

Review

Artificial Intelligence in Neuroradiology: A Review of Current Topics and Competition Challenges

Daniel T. Wagner ¹, Luke Tilmans ¹, Kevin Peng ², Marilyn Niedermeier ², Matt Rohl ³, Sean Ryan ¹, Divya Yadav ², Noah Takacs ² , Krystle Garcia-Fraley ¹, Mensur Koso ², Engin Dikici ¹, Luciano M. Prevedello ¹  and Xuan V. Nguyen ^{1,*}

¹ Department of Radiology, The Ohio State University Wexner Medical Center, Columbus, OH 43210, USA; luciano.prevedello@osumc.edu (L.M.P.)

² College of Medicine, The Ohio State University, Columbus, OH 43210, USA

³ College of Arts and Sciences, The Ohio State University, Columbus, OH 43210, USA

* Correspondence: xuan.nguyen@osumc.edu

Abstract: There is an expanding body of literature that describes the application of deep learning and other machine learning and artificial intelligence methods with potential relevance to neuroradiology practice. In this article, we performed a literature review to identify recent developments on the topics of artificial intelligence in neuroradiology, with particular emphasis on large datasets and large-scale algorithm assessments, such as those used in imaging AI competition challenges. Numerous applications relevant to ischemic stroke, intracranial hemorrhage, brain tumors, demyelinating disease, and neurodegenerative/neurocognitive disorders were discussed. The potential applications of these methods to spinal fractures, scoliosis grading, head and neck oncology, and vascular imaging were also reviewed. The AI applications examined perform a variety of tasks, including localization, segmentation, longitudinal monitoring, diagnostic classification, and prognostication. While research on this topic is ongoing, several applications have been cleared for clinical use and have the potential to augment the accuracy or efficiency of neuroradiologists.

Keywords: artificial intelligence; neuroradiology; AI-based challenge competitions; machine learning; deep learning



Citation: Wagner, D.T.; Tilmans, L.; Peng, K.; Niedermeier, M.; Rohl, M.; Ryan, S.; Yadav, D.; Takacs, N.; Garcia-Fraley, K.; Koso, M.; et al. Artificial Intelligence in Neuroradiology: A Review of Current Topics and Competition Challenges. *Diagnostics* **2023**, *13*, 2670. <https://doi.org/10.3390/diagnostics13162670>

Academic Editor: Frank Weber

Received: 16 May 2023

Revised: 7 August 2023

Accepted: 9 August 2023

Published: 14 August 2023



Copyright: © 2023 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/>).

1. Introduction

Artificial intelligence (AI), including its subsets of machine learning (ML) and deep learning (DL), is currently one of the most heavily researched fields in radiology, including neuroradiology. The results of a recent PubMed database search for “artificial intelligence” and “neuroradiology” revealed that more papers were published in 2021 and 2022 than all years prior (Figure 1). In fact, neuroradiology accounts for nearly one-third of all AI-related papers in radiology [1]. Our analysis of the text in the titles of all PubMed entries from this query from 2017 onwards shows a relatively high proportion of deep learning articles, with brain imaging, particularly imaging for stroke, representing a common topic (Figure 2). With continued research and growing capital investment in the industry, AI has the potential to revolutionize the medical imaging industry by providing improved diagnostic accuracy, increased efficiency, reduced costs, and better patient outcomes. In contrast to the initial claims that AI would render radiologists obsolete, this outlook has recently been redefined into a new paradigm—that AI will augment the modern-day radiologist and facilitate higher-quality and more efficient patient care [2,3].

This paper will provide an overview of the current topics and advancements in artificial intelligence related to the field of neuroradiology, including ischemic stroke, intracranial hemorrhage, brain tumors, neurocognitive imaging, white matter lesions, spinal imaging, and head and neck imaging. It will primarily focus on the performance of new and innovative research tied to current or recent AI-based challenge competitions,

with particular attention paid to tasks that are central to neuroradiologists and with an emphasis on the literature published in 2017 or later.

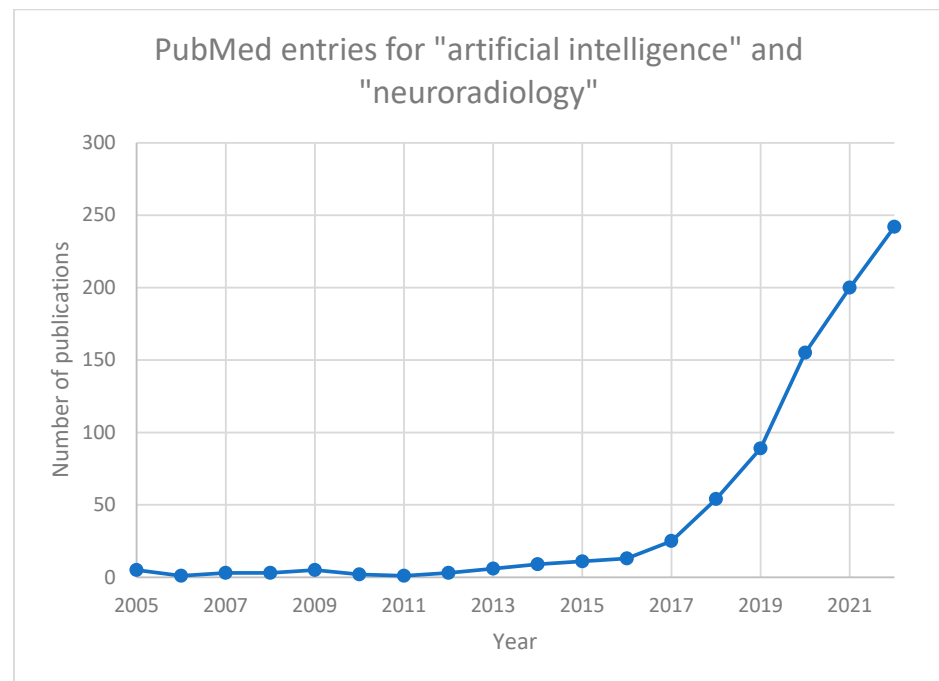


Figure 1. Counts of PubMed entries obtained when searching for “artificial intelligence” and “neuroradiology” by calendar year of publication.

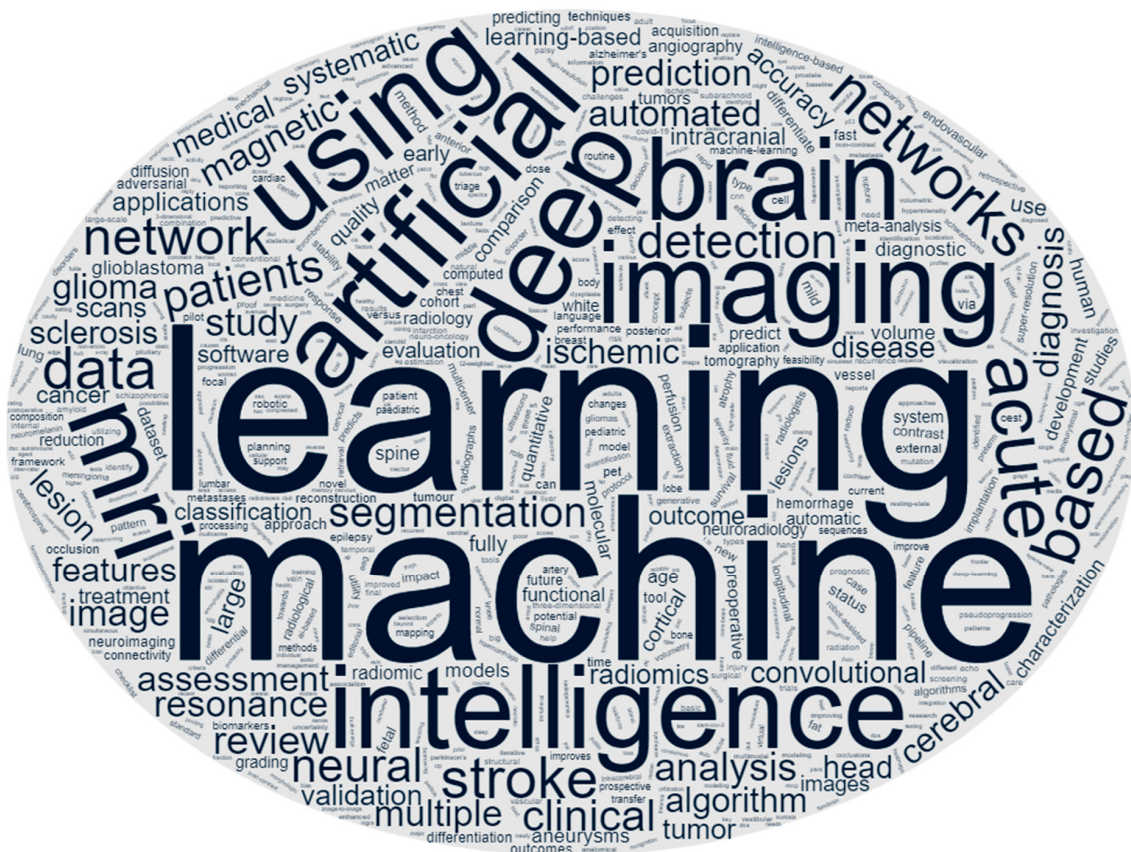


Figure 2. Word cloud depiction of terms from titles of articles from a PubMed query for “artificial intelligence” and “neuroradiology” published in 2017 or later.

