Kang, B.K.; Kim, H.S. Deep learning algorithm for automated segmentation and volume measurement of the liver and spleen using portal venous phase computed tomography images. *Korean J. Radiol.* **2020**, *21*, 987–997. [[CrossRef](#)]
22. Bousabarah, K.; Letzen, B.; Tefera, J.; Savic, L.; Schobert, I.; Schlachter, T.; Staib, L.H.; Kocher, M.; Chapiro, J.; Lin, M. Automated detection and delineation of hepatocellular carcinoma on multiphase contrast-enhanced MRI using deep learning. *Abdom. Radiol. N. Y.* **2020**. [[CrossRef](#)]
23. Bagheri, M.H.; Roth, H.; Kovacs, W.; Yao, J.; Farhadi, F.; Li, X.; Summers, R.M. Technical and Clinical Factors Affecting Success Rate of a Deep Learning Method for Pancreas Segmentation on CT. *Acad. Radiol.* **2020**, *27*, 689–695. [[CrossRef](#)] [[PubMed](#)]
24. Chen, Y.; Ruan, D.; Xiao, J.; Wang, L.; Sun, B.; Saouaf, R.; Yang, W.; Li, D.; Fan, Z. Fully automated multiorgan segmentation in abdominal magnetic resonance imaging with deep neural networks. *Med. Phys.* **2020**. [[CrossRef](#)] [[PubMed](#)]
25. Weston, A.D.; Korfiatis, P.; Philibrick, K.A.; Conte, G.M.; Kostandy, P.; Sakinis, T.; Zeinodini, A.; Boonrod, A.; Moynagh, M.; Takahashi, N.; et al. Complete abdomen and pelvis segmentation using U-net variant architecture. *Med. Phys.* **2020**. [[CrossRef](#)]
26. Fang, X.; Xu, S.; Wood, B.J.; Yan, P. Deep learning-based liver segmentation for fusion-guided intervention. *Int. J. Comput. Assist. Radiol. Surg.* **2020**, *15*, 963–972. [[CrossRef](#)] [[PubMed](#)]
27. Shin, T.Y.; Kim, H.; Lee, J.; Choi, J.; Min, H.; Cho, H.; Kim, K.; Kang, G.; Kim, J.; Yoon, S.; et al. Expert-level segmentation using deep learning for volumetry of polycystic kidney and liver. *Investig. Clin. Urol.* **2020**, *61*, 555–564. [[CrossRef](#)]
28. NVIDIA CLARA TRAIN SDK: AI-ASSISTED ANNOTATION; DU-09358-002 _v2.0; NVIDIA: Santa Clara, CA, USA, 2019.
29. Treacher, A.; Beauchamp, D.; Quadri, B.; Fetzer, D.; Vij, A.; Yokoo, T.; Montillo, A. Deep Learning Convolutional Neural Networks for the Estimation of Liver Fibrosis Severity from Ultrasound Texture. In Proceedings of the International Society for Optical Engineering, SPIE Medical Imaging, San Diego, CA, USA, 16–21 February 2019; Volume 10950.
30. Schawkat, K.; Ciritsis, A.; von Ulmenstein, S.; Honcharova-Biletska, H.; Jüngst, C.; Weber, A.; Gubler, C.; Mertens, J.; Reiner, C.S. Diagnostic accuracy of texture analysis and machine learning for quantification of liver fibrosis in MRI: Correlation with MR elastography and histopathology. *Eur. Radiol.* **2020**, *30*, 4675–4685. [[CrossRef](#)]
31. Hectors, S.J.; Kennedy, P.; Huang, K.; Stocker, D.; Carbonell, G.; Greenspan, H.; Friedman, S.; Taouli, B. Fully automated prediction of liver fibrosis using deep learning analysis of gadoteric acid-enhanced MRI. *Eur. Radiol.* **2020**. [[CrossRef](#)]
32. Cao, W.; An, X.; Cong, L.; Lyu, C.; Zhou, Q.; Guo, R. Application of Deep Learning in Quantitative Analysis of 2-Dimensional Ultrasound Imaging of Nonalcoholic Fatty Liver Disease. *J. Ultrasound Med. Off. J. Am. Inst. Ultrasound Med.* **2020**, *39*, 51–59. [[CrossRef](#)]
33. Pickhardt, P.J.; Blake, G.M.; Graffy, P.M.; Sandfort, V.; Elton, D.C.; Perez, A.A.; Summers, R.M. Liver Steatosis Categorization on Contrast-Enhanced CT Using a Fully-Automated Deep Learning Volumetric Segmentation Tool: Evaluation in 1204 Healthy Adults Using Unenhanced CT as Reference Standard. *Am. J. Roentgenol.* **2020**. [[CrossRef](#)]
34. Maor, Y.; Malnick, S. Liver Injury Induced by Anticancer Chemotherapy and Radiation Therapy. *Int. J. Hepatol.* **2013**, *2013*, 1–8. [[CrossRef](#)] [[PubMed](#)]
35. White, M.A. Chemotherapy-Associated Hepatotoxicities. *Surg. Clin. N. Am.* **2016**, *96*, 207–217. [[CrossRef](#)] [[PubMed](#)]