2. AI Challenge Competitions

AI challenge competitions are events organized around a specific task or problem. These competitions are designed to promote the research and development of AI by providing a platform for testing various algorithms and models against current best practices and other competing AI methodologies. They provide unique opportunities for researchers and vendors to benchmark their AI-based algorithms. Many current AI biomedical imaging challenges can be found on the website Grand Challenge (<https://grand-challenge.org/>, accessed on 15 May 2023). Radiology AI challenge competitions often make use of datasets of medical images of a given modality; for neuroradiology, the most commonly used radiologic modalities are computed tomography (CT) and magnetic resonance (MR) imaging, although radiography and positron emission tomography are occasionally included. In most cases, the datasets consist of images obtained via standard clinical protocols for processing and may also include metadata related to image acquisition parameters or other demographic or clinical data. Teams that submit candidate algorithms are typically permitted to perform any additional image processing steps they deem appropriate, such as masking or scaling. The design of these competitions varies depending on the scope and goals of each competition. Typically, candidate algorithms are expected to perform a specific computer vision task, such as lesion detection, segmentation, or classification, and public training datasets are made available to participants for this purpose; their performance is evaluated using a test set that is not made public during the competition.

A critical problem with AI in radiology is the lack of large, public, high-quality, and well-annotated datasets [4,5]. Research from single institutions or individuals is often based on algorithms that are optimized to relatively small home-grown datasets with limited generalizability or reproducibility. In some situations, large amounts of data are available but may not have been curated to ensure high data quality or adequate data standardization. Differing scanner types, imaging protocols, pre-processing techniques, and patient populations make validation and large-scale implementation difficult. A systematic review by Yu et al. demonstrated that a vast majority of published DL algorithms saw a drop in performance when deployed on external data [5]. AI-based challenges attempt to circumvent this problem and provide large, well-annotated datasets with which to test and validate various algorithms. For example, when planning a challenge related to tumor segmentation, the organizers would typically seek to acquire a large number of imaging exams of tumors, ideally from different institutions and geographic areas, and have experts annotate the pixels on the images that correspond to tumor. Similarly, for a lesion detection challenge, the organizers would obtain multiple imaging exams that were annotated by experts as positive or negative for the lesion of interest. While many of these datasets are still a work in progress, in our opinion, they are currently one of the most reliable methods with which to evaluate and compare various artificial intelligence methods and advancements.

3. Definitions

Artificial intelligence is an umbrella term for any machine or system that performs tasks typically associated with human intelligence. Machine learning (ML) is a subtype of AI and refers to a technique by which programs learn patterns and/or adapt based on experience, typically through the use of large training datasets. ML can be categorized into supervised, unsupervised, and semi-supervised versions. In supervised learning, the datasets are labeled by an expert to “train” the algorithm to produce a desired output, with each piece of data refining the algorithm further. This is one of the primary AI methods used in radiologic imaging. In unsupervised learning, the datasets are not labeled. Patterns within the dataset are inferred by the AI algorithm without prior information or guidance. As its name suggests, semi-supervised learning incorporates components from both supervised and unsupervised ML by utilizing both the labeled and unlabeled datasets during the AI algorithm’s training. Deep learning (DL) is a subtype of ML that uses multi-layer artificial neural networks, which are often greater than 20 layers deep [6]. A common

DL method used in image classification and lesion detection tasks utilizes convolution neural networks (CNN), i.e., DL architectures designed for processing and analyzing images and organized like the structure of a human brain. AI approaches to segmentation often make use of variants of the U-Net architecture, which contains convolutional neural network components arranged such that an encoding arm extracts relevant features from an input image and a decoding arm subsequently converts the features into an image of similar size to the input image.

The evaluation and comparison of AI algorithms is inherently challenging. This is not only difficult due to variation in AI architectures, but also due to variation in test datasets. Fortunately, there are several performance metrics that allow for the standardization and comparison of AI algorithms, several of which are presented in this paper. Some of the metrics commonly used when discussing diagnostic accuracy include true positives (TP), false positives (FP), true negatives (TN), false negatives (FN), sensitivity $[TP/(TP+FN)]$, specificity $[TN/(FP+TN)]$, accuracy $[(TP+TN)/(TP+TN+FP+FN)]$, precision $[TP/(TP+FP)]$, Dice/F1 score $[2TP/(2TP+FP+FN)]$, and the area under the receiver operating characteristic (ROC) curve (AUC). For image data, the Dice score, equivalent to the F1 score, is a common metric for evaluating model segmentation performance by evaluating the spatial overlap (pixelwise agreement) between a gold-standard segmentation and model-derived segmentation, with scores ranging from 0 (no overlap) to 1 (exact overlap). It represents the harmonic mean of precision and sensitivity and therefore is commonly used to capture both values within one metric.

4. Ischemic Stroke

Stroke is a leading cause of death and the number one cause of serious long-term disability in the United States, with ischemic stroke accounting for most stroke types [7]. Over the last decade, several large clinical trials have increased the time window for stroke intervention up to 24 h, significantly increasing the pool of potentially treatable stroke patients, with subsequent demand for the faster and more accurate interpretation of images [8–10].

AI advancements in ischemic stroke imaging have primarily focused on optimizing stroke workflows, the detection of large vessel occlusions (LVO), quantifying stroke scoring metrics, the segmentation and quantification of ischemic or at-risk tissue, and clinical stroke outcome prediction. This section will center on stroke segmentation, ischemic stroke detection, and the automation of stroke metrics. Current AI-based competitions and challenges related to stroke segmentation will also be discussed.

ML and DL have already made a significant impact on current clinical stroke imaging practices. Currently, there are at least 14 different commercially available AI-based software packages related to ischemic stroke imaging [11–14]. Most packages have received FDA (Food and Drug Administration) or EU (European Union) approval for use in tasks such as LVO detection, ASPECTS scoring, and perfusion analysis. These software packages primarily use ML and DL methods, including convolution neural networks (CNN), with many of them built upon established well-trained deep neural network architectures such as AlexNet, GoogleNET, ResNET, or DenseNet-121 [15,16].

4.1. Segmentation and Perfusion Imaging

A major goal of stroke imaging is to evaluate and accurately segment ischemic penumbra volume from core infarct volume. MRI/DWI is the gold standard for evaluating infarct core volumes. However, it is typically slow, expensive, and not always readily available. Additionally, manual segmentation of ischemic changes is cumbersome, time-consuming, and affected by inter-rater variability [17]. CT perfusion (CTP) is commonly performed for acute ischemic workup, often in tandem with computed tomography angiography (CTA), with relative cerebral blood flow (rCBF) and time-to-maximum (Tmax) among the most commonly used CTP metrics [9,10,18]. These metrics are derived from time-attenuation curves, with applied linear statistical models and threshold values used to

develop perfusion maps and volumetric segmentation. Threshold cut-offs for core infarct and penumbra vary in the literature, but typical values include $rCBF < 30\%$ and $Tmax > 6$ s, respectively [18,19]. The prediction and segmentation of DWI core infarct volumes using CTP data is a heavily researched area in AI and remains an ongoing challenge [20,21].

Many major AI commercial software packages provide threshold-based perfusion maps and segmentation analysis via DL models. Over the last several years, multiple new architectures and methods have been developed for stroke segmentation for both CT and MR, many of which have been developed and tested during recent AI competitions and challenges.

The Ischemic Stroke Lesion Segmentation (ISLES) challenge (<https://www.isles-challenge.org/>, accessed on 15 May 2023) was created in 2015 as an open competition to develop and encourage the design of advanced tools for use in ischemic stroke analysis. The most recently completed challenge in 2018 evaluated 24 teams' ability to segment infarcted tissue from CTP images, with corresponding DWI images being the referenced standard [22]. The dataset consisted of 103 cases, including 63 training cases and 40 test cases. A comparison was also made to a conventional threshold-based method ($rCBF < 38\%$).

All the top-performing teams used various DL U-net architectures, with the winning team of Song et al. using a 3D multi-slice U-shaped network [22–25]. Nearly every team outperformed the traditional threshold-based model using $rCBF < 38\%$ (mean Dice Score 0.34). However, the mean Dice score for the top-performing team was only 0.51, illustrating the persistent gap in performance and accuracy compared to the manual segmentation of DWI imaging. Soltanpour et al. recently proposed a MultiRes U-Net technique, in which contralateral CTP imaging and Tmax heatmaps were used as additional data with which to supplement the CTP input images. This study used the ISLES 2018 dataset and achieved a Dice score of 0.69, a significant improvement from the original competition [23]. A recent meta-analysis on the performance of ML segmentation in stroke found that top-performing algorithms used DL methods as opposed to conventional ML classifiers, with a pooled Dice score of 0.50 [26].

The 2022 ISLES challenge (<https://isles22.grand-challenge.org/>, accessed on 15 May 2023) is currently underway and has the goal of segmenting infarcted brain lesions using DWI, ADC, and FLAIR images. There have been over 90 entries from multiple countries, with the current highest DICE score of 0.78. An additional separate segmentation challenge using a larger standardized dataset of T1-weighted images known as ATLAS 2.0—Anatomical Tracing of Lesions After Stroke (<https://atlas.grand-challenge.org/>, accessed on 15 May 2023)—is also underway [27]. This challenge only has several entries, with the highest DICE score currently 0.61.

4.2. Large Vessel Occlusion (LVO) Detection and Stroke Scoring Metrics

Large-vessel occlusion (LVO) and a corresponding hyperdense artery sign are strong indicators of stroke on the CT angiogram (CTA) and non-contrast CT (NCCT), respectively. Many commercially available AI-based software packages provide LVO detection [11]. These are primarily confined to assessing only anterior circulation using single-phase CTA images and often use CNN-based algorithms [15]. Over the last few years, multiple studies have evaluated the performance of LVO detection for several available software packages with varying results. Sensitivities ranged from 73–98% and specificities ranged from 52–98% [28–36]. As a reference standard, in one study, neuroradiologists' ability to detect anterior LVO using CTA and NCCT had sensitivities ranging from 75 to 88% and specificities ranging from 88 to 97% [37]. Those figures significantly increased when supplemented with CTP data. Recently, Stib et al. used multi-phase CTA and a CNN using a DenseNet-121 architecture, with sensitivities of up to 100% and specificities of 77% when using all three phases [38]. In contrast to other studies, posterior circulation and cervical ICA occlusions were included in the analysis. Just as it is for radiologists, AI-based algorithms have low sensitivity for detection of distal occlusions (i.e., M2 or M3) [31,33,35].

The Alberta Stroke Program Early CT Score (ASPECTS) is a quantitative 10-point scoring system used to assess patients who would benefit from endovascular therapy. The score is determined based on early ischemic changes visible on NCCT [39]. The scale assesses 10 regions of the MCA territory, subtracting (from a normal score of 10) one point for every affected area, and patients with a resultant score of greater than or equal to six are given priority for thrombectomy [8]. Unfortunately, an inherent problem with ASPECTS is poor inter-rater variability between radiologists [40]. Many commercially available packages provide ASPECTS scoring, several of which employ classical ML methods like random forests [14]. DL methods are also being explored, with encouraging results [41]. Several studies have demonstrated AI ASPECTS scoring performance to be equal to that of experienced neuroradiologists [42–47]. Albers et al. showed that RAPID ASPECTS was more accurate than experienced readers in identifying early ischemia when compared to the corresponding DWI results [48]. Unfortunately, AI software for LVO detection and ASPECTS scoring have difficulty interpreting images of patients with findings of chronic underlying abnormalities, such as remote infarcts, chronic white matter changes, and post-operative changes [11,49].

The various available automated tasks described in this section have potential to facilitate the objective characterization of ischemic stroke to guide emergent treatment. While the overall performance of these methods is encouraging, the role of the experienced radiologist remains vital to the final interpretation and diagnosis.

5. Intracranial Hemorrhage

In the acute clinical setting, intracranial hemorrhage (ICH) represents a potentially life-threatening situation that demands prompt and accurate detection with imaging. Failure to detect and diagnose ICH in a timely manner can result in a delay of treatment and thus higher morbidity and mortality, underscoring the importance of early and accurate detection [50–52]. Coupled with increasing imaging workloads and a shortage of qualified radiologists, the need to streamline and expedite reads is paramount. Current AI research on ICH is aimed at ICH detection, segmentation, and classification, the prediction of hemorrhagic expansion, and even workflow prioritization. This section will focus on recent AI advancements in the detection and segmentation of ICH.

5.1. Detection

Due to its high sensitivity and specificity, CT is the modality of choice for ICH detection. Many recent studies on ICH detection examined DL neural networks, and most studies evaluating specific algorithms were retrospective, often comparing the results to those obtained with experienced neuroradiologist(s). Overall, current software packages have performed well with ICH detection. For example, Heit et al. evaluated RAPID ICH detection of ICH compared to the results of 3 trained neuroradiologists, with sensitivities and specificities of 96 and 95%, respectively [53]. This study included all subtypes of hemorrhage, including subdural, epidural, subarachnoid, intraparenchymal, and intraventricular. False-negative cases were primarily associated with small-volume hemorrhaging (<1.5 mL). False positives were associated with the observation of calcifications and other hyperdense structures on CT. Colasurdo et al. used the Viz.ai software package to detect subdural hemorrhage (SDH). The package had a sensitivity and specificity of 91% and 96%, respectively, with sensitivities smaller for small chronic SDH [54]. Another study, by Seyam et al., used Aldoc Medical to detect subtypes of hemorrhage. A sensitivity of 87.2% was observed with a 97.8% NPV. The program yielded lower detection rates for subdural hemorrhage (69.2%) and subarachnoid hemorrhage (77.4%) [55]. Whereas most studies to date have used small datasets from a single institution, Matsoukas et al. in 2022 performed the first systemic review of ICH detection, summarizing and pooling the performance results of various AI algorithms over the last couple decades. This review included approximately 40 relevant studies and reported the sensitivity and specificity for ICH detection to be 92% and 94%, respectively [56].

Numerous software tools have been developed in the last decade, many with high-performance and expert-level accuracy in terms of the diagnosis of ICH. Implementing them into clinical practice could improve the quality of radiology workflow, effectively providing a “second” read or quality assurance to the reading radiologist. One major obstacle to robust performance, however, lies in single-center design and the small size of single-institution training and test datasets, limiting external validity and applicability to a larger-scale setting. One initiative to improve AI-based algorithms would be the creation of large-scale annotated datasets. AI challenge competitions provide this level playing field, allowing the nonbiased and validated performance comparison of algorithms’ performance.

The first AI ICH detection challenge took place in 2019 when the Radiological Society of North America (RSNA) collaborated with the American Society of Neuroradiology (ASNR) [57]. An 874,035-image brain hemorrhage CT dataset was pooled from historical imaging from Stanford University, Universidad Federal de Sao Paulo, and Thomas Jefferson University Hospital [58]. This dataset was annotated by 60 volunteer neuroradiologists, serving as the reference standard, with each annotator labeling each imaging set as one of the following: ICH (and its subtype), normal, or abnormal but no hemorrhage. Each participating team’s algorithm was subsequently evaluated and ranked based on its ability to detect and classify ICH. Using the 2019-RSNA batch-1 test set, the winning team had sensitivity and specificity scores of 0.950 and 0.944, respectively, for all subtypes of ICH [59]. Detection metrics were worst for SDH and best for IVH. The 2019 ICH challenge provided scale, the opportunity for comparisons, and external validity to ICH detection algorithms. It also offered full transparency into the details of the various algorithms.