36. Ramadori, G.; Cameron, S. Effects of systemic chemotherapy on the liver. *Ann. Hepatol.* **2010**, *9*, 133–143. [[CrossRef](#)]
37. Reddy, S.K.; Reilly, C.; Zhan, M.; Mindikoglu, A.L.; Jiang, Y.; Lane, B.F.; Alezander, R.H.; Culpepper, W.J.; El-Kamary, S.S. Long-term influence of chemotherapy on steatosis-associated advanced hepatic fibrosis. *Med. Oncol.* **2014**, *31*, 971. [[CrossRef](#)]
38. Simionato, F.; Zecchetto, C.; Merz, V.; Cavaliere, A.; Casalino, S.; Gaule, M.; D’Onofrio, M.; Malleo, G.; Landoni, L.; Esposito, A.; et al. A phase II study of liposomal irinotecan with 5-fluorouracil, leucovorin and oxaliplatin in patients with resectable pancreatic cancer: The nITRO trial. *Ther. Adv. Med. Oncol.* **2020**, *12*, 1–21. [[CrossRef](#)] [[PubMed](#)]
39. Vivanti, R.; Szeskin, A.; Lev-Cohain, N.; Sosna, J.; Joskowicz, L. Automatic detection of new tumors and tumor burden evaluation in longitudinal liver CT scan studies. *Int. J. Comput. Assist. Radiol. Surg.* **2017**, *12*, 1945–1957. [[CrossRef](#)] [[PubMed](#)]
40. Goehler, A.; Hsu, T.H.; Lacson, R.; Gujrathi, I.; Hashemi, R.; Chlebus, G.; Szolovits, P.; Khorasani, R. Three-Dimensional Neural Network to Automatically Assess Liver Tumor Burden Change on Consecutive Liver MRIs. *J. Am. Coll. Radiol.* **2020**, *17*, 1475–1484. [[CrossRef](#)]
41. Hamm, C.A.; Wang, C.J.; Savic, L.J.; Ferrante, M.; Schobert, I.; Schlachter, T.; Lin, M.; Duncan, J.S.; Weinreb, J.C.; Chapiro, J.; et al. Deep learning for liver tumor diagnosis part I: Development of a convolutional neural network classifier for multi-phasic MRI. *Eur Radiol.* **2019**, *29*, 3338–3347. [[CrossRef](#)]
42. Chernyak, V.; Fowler, K.J.; Kamaya, A.; Kielar, A.Z.; Elsayes, K.M.; Bashir, M.R.; Kono, Y.; Do, R.K.; Mitchell, D.G.; Singal, A.G.; et al. Liver Imaging Reporting and Data System (LI-RADS) version 2018: Imaging of hepatocellular carcinoma in at-risk patients. *Radiology* **2018**, *289*, 816–830. [[CrossRef](#)]

43. Wu, Y.; White, G.M.; Cornelius, T.; Gowdar, I.; Ansari, M.H.; Supanich, M.P.; Deng, J. Deep learning LI-RADS grading system based on contrast enhanced multiphase MRI for differentiation between LR-3 and LR-4/LR-5 liver tumors. *Ann. Transl. Med.* **2020**, *8*, 701. [[CrossRef](#)] [[PubMed](#)]
44. Cao, S.-E.; Zhang, L.; Kuang, S.; Shi, W.; Hu, B.; Xie, S.; Chen, Y.; Liu, H.; Chen, S.; Jiang, T.; et al. Multiphase convolutional dense network for the classification of focal liver lesions on dynamic contrast-enhanced computed tomography. *World J. Gastroenterol.* **2020**, *26*, 3660–3672. [[CrossRef](#)]
45. Shi, W.; Kuang, S.; Cao, S.; Hu, B.; Xie, S.; Chen, S.; Chen, Y.; Gao, D.; Chen, Y.; Zhu, Y.; et al. Deep learning assisted differentiation of hepatocellular carcinoma from focal liver lesions: Choice of four-phase and three-phase CT imaging protocol. *Abdom. Radiol.* **2020**, *45*, 2688–2697. [[CrossRef](#)]
46. Pereira, S.P.; Oldfield, L.; Ney, A.; Hart, P.A.; Keane, M.G.; Pandol, S.J.; Li, D.; Grrenhalf, W.; Jeon, C.Y.; Koay, E.J.; et al. Early detection of pancreatic cancer. *Lancet Gastroenterol. Hepatol.* **2020**, *5*, 698–710. [[CrossRef](#)]
47. Liu, K.L.; Wu, T.; Chen, P.; Tsai, Y.; Roth, H.; Wu, M.; Liao, W.; Wang, W. Deep learning to distinguish pancreatic cancer tissue from non-cancerous pancreatic tissue: A retrospective study with cross-racial external validation. *Lancet Digit. Health* **2020**, *2*, e303–e313. [[CrossRef](#)]
48. Kuwahara, T.; Hara, K.; Mizuno, N.; Okuno, N.; Matsumoto, S.; Obata, M.; Kurita, Y.; Koda, H.; Toriyama, K.; Onishi, S.; et al. Usefulness of deep learning analysis for the diagnosis of malignancy in intraductal papillary mucinous neoplasms of the pancreas. *Clin. Transl. Gastroenterol.* **2019**, *10*, 1–8. [[CrossRef](#)] [[PubMed](#)]
49. Fabiano, F.; Dal, P. An asp approach for arteries classification in CT-scans? *CEUR Workshop Proc.* **2020**, *2710*, 312–326.