5.2. Segmentation

In addition to detection, hemorrhagic volume and expansion are important predictors of outcomes and treatment responses [60]. Accurate measurements of ICH, IVH, and perihematoma edema (PHE) facilitate the prediction of morbidity and mortality [60]. A common method for ICH volume calculation is known as the ellipsoid approximation technique. This involves the product of 0.52 and manual measurements by the radiologist for the 3 orthogonal axes. One study found that this method leads to the overestimation of hemorrhage size by nearly 20%, particularly when dealing with large or irregularly shaped hematomas [61]. Other limitations include significant interrater variability, limiting reliability, and the temporal demands of manual ICH measurements.

In recent years, researchers have investigated the ability of ML models to mitigate the aforementioned shortcomings of manual ICH segmentation. Patel et al. 2019 found that a convolutional neural network was able to match radiologists’ estimation of ICH volume [61]. Islam et al. in 2019 developed a novel DL model known as ICHNet. This achieved a Dice score of 0.89, comparable to those obtained by radiologists [62]. Heit et al., using RAPID, demonstrated results that strongly correlated with the expert consensus ($r = 0.983$) [53]. DL models based on the U-Net architecture or derivatives have also been used in segmentation. Zhao et al. utilized a no-new-U-Net (nn-Unet) framework, a type of DL model without a predetermined internal organization. Instead, this type of model continuously adapts to its training data and is therefore better optimized for a wider variety of data cohorts. Zhao reported Dice scores of 0.92 for ICH, 0.79 for IVH, and 0.71 for PHE [60]. They suspected that lower scores for IVH might be related to limited test data. Kok et al. evaluated various types of algorithms used in segmentation using a large-scale multicenter database from the Tranexamic Acid for Hyperacute Primary Intracerebral Hemorrhage (TICH-2) trial [63]. The types of algorithms studied and compared were U-Net, nnUnet, BLAST-CT, and DeepLab3+. All these DL algorithms represent well-qualified, state-of-the-art technology with excellent track records in medical AI. They concluded that U-Net-based networks achieved significantly better performances than others for ICH and intraventricular hemorrhage (IVH) segmentations ($p < 0.05$). The top-performing models in their study had median Dice scores of 0.92 for ICH, 0.66 for perihematoma edema (PHE), and 1.00 for IVH. Notably, a nnU net algorithm, named Focal, achieved the highest Dice

score with IVH. Not unexpectedly, a worse performance was noted for the detection of smaller hemorrhages.

6. Brain Tumors

Primary central nervous system (CNS) tumors are relatively rare neoplasms, with potentially devastating outcomes. Intracranial metastatic disease is the most common CNS tumor overall, accounting for more than half of all brain malignancies. Gliomas and meningiomas account for two thirds of all primary CNS tumors. High-grade gliomas (glioblastomas) are an often-catastrophic malignancy, with persistently low survival rates despite our improved understanding of the disease [64,65]. On the other hand, brain metastasis survival rates and incidence rates are continuing to increase, which is likely related to a combination of improved therapeutic regimens and tumor detection [66–68]. Imaging is therefore continuing to play an ever-increasing and vital role in both the detection and monitoring of disease.

Advancements in AI for brain tumors are broad and include research into tumor detection (identifying and/or localizing a tumor within an image), segmentation (defining tumor boundaries), grading (determining aggressiveness of a tumor), prognostication (predicting future outcomes such as survival), and treatment response assessment (evaluation of pseudo-progression and post-treatment necrosis). Several research areas are related to an emerging field called radiogenomics (imaging genomics), which explores the correlations between imaging characteristics and genetic/mutational patterns, essentially providing a “virtual biopsy” based on imaging patterns and other clinical data [69]. Although radiogenomics holds great potential, this section will primarily focus on advancements in AI that are central to a neuroradiologist’s role, such as tumor detection, segmentation, and post-treatment disease evaluation.

6.1. Tumor Detection

Tumor detection and surveillance of disease is a common and important task for radiologists. Targeted therapies like stereotactic radiosurgery (SRS) have pushed radiologists to meticulously detect and track numerous CNS lesions, often adding significant time and effort for each study. In a time-constrained work environment, the potential to augment tumor detection via AI is appealing. Two recent meta-analyses by Ozkara et al. and Cho et al. demonstrated that classical ML and DL algorithms overall performed well in the detection of brain metastasis, with a pooled sensitivity of 89–90% [70,71]. As a reference, 7 board-certified radiologists and 5 resident radiologists in one study had a similar pooled sensitivity of 89% [72]. However, sensitivity was significantly worse with small lesions. This is discouraging as the detection of smaller lesions is likely to provide the most benefit for radiologists [73]. Most DL algorithms within this analysis used 3D U-net or Deep Medic architectures. While the sensitivities are encouraging, it is unlikely this technology could replace trained radiologists, but rather augment their work. In one multi-center study performed using a multi-scale CNN detection algorithm to assist radiologists, the sensitivity for radiologists to detect brain metastasis increased by 21%, with a decreased reading time of 40% [74]. This type of result is likely to be the most realistic benefit of AI in this field, namely, the development of a tool to improve the quality and efficiency of radiologists.

6.2. Tumor Segmentation

The segmentation of brain tumors involves accurately and reliably delineating brain tumors from normal tissue, other disease pathologies, and imaging artifacts. The segmentation of brain tumors is difficult due to the variable size and heterogenous appearance of CNS malignancies. Additionally, some tumors demonstrate poor contrast enhancement and can easily overlap in intensity with the surrounding healthy brain parenchyma. Tumor grading, disease prognosis, treatment options, and post-treatment monitoring all depend on accurate segmentation. To explore AI facilitation of these efforts, the Brain Tumor

Segmentation (BraTS) challenges have been ongoing for 10 years, representing the crucial joint efforts of the Radiological Society of North America (RSNA), the American Society of Neuroradiology (ASNR), and the Medical Image Computing and Computer Assisted Interventions (MICCAI) society to characterize gliomas and to improve algorithms designed for brain glioma segmentation, tumor compartmentalization, and tumor molecular characterization. New dataset analytics, including the most recent iterations of the BraTS challenge, have expanded case pathology and segmentation analytics beyond just gliomas to include meningiomas, pediatric brain, tumors, brain metastases, schwannomas, and other brain tumors [75].

While there are several recognized brain tumor datasets, the BraTS datasets are some of the most widely used for training and testing AI-based algorithms that focus on tumor segmentation. These multi-institution datasets have been growing annually since 2012 and comprise hundreds of well-annotated images obtained using common MR sequences (T1W, T2W, T1c, and FLAIR) of both high- and low-grade tumors. The datasets also consist of pooled ground truths determined by a set of expert raters who segment various components including edema, necrotic tumor core, and enhancing tumor. Numerous studies and papers related to tumor segmentation have been published in relation to the BraTS dataset over the years using various AI methods, including DL and other ML algorithms, with performance improving over the years [76,77]. Initial algorithms using the 2012 dataset primarily used conventional ML methods, with Dice scores ranging from 0.14 to 0.70. Recently, DL models have dominated the landscape, with significant performance increases. One recent meta-analysis of 10 ML-based segmentation studies demonstrated Dice scores of 0.84 [78]. Another meta-analysis of DL algorithms demonstrated slightly improved performance, with a pooled Dice score of 0.89 [76]. A recent study using the BraTS datasets and a 3D U-Net DL method achieved a Dice score of 0.95 [79]. There are countless additional algorithms that have achieved results above 90%, illustrating the significant progress made in such a short period of time.

6.3. Post Treatment Evaluation

Radiomics studies are using ML to assess brain tumor responses to clinical treatment, differentiate progression and pseudoprogression, and predict the recurrence and infiltration of neoplastic disease. It can often be hard to delineate tumor response and post-chemotherapy and radiation changes upon subsequent imaging. Further complicating the picture are new immunotherapies that incite complex inflammatory responses and antiangiogenic agents like bevacizumab that can cause a reduction in enhancement by reducing microvascular growth, but without necessarily offering any improvement in the overall survival rate [4].

Pseudoprogression is characterized by increased extent of imaging abnormalities after treatment, including abnormal enhancement and T2/FLAIR signal changes, that typically resolve without intervention. Differentiating pseudoprogression from true tumor progression is vital to treatment success and clinical outcomes. AI methods and techniques, including radiomics, are currently being explored to differentiate true progression from pseudoprogression by incorporating features from advanced MRI techniques, including diffusion-weighted imaging (DWI), perfusion-weighted imaging (PWI), and MR spectroscopy [80]. For example, Sun et al. performed a retrospective study on 77 post-treatment GBM patients, 26 of whom had known pseudoprogression [81]. Clinical data, patient outcomes, and multiple features from T1-weighted imaging were used to develop a radiomics model in order to evaluate true progression from pseudoprogression in comparison to the results of 3 trained radiologists. The model demonstrated sensitivity and specificity of 78% and 61%, respectively. The 3 radiologists had sensitivities of 62–69% and specificities of 47–68%.

Differentiating infiltrating neoplasms from adjacent edema is difficult using conventional imaging approaches. However, ML has the capability to identify the margins of infiltrative tissue from normal tissue on pre-treatment MRI images. Identifying infiltrative

margins is important in resection planning, biopsy site selection, and monitoring treatment response. Approaches that incorporate ML have been successful in generating spatial maps of infiltrated tissue, with approximately 90% cross-validated accuracy [82]. A fully automated CNN system has been created for the purpose of registering biopsy sites using MR images. The system generates noninvasive maps of cell density to identify the infiltrative margins of gliomas rather than just relying on the margins of the enhancing tumor alone [83]. Determining the level of surrounding infiltration can help to stratify patients into the appropriate invasive or noninvasive treatment options.

7. White Matter Disease

7.1. White Matter Hyperintensities

White matter lesions (WMLs), or white matter hyperintensities (WMHs), are areas of abnormal myelination that can be best appreciated as increased signal intensity on T2/FLAIR sequences. They are widely utilized neuroradiologic biomarkers of brain parenchymal pathologies. As a group, they include but are not limited to small-vessel ischemic disease, demyelinating disease, or other inflammatory processes [84,85]. The presence of these lesions has been shown to incur an increased risk of stroke, dementia, and death [86,87]. Furthermore, WMLs are associated with grey matter atrophy, accelerated neurodegeneration, and cerebrovascular incidents [88].

Due to the impact and location-dependent nature of these lesions, accurate detection and quantification is paramount. Traditional practices involve the manual delineation of regions via analysis dependent on visual, qualitative, or semiquantitative inspection of imaging, although even early studies recognized the benefit of automatic or semiautomatic image analysis. The utilization of neural networks and other AI approaches in quantifying white matter signal abnormalities and related findings such as brain volume changes has evolved over the past five years, with numerous methods already developed and commercially available [89–96]. Currently available software packages can include classification algorithms (e.g., diagnosis), regression algorithms (e.g., linking clinical scores or liquid biomarker levels to images), detection algorithms, segmentation algorithms, or a mix [97]. Several of the commercially available software packages are listed on the Grand Challenge website (<https://grand-challenge.org/aiforradiology/>, accessed on 15 May 2023).

Some of the most promising clinical applications of these algorithms have included use in the automatic and semiautomatic segmentation of multiple-sclerosis (MS) lesions, age-related WMH, and in aspects of radiomics [4,98]. All of these rely on the accurate quantification of lesions to determine prognosis, monitor treatment, and develop quantitative imaging biomarkers. Recent studies have found that automatic segmentation of these lesions has excellent accuracy and fares well compared to manual segmentation, with processing times ranging from 2–20 min. However, methods tend to underestimate the volume of lesions, and no current method factors in the wide variability of MRI contrast within subjects and protocols [98–100]. Furthermore, given the numerous available algorithms and packages, comparing options can be challenging. For this reason, comparison generally relies in part on the ECLAIR guidelines (Evaluation of Commercial AI Solutions in Radiology). However, issues remain on account of the variability of training datasets, patient populations, and many other factors. Perhaps most notably, segmentation of WMLs, despite being the most common primary task performed by AI, has not been well standardized for large-scale studies and has therefore historically struggled with automation [101,102]. Accurate quantification is a key area of research and exploration given that segmentation is a key factor in determining the extent of disease, the changes in clinical picture, and patient outcomes.

Similar to early public databases such as BrainWeb computational phantom, competitions like the International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI) and WMH Segmentation Challenge serve as catalysts for the development of novel methods and provide multi-institutional comparisons of algorithms using standardized evaluation criteria [103,104]. Several similar challenges

exist that focus on different tasks, including multiple sclerosis (MS) lesions, tumors, and strokes [105–108]. Although the results of these varying competitions are largely incomparable due to differences in evaluation criteria and datasets, they provide valuable timelines for advancements in AI and ML in terms of radiological detection, segmentation, and quantification. Furthermore, some challenges such as the MS Lesion Segmentation Challenge occurred over multiple years (2015 and 2021) and were therefore able to provide insights into advancements within a specific application.

The 2017 MICCAI WMH Segmentation Challenge was held in Quebec, Canada, and originally included 20 participating teams. Teams were given a training dataset, containing 60 brain MR cases from three separate scanners, to develop their methods with and were then tested using a separate test dataset containing 110 MRI brain cases from five different scanners [103]. Though new methods are still being submitted, the winning method of the challenge demonstrated an F1 score of 0.76 and a Dice score of 0.80. This accuracy was achieved using an algorithm based on an ensemble of convolution–deconvolution architectures with 19-layer implementation optimized for classifying and localizing WMHs [25,102,109]. However, the algorithm did not perform well with multi-scale features, and in 2020 Liu and colleagues created a deep convolutional neural network, M2DCNN. This addressed the problem using two subnets that rely on a set of novel multi-scale features and a novel architecture designed to reduce the loss of receptive fields [110]. In addition to improved detection of large and small lesions, M2DCNN demonstrated fewer false positives and less variability in the predictive performance than previously described segmentation methods. Clinically, this advancement allows radiologists the ability to apply AI algorithms to images with minimal spatial information loss, improved image classification, and better localized segmentation [111]. Most recently, a new method built on the U-Net architecture achieved F1 scores as high as 0.93 using the MICCAI WMH challenge training dataset by introducing dense connections to allow for better utilization of multi-scale features [110].

7.2. Multiple Sclerosis (MS)

Multiple sclerosis is a debilitating disease that often affects people in the prime of their lives. The prevalence of MS is growing, and the need to have automated accurate detection and quantification of lesion burden is crucial to disease management and prognosis [112].

The first major MS lesion segmentation challenge occurred at the MICCAI 2008 conference. A second MS lesion segmentation challenge was conducted at the 2015 International Symposium on Biomedical Imaging (ISBI) in New York. The 2015 challenge and dataset are still being used and currently have over 2000 submissions. A subsequent 2016 segmentation challenge and dataset attempted to address multiple potential issues faced during the 2008 and 2015 challenges by using high-quality patient cases, drawn from four different scanners after delineation by seven expert evaluators, to reduce inter-rater variability. Evaluations utilized a distributed Web platform for the automatic, fair comparisons of algorithms. The challenge evaluated candidate algorithms' performance for both detection (correct identification of all lesions in an image) and segmentation (precisely outline lesions). Only 13 teams participated in the original challenge in 2016. In this event, F1 lesion detection scores ranged from 0.13 to 0.49 (avg 0.32), with an expert range of 0.66 to 0.89 (avg 0.77). Dice segmentation scores for the algorithms ranged from 0.27 to 0.59 (avg 0.46), with the expert Dice scores ranging from 0.69 to 0.78 (avg 0.71) [108,113]. It was also noted that algorithm performance diminished with an increased number of lesions and a decreased size of lesions. The challenge was modified in 2021 to delineate new MS lesions with similar parameters and larger datasets. Although all experts still scored significantly higher than any submitted method, the difference was much smaller for detection, with combined average F1 scores of 0.61 for the experts compared to 0.42 for automatic methods [114]. Combined average Dice scores for the experts were 0.56 compared to 0.39 for the automatic methods. Detailed descriptions of each method and its

results have been published in HAL Open Science, with comprehensive results available on the website (<https://zenodo.org/record/5775523>, accessed on 15 May 2023) [108,114].