50. Reina, G.A.; Stassen, M.; Pezzotti, N. *White Paper Philips Healthcare Uses the Intel®Distribution of OpenVINO™ Toolkit and the Intel®DevCloud for the Edge to Accelerate Compressed Sensing Image Reconstruction Algorithms for MRI*; White paper; Intel: Santa Clara, CA, USA, 2020.
51. Hsieh, J.; Liu, E.; Nett, B.; Tang, J.; Thibault, J.; Sahney, S. *A New Era of Image Reconstruction: TrueFidelity™ Technical White Paper on Deep Learning Image Reconstruction*; White Paper (JB68676XX); GE Healthcare: Chicago, IL, USA, 2019.
52. Boedeker, K. *AiCE Deep Learning Reconstruction: Bringing the Power of Ultra-High Resolution CT to Routine Imaging*; Aquilion Precision Ultra-High Resolution CT: Quantifying diagnostic image quality; Canon Medical Systems Corporation: Otawara, Tochigi, Japan, 2017.
53. Hammernik, K.; Knoll, F.; Rueckert, D. Deep Learning for Parallel MRI Reconstruction: Overview, Challenges, and Opportunities. *MAGNETOM Flash* **2019**, *4*, 10–15.
54. Greffier, J.; Hamard, A.; Pereira, F.; Barrau, C.; Pasquier, H.; Beregi, J.P.; Frandon, J. Image quality and dose reduction opportunity of deep learning image reconstruction algorithm for CT: A phantom study. *Eur. Radiol.* **2020**, *30*, 3951–3959. [[CrossRef](#)]
55. Ichikawa, Y.; Kani, Y.; Yamazaki, A.; Nagasawa, N.; Nagata, M.; Ishida, M.; Kitagawa, K.; Sakuma, H. Deep learning image reconstruction for improvement of image quality of abdominal computed tomography: Comparison with hybrid iterative reconstruction. *Jpn. J. Radiol.* **2021**. [[CrossRef](#)] [[PubMed](#)]
56. Park, C.; Choo, K.S.; Jung, Y.; Jeong, H.S.; Hwang, J.; Yun, M.S. CT iterative vs. deep learning reconstruction: Comparison of noise and sharpness. *Eur. Radiol.* **2020**. [[CrossRef](#)]
57. Akagi, M.; Nakamura, Y.; Higaki, T.; Narita, K.; Honda, Y.; Zhou, J.; Zhou, Y.; Akino, N.; Awai, K. Deep learning reconstruction improves image quality of abdominal ultra-high-resolution CT. *Eur. Radiol.* **2019**, *29*, 6163–6171. [[CrossRef](#)] [[PubMed](#)]
58. Narita, K.; Nakamura, Y.; Higaki, T.; Akagi, M.; Honda, Y.; Awai, K. Deep learning reconstruction of drip-infusion cholangiography acquired with ultra-high-resolution computed tomography. *Abdom. Radiol.* **2020**. [[CrossRef](#)]
59. Nakamura, Y.; Narita, K.; Higaki, T.; Akagi, M.; Honda, Y.; Awai, K. Diagnostic value of deep learning reconstruction for radiation dose reduction at abdominal ultra-high-resolution CT. *Eur. Radiol.* **2021**. [[CrossRef](#)] [[PubMed](#)]
60. Lee, S.; Choi, Y.H.; Cho, J.Y.; Lee, S.B.; Cheon, J.; Kim, W.S.; Ahn, C.K.; Kim, J.H. Noise reduction approach in pediatric abdominal CT combining deep learning and dual-energy technique. *Eur. Radiol.* **2020**. [[CrossRef](#)]
61. Cao, L.; Liu, X.; Li, J.; Qu, T.; Chen, L.; Cheng, Y.; Hu, J.; Sun, J.; Guo, J. A study of using a deep learning image reconstruction to improve the image quality of extremely low-dose contrast-enhanced abdominal CT for patients with hepatic lesions. *Br. J. Radiol.* **2021**. [[CrossRef](#)] [[PubMed](#)]
62. Herrmann, J.; Gassenmaier, S.; Nickel, D.; Arberet, S.; Afat, S.; Lingg, A.; Kündel, M.; Othman, A.E. Diagnostic Confidence and Feasibility of a Deep Learning Accelerated HASTE Sequence of the Abdomen in a Single Breath-Hold. *Investig. Radiol.* **2020**. [[CrossRef](#)]