8. Neurocognitive

Dementia is an acquired syndrome characterized by a significant decline in cognitive function that leads to difficulty in daily functioning/independence, with an estimated prevalence in the United States of 11% in individuals over 65 years old [115]. Mild cognitive impairment (MCI) represents cognitive impairment that is more severe than normal aging but which does not interfere with independent daily functioning. The screening, diagnosis, and monitoring of neurocognitive disorders are typically guided by a patient's history and clinical symptoms, including an emphasis on the clinical interview [116]. Neuroimaging, specifically structural MRI, represents a widely available and non-invasive test when evaluating for cognitive dysfunction that is commonly used to support a diagnosis of cognitive impairment [117]. However, dementia and MCI are significantly underdiagnosed in the community setting, with up to 60% of cases not being detected [118]. The National Institute on Aging-Alzheimer's Association Framework predicts that there also exists a silent preclinical stage of Alzheimer's disease before cognitive symptoms emerge, with the possibility of novel interventions at this preclinical stage [119].

Because early and accurate diagnosis of dementia and other neurocognitive disorders is imperative to allow access to supportive therapies that can help patients maintain their independence, ML methods based on neuroimaging have great potential to promote earlier and more sensitive diagnosis of neurocognitive disorders [120]. For instance, ML algorithms utilizing electronic health record (EHR) data for early detection of cognitive impairment risk are being evaluated in an active clinical trial [121]. Because of the potential benefits of neuroimaging-based ML applications in the detection of dementia, there have been numerous studies investigating the applications of ML in this field.

Numerous challenges have been conducted to investigate the use of structural-MRI-based ML in the screening and diagnosis of dementia. While several challenges addressed research questions of aiding prediction of future outcomes and cognitive scores, the challenges that most pertain to the practicing neuroradiologist largely correspond to three clinical questions in dementia: screening, clinical status classification, and monitoring of disease progression [122–124]. The remainder of this section will briefly review current clinical practices in cognitive impairment screening, classification, and monitoring and discuss the Predictive Analytics Competition (PAC) 2019, Computer-aided diagnosis of Dementia (CADDementia) challenge, and Minimal Interval Resonance Imaging in Alzheimer's Disease (MIRIAD) challenge, as well as their implications in their respective clinical tasks.

8.1. Screening

Currently, approaches to screening for cognitive impairment include screening tests such as the mini-mental state examination (MMSE), clock drawing test, and Montreal cognitive assessment (MoCA), among others, as well as biological markers [121]. However, in the most recent 2020 US Preventative Services Task Force (USPSTF) recommendation statement, they concluded that there existed insufficient evidence to assess the benefits and harms of widespread screening of asymptomatic adults for cognitive impairment [125]. Most primary care systems are not equipped to routinely detect dementia, especially in multicultural populations or those with lower levels of educational attainment [118]. These gaps in screening for cognitive impairment are an opportunity for neuroimaging-based AI algorithms to aid in clinical decision making.

The Predictive Analytics Competition (PAC) 2019 sought to improve ML models using whole-brain structural MRI scans from healthy individuals to predict brain age, a cumulative screening marker of functional capacity, residual lifespan, and the risk of progression to neurocognitive disease [126]. Participants were given the following tasks: 1. minimize the mean absolute error (MAE) between chronological and predicted age (brain-age gap) and 2. minimize brain-age gap while keeping the Spearman correlation

between the brain–age gap and chronological age below $r = 0.10$ in order to reduce bias. Seventy-nine participating teams were provided with a large structural MRI training dataset of healthy individuals with ages provided ($N = 2640$) and tested on a test dataset obtained from the same institutions ($N = 660$). The winning team of Gong et al. used lightweight 3D convolutional neural networks (CNNs), combined with preprocessing steps and pre-trained on UK Biobank data (MAE = 2.95 years after bias correction) [127].

PAC 2019 revealed that current ML models are increasingly effective at predicting brain age without incurring significant bias. DL models also outperformed classic ML algorithms [126]. CNNs using 3D kernels instead of 2D kernels were useful in efforts to exploit features across all spatial dimensions. The top-performing models also utilized shallower CNN architectures in contrast to the deep 2D CNN architectures used in slice-level modeling [127], suggesting that either the sample size used in this challenge was too small for deeper structures to confer advantages, or that age-related brain morphology changes are relatively simple to detect [127].

While this challenge demonstrated the efficacy of ML models in the specific task of brain age prediction, the cost and availability of structural MRI acquisition for the general population of healthy adults as a screening test for cognitive impairment remains a barrier to the application of these models in clinical practice. Brain age prediction through these ML models may be more useful for populations that are predisposed to cognitive impairment, such as those with a family history or other known biomarkers.

8.2. Classification

In current clinical practice, the initial evaluation of cognitive impairment includes elements of clinical history, neurologic examinations with an emphasis on mental status, labs to screen for reversible causes of cognitive dysfunction (i.e., chemistries, thyroid panel, B12), and structural brain imaging, with MRI being the preferred method over CT [116]. Classification of cognitive impairment, which refers in this article to the distinction between cognitively normal patients (CN), those with MCI, and those with dementia, is primarily performed using elements of patient history and an assessment of cognitive function.

The CADDementia challenge was conducted in 2014 and gave participants the task of classifying baseline MRI scans into three diagnostic classes: Alzheimer’s disease, mild cognitive impairment (MCI) and cognitively normal (CN) [128]. Fifteen research teams submitted a total of 29 algorithms. The challenge provided a small training set of multi-center T1-weighted MRI scans ($N = 30$) of patients that were equally representative of the 3 diagnostic classes. Participants were also allowed to use extra training data, and most used data from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database and/or the Australian Imaging Biomarker and Lifestyle flagship study of aging (AIBL). Algorithms were tested against a previously unseen multi-center CADDementia dataset ($N = 354$) with clinical diagnosis established via a multi-disciplinary consensus that was blinded to the participants. The performance of the algorithms was quantified by classification accuracy, area under the receiver operating characteristic (ROC) curve (AUC), and the true positive fraction for the three classes.

The challenge submissions utilized a wide range of approaches, with most methods accounting for input features such as volume (most common feature, used by $N = 19$ algorithms), cortical thickness, intensity, and shape. Algorithms also utilized a variety of classifiers, such as SVM classifier, random forest classifier, and linear discriminant analysis (LDA). Only one algorithm (Folego-ADNet) used a convolutional neural network for classification and achieved a relatively low rank [129]. The best-performing algorithm used a LDA of features measuring volume, thickness, shape, and intensity relations of brain regions (accuracy = 63%, AUC = 78%) [130].

The results of the CADDementia challenge showed that the best-performing algorithms incorporated multiple features and utilized additional larger training datasets. While the performance of algorithms in this challenge was deemed too low for clinical

application [122], multiple groups have applied DL models to ADNI data for this 3-class classification and achieved accuracies as high as 90% [131].

8.3. Assessing Disease Progression

In current clinical practice, monitoring for disease progression in dementia is largely centered on factors such as the loss of additional cognitive function, which may be quantified using the same screening instruments discussed earlier. Serial neuroimaging is not routinely performed in patients with dementia unless there is a new rapid loss of cognition, focal neurological signs, or seizure. There exist ongoing datasets of longitudinal biomarkers of dementia, such as the ADNI, although many of these biomarkers are obtained primarily in research settings and not part of standard clinical practice [132,133].

The Minimal Interval Resonance Imaging in Alzheimer's Disease (MIRIAD) challenge was conducted to develop and compare methods of estimating atrophy and rates of atrophy from structural MRI [134]. In particular, the challenge looked at volumetric measurements of key structures: the whole brain, lateral ventricles, and hippocampus.

The MIRIAD dataset consisted of 708 T1-weighted volumetric scans acquired from 69 subjects (46 patients with clinical diagnosis of AD and 23 cognitively normal controls), with each subject undergoing 1–12 scans over 1–2 years at various time intervals. Submissions were graded based on the predicted sample size requirements for a hypothetical clinical trial, assuming a putative treatment effect of a 25% atrophy rate reduction [134]. The rationale for this approach was that the methodology with the most utility would require the smallest sample size to provide sufficient power, given that all other aspects of the design are fixed.

Challenge participants produced consistent and repeatable measures of change in brain regions and ventricles; however, hippocampal measures were the most variable among the submissions. Cash et al. attribute this variance to the differing definitions of the hippocampus used in segmentation protocols, as well as the possibility that the hippocampus is a small structure of the brain susceptible to MRI acquisition artifacts [134]. The best methods, i.e., those requiring the smallest sample sizes, were the boundary shift integral for whole-brain atrophy and the combination of diffeomorphic registration (Demons-LCC) and regional flux analysis for ventricle and hippocampus atrophy.

The application of the results of the MIRIAD challenge to clinical practice would be limited to the occasional cases of cognitive impairment in which serial neuroimaging is acquired. Several studies have attempted to predict the conversion of MCI into AD using baseline data, with accuracies as high as 83% when incorporating the results of multiple modalities, including structural MRI, PET, CSF, and clinical metrics [135]. This predictive task has the potential to permit the earlier identification of patients at highest risk for future progression to full-fledged dementia.

9. Spine

Back pain is a common problem and one of the leading causes of disability in both developed and developing countries [136,137]. Degenerative processes and spinal injuries also have increased in prevalence across all age groups, and the demand for imaging has followed suit [138,139]. MRI and CT are heavily relied upon to accurately diagnose spinal degenerative changes, deformities, instability, and fractures. This section will discuss the fundamentals of spine AI and current research on clinical applications.

Accurate spine labeling is a critically important aspect of imaging interpretation as it establishes the numerical and categorical relationship of one vertebra to another. Manual labeling can be time-consuming and error-prone, particularly when variant anatomy exists. In the setting of AI, several definitions pertain to the process of assigning a specific identifier to a vertebral body. Localization is the detection of a distinct three-dimensional unit in space (such as the vertebral body or the intervertebral disc), and labeling assigns an identifier or class to the three-dimensional unit. Segmentation can be thought of as a voxel-level labeling task [140]. Automated localization and segmentation will greatly increase

accuracy and efficiency while also serving as a fundamental tool for additional AI-related spinal applications [140,141]. This includes diagnosis and evaluation of vertebral fractures, scoliosis, kyphosis, and degenerative intervertebral discs, which will be discussed later in this section.

Segmentation is a demanding task, both for radiologists and algorithms. While manual segmentation is possible, it is not practical in the clinical setting. For example, it is very time-consuming to segment the posterior elements manually. AI-based segmentation algorithms may suffer from an insufficient volume of adequate training datasets. In conjunction with the International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), the Large Scale Vertebrae Segmentation Challenge (VerSe) was organized in 2019 and 2020 to tackle this ongoing problem. A large training dataset was created that included 374 spine CT exams annotated through semi-automated techniques, with additional manual refinement [140]. A total of 26 different algorithms were evaluated and compared based on the metrics of labeling accuracy and segmentation. In 2019, the best algorithms produced an identification (labeling) rate of 94.3% and a Dice score of 89.9%, which improved to 96.6% and 91.7% in 2020, respectively. Several specific impediments to accurate AI labeling and segmentation were fractures, metal implants, cement, and transitional vertebrae [140].

Intervertebral discs (IVD) have also been targets for automated localization and segmentation in efforts to help identify and quantify degenerative disc disease with the goals of minimizing error and time spent on manual interpretation [141,142]. Li et al. created an automated method that received first place in the MICCAI 2016 automatic IVD challenge, with a mean segmentation Dice coefficient of 91.2% and a mean localization error of 0.62 mm on multi-modal magnetic resonance images (MRI) [141]. The low performance of the segmentation algorithms primarily involved disc margins, which commonly can have variable and irregular appearances, which the authors suggested could be improved with higher image resolution [143]. As with variation in vertebral anatomy, variations in disc morphology may also affect the segmentation of discs.

There are numerous applications once accurate segmentation is achieved, including the evaluation of degenerative changes. In 2017, Jamaluden et al. developed a localization algorithm using T2-weighted sagittal spine MRI images with an accuracy of 95.6% [144]. Their algorithm was also able to grade canal stenosis, Modic type changes, and disc narrowing, as well as perform Pfirrmann grading at a level comparable to a radiologist [144,145]. Hallinan et al. developed a DL model using a set of 446 MRI lumbar spine studies (T2W axial and T1W sagittal) that classified central spinal canal, lateral recess, and foraminal stenosis using a dichotomous scale (normal/mild vs. moderate/severe [146]), with performance comparable to 2 radiologists [146]. The DL algorithm was slightly worse in terms of agreement when tasked with assigning four distinct levels of severity (normal, mild, moderate, severe).

Vertebral fractures in the thoracic and lumbar spine have also been targets of AI automation and assistance. Compression fractures may be overlooked by radiologists, especially those who do not frequently read spine imaging [147]. Burns et al. reported an algorithm with the ability to detect and localize thoracic and lumbar spinal fractures on CT scans from 150 patients with a sensitivity of 95.7% and a false positive rate of 0.29 per patient [147,148]. The authors were also able to classify by Gerant type (anterior, middle, and posterior height loss) with an accuracy of 95%. Murata et al. trained a DL model with AP and lateral thoracolumbar radiographs on 300 patients that could localize fractures with the accuracy and sensitivity of 86% and 84.7%, respectively [149]. They demonstrated their model's ability to detect vertebral fractures to be equivalent to that of orthopedic surgeons and residents [149,150]. In conjunction with the American Society of Neuroradiology (ASNR) and American Society of Spine Radiology (ASSR), the Radiological Society of North America (RSNA) created a challenge in 2022 to encourage the creation of AI-based algorithms to detect and localize cervical spinal fractures. Their dataset included

around 3000 normal and fracture-positive CT examinations annotated by expert spinal radiologists from ASNR and ASSR.

The evaluation of scoliosis is another prime target for AI, including Cobb angle measurements. Issues of reproducibility, accuracy, and the inter-rater reliability of clinicians manually measuring these angles could potentially be addressed using AI. Wang et al. proposed a multi-view extrapolation net (MVE-Net) to estimate Cobb angles using AP and lateral radiographs [151]. The authors obtained a circular mean absolute error of 7.81 degrees on AP and 6.26 degrees on lateral x-ray angle estimation on 526 images, presenting a reasonably accurate estimation of the degree of scoliosis. Large-scale competitions, such as the Accurate Automated Spinal Curvature Estimate (AASCE2019) challenge, have demonstrated strong performances among many algorithms [152]. The AASCE2019 challenge employed a dataset of 707 spine AP radiographs and evaluated algorithms' performance at producing accurate and reliable Cobb angles. Zhang et al. also used an artificial neural network to measure Cobb angles, examining 65 in vivo coronal radiographs (patients with idiopathic scoliosis) and 40 model radiographs (from a spine model positioned in different poses) [153]. Their model had an absolute error of less than 3 degrees for the spine model radiographs but performed worse on the in vivo images.

10. Head and Neck

10.1. Tumors

Artificial intelligence (AI) has the potential to revolutionize head and neck imaging by augmenting image quality and improving its ability to perform clinically relevant tasks such as tumor volume segmentation, tumor characterization, tumor prognostication, treatment response assessment, and the prediction of metastatic lymph node disease [154,155]. Head and neck oncology care is well positioned for the application of imaging AI since treatment is guided by a wealth of information derived from US, CT, and MRI imaging data. DeJohn et al. conducted a literature review of the current state of the field and identified several areas where ML could potentially be applied, including image quality enhancement, automatic feature extraction, and automated diagnosis [156]. ML and DL models can improve patient care throughout the clinical workflow from the time of imaging to interpretation and through quality improvement via standardization of automated tools.

AI has also been applied in the prognostication of responses to chemotherapy or radiation in head and neck cancer [157]. In addition to auto-segmentation for treatment planning, AI tools can also be beneficial in oncological outcome prediction and toxicity prediction in radiation treatment [158]. The Head and Neck Organ-at-Risk Multi-Modal Segmentation Challenge (<https://han-seg2023.grand-challenge.org/>, accessed on 15 May 2023) was launched recently to promote the development of new and existing applications of fully automated techniques for OAR (organ-at-risk) segmentation in the head and neck regions of CT images. The goal of this challenge is to exploit the information of multiple imaging modalities in order to improve the accuracy of segmentation results. The HEad and neCK TumOR (HECKTOR) challenges (<https://hecktor.grand-challenge.org/>, accessed on 15 May 2023) focused on establishing best-performing methods in order to predict patient outcomes from FDG-PET/CT and clinical data and conduct the automatic segmentation of head and neck primary tumors and lymph nodes on FDG-PET/CT images.