63. Kromrey, M.-L.; Tamada, D.; John, H.; Funayama, S.; Nagata, N.; Ichikawa, S.; Kühn, J.P.; Onishi, H.; Motosugi, U. Reduction of respiratory motion artifacts in gadopentate-enhanced MR with a deep learning-based filter using convolutional neural network. *Eur. Radiol.* **2020**, *30*, 5923–5932. [[CrossRef](#)]
64. Antun, V.; Renna, F.; Poon, C.; Adcock, B.; Hansen, A.C. On instabilities of deep learning in image reconstruction and the potential costs of AI. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 30088–30095. [[CrossRef](#)]
65. The new EU General Data Protection Regulation: What the radiologist should know. *Insights Imaging* **2017**. [[CrossRef](#)]
66. Toll, D.B.; Janssen, K.J., M.; Vergouwe, Y.; Moons, K.G.M. Validation, updating and impact of clinical prediction rules: A review. *J. Clin. Epidemiol.* **2008**. [[CrossRef](#)] [[PubMed](#)]
67. Zech, J.R.; Badgeley, M.A.; Liu, M.; Costa, A.B.; Titano, J.J.; Oermann, E.K. Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: A cross-sectional study. *PLoS Med.* **2018**, *15*, e1002683. [[CrossRef](#)] [[PubMed](#)]

68. Geijer, H.; Geijer, M. Added value of double reading in diagnostic radiology, a systematic review. *Insights Imaging* **2018**. [[CrossRef](#)] [[PubMed](#)]
69. Rajpurkar, P.; Irvin, J.; Ball, R.L.; Zhu, K.; Yang, B.; Mehta, H.; Duan, T.; Ding, D.; Bagul, A.; Langlotz, C.P.; et al. Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists. *PLoS Med.* **2018**. [[CrossRef](#)]
70. Mutasa, S.; Sun, S.; Ha, R. Understanding artificial intelligence based radiology studies: What is overfitting? *Clin. Imaging* **2020**. [[CrossRef](#)] [[PubMed](#)]
71. Liu, X.; Faes, L.; Kale, A.U.; Wagner, S.K.; Fu, D.J.; Bruynseels, A.; Mahendiran, T.; Moraes, G.; Shamdas, M.; Kern, C.; et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: A systematic review and meta-analysis. *Lancet Digit. Health* **2019**. [[CrossRef](#)]
72. Kim, D.W.; Jang, H.Y.; Kim, K.W.; Shin, Y.; Park, S.H. Design characteristics of studies reporting the performance of artificial intelligence algorithms for diagnostic analysis of medical images: Results from recently published papers. *Korean J. Radiol.* **2019**. [[CrossRef](#)]
73. Arrieta, A.B.; Rodriguez, N.D.; Del Ser, J.; Bennetot, A.; Tabik, S.; Barbado, A.; Garcia, S.; Gil-Lopez, S.; Molina, D.; Benjamins, R.; et al. *Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI*; Elsevier: Amsterdam, The Netherlands, 2019.
74. Samek, W.; Müller, K.R. Towards Explainable Artificial Intelligence. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Berlin/Heidelberg, Germany, 2019. [[CrossRef](#)]
75. Singh, A.; Sengupta, S.; Lakshminarayanan, V. Explainable deep learning models in medical image analysis. *J. Imaging* **2020**, *6*, 52. [[CrossRef](#)]
76. Couteaux, V.; Nempont, O.; Pizaine, G.; Bloch, I. Towards interpretability of segmentation networks by analyzing deepDreams. In *Interpretability of Machine Intelligence in Medical Image Computing and Multimodal Learning for Clinical Decision Support, Proceedings of the Second International Workshop, iMIMIC 2019, and 9th International Workshop, ML-CDS 2019, Shenzhen, China, 17 October 2019*; Lecture Notes in Computer Science; Suzuki, K., Reyes, M., Eds.; Springer: Berlin/Heidelberg, Germany, 2019; Volume 11797. [[CrossRef](#)]