AI has the potential to significantly improve the accuracy and efficiency of ultrasound use in head and neck oncology. A systematic review by Santer et al. found that 74% of studies on the use of AI in ultrasound for head and neck oncology addressed disease diagnosis, with 56% examining the ability to distinguish benign and malignant thyroid nodules and 44% seeking to identify metastatic lymph nodes [157]. Radiomics-based MRI features have been used in the assessment of various head and neck cancer (HNC) lesions. In several studies, traditional ML techniques have been used for the automatic segmentation of HNC lesions using MRI, with promising results in terms of accuracy ($86 \pm 8\%$) and overlap measures ($0.76+/-0.08$) [159,160]. The textural analysis of MRI and CT images has also been used to differentiate between different types of HNC lesions, with

accuracies ranging from 75.7% to 100% [161–163]. Parameters obtained via histogram and texture analysis of MRI T2WI can even serve as noninvasive predictors of histological type and grade in head and neck malignancy [164]. In addition, textural features derived from intraoral X-ray images have been used to predict the early onset of oral squamous cell carcinoma, with accuracies of 99.2% [165]. Overall, these findings suggest that radiomics-based prediction can be a useful tool for the assessment and diagnosis of HNC lesions.

The increasing number of independent prognostic and predictive markers has sparked interest in the use of AI-based prediction models. AI-based methods can integrate complex imaging, histologic, molecular, and clinical data to model tumor biology and behavior, and potentially identify associations far beyond what conventional qualitative imaging can provide alone. DL-based models can more accurately predict oncological outcomes using pre-treatment data than existing models. Likewise, they are better at predicting treatment toxicity prior to the start of treatment as well as the prediction of pathological data from imaging data [166].

10.2. Vascular Lesions

CTA is a widely used and cost-effective imaging modality for the diagnosis of cerebrovascular disease in the head and neck region. However, manual postprocessing of CTA images can be time-consuming and subject to human error. DL-based segmentation approaches have been proposed as having the potential to improve the efficiency of CTA analysis by reducing the need for manual post-processing. One major challenge in CTA image postprocessing is the accurate segmentation of vessels, exacerbated by their branching morphology, variable anatomy, and overlap in density with other tissues. To address these challenges, Fu et al. developed an automatic imaging reconstruction system called CerebralDoc, which uses a 3D-CNN containing modified U-net components for the reconstruction of original head and neck CTA images [167]. This system has the potential to assist CT technologist or radiologist workflow and improve efficiency by removing time-consuming steps in CTA post-processing.

Traditionally, the noninvasive diagnosis of cerebral aneurysms has relied on imaging modalities such as CTA or magnetic resonance angiography (MRA). However, these techniques can be limited in their ability to accurately detect and classify small or complex aneurysms. AI has the potential to improve the accuracy of cerebral aneurysm diagnosis by automating the analysis of imaging data and identifying subtle features that may be overlooked by human observers. In a study by Park et al., a DL-based model called HeadXNet was developed for the diagnosis of cerebral aneurysms using CTA images [168]. The model alone had a sensitivity of 0.95 and specificity of 0.66 for the detection of aneurysms. When paired with trained radiologists, the model improved a radiologist's sensitivity, accuracy, and interrater agreement. Chen et al. developed a DL-based method for the segmentation of cerebral aneurysms in 3D TOF-MRA images using a coarse-to-fine framework [155]. The method was able to accurately identify and segment aneurysms, with a Dice coefficient of 0.87.

DL-based models have also been developed in order to improve the accuracy of cerebral aneurysm rupture risk prediction. In a study by Yang et al., a CNN-based DL model was developed for the prediction of cerebral aneurysm rupture risk using 3D time-of-flight magnetic resonance angiography (TOF-MRA) images [169]. The model was able to achieve high accuracy in predicting rupture risk, with an AUC of 0.95 [169]. Similarly, AI has the potential to improve the accuracy and efficiency of venous malformation diagnosis and treatment planning. In a study by Ryu et al. (2022), a DL-based method called 3D U-Net was used for the automatic segmentation of extracranial venous malformations in the head and neck region from MRI images [170]. The method was able to accurately identify and segment venous malformations with a high degree of accuracy, producing a Dice coefficient of 0.87.

11. Conclusions

Artificial intelligence has numerous applications throughout the field of neuroradiology, with great promise to augment the work of the modern-day radiologist. AI-based methods have taken large strides in accuracy and efficiency over the past decade, some of which can be linked to the AI challenge competitions conducted to benchmark leading AI methodologies using well-annotated datasets. As demonstrated in this article, AI applications have been developed and evaluated for use in detecting or quantifying intracranial hemorrhage and stroke, brain and head/neck tumors, spinal fractures, degenerative spinal disease, and inflammatory or neurodegenerative brain disorders. While many AI methods demonstrate remarkable performance in specific tasks and several software packages have been approved for clinical use, there remains a continuous need to push the field further to develop improved tools with which to better augment the work of the practicing neuroradiologist. AI-based challenges will likely continue to showcase the latest advancements in AI methods and provide an impetus for improvement, ultimately leading to higher-quality patient care.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Pesapane, F.; Codari, M.; Sardanelli, F. Artificial intelligence in medical imaging: Threat or opportunity? Radiologists again at the forefront of innovation in medicine. *Eur. Radiol. Exp.* **2018**, *2*, 35–36. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Wiggins, W.F.; Magudia, K.; Schmidt, T.M.S.; O'Connor, S.D.; Carr, C.D.; Kohli, M.D.; Andriole, K.P. Imaging AI in practice: A demonstration of future workflow using integration standards. *Radiol. Artif. Intell.* **2021**, *3*, e210152. [\[CrossRef\]](#)
3. Langlotz, C.P. Will artificial intelligence replace radiologists? *Radiol. Artif. Intell.* **2019**, *1*, e190058. [\[CrossRef\]](#)
4. Rudie, J.D.; Rauschecker, A.M.; Bryan, R.N.; Davatzikos, C.; Mohan, S. Emerging applications of artificial intelligence in neuro-oncology. *Radiology* **2019**, *290*, 607–618. [\[CrossRef\]](#)
5. Yu, A.C.; Mohajer, B.; Eng, J. External validation of deep learning algorithms for radiologic diagnosis: A systematic review. *Radiol. Artif. Intell.* **2022**, *4*, e210064. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Erickson, B.J.; Korfiatis, P.; Akkus, Z.; Kline, T.L. Machine learning for medical imaging. *Radiographics* **2017**, *37*, 505–515. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Tsao, C.W.; Aday, A.W.; Almarzooq, Z.I.; Alonso, A.; Beaton, A.Z.; Bittencourt, M.S.; Boehme, A.K.; Buxton, A.E.; Carson, A.P.; Commodore-Mensah, Y.; et al. Heart disease and stroke statistics-2022 update: A report from the American Heart Association. *Circulation* **2022**, *145*, e153–e639. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Powers, W.J.; Rabinstein, A.A.; Ackerson, T.; Adeoye, O.M.; Bambakidis, N.C.; Becker, K.; Biller, J.; Brown, M.; Demaerschalk, B.M.; Hoh, B.; et al. Guidelines for the early management of patients with acute ischemic stroke: 2019 update to the 2018 guidelines for the early management of acute ischemic stroke: A guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke* **2019**, *50*, e344–e418. [\[CrossRef\]](#)
9. Albers, G.W.; Marks, M.P.; Kemp, S.; Christensen, S.; Tsai, J.P.; Ortega-Gutierrez, S.; McTaggart, R.A.; Torbey, M.T.; Kim-Tenser, M.; Leslie-Mazwi, T.; et al. Thrombectomy for stroke at 6 to 16 hours with selection by perfusion imaging. *N. Engl. J. Med.* **2018**, *378*, 708–718. [\[CrossRef\]](#)
10. Nogueira, R.G.; Jadhav, A.P.; Haussen, D.C.; Bonafe, A.; Budzik, R.F.; Bhuva, P.; Yavagal, D.R.; Ribo, M.; Cognard, C.; Hanel, R.A.; et al. Thrombectomy 6 to 24 hours after stroke with a mismatch between deficit and infarct. *N. Engl. J. Med.* **2018**, *378*, 11–21. [\[CrossRef\]](#)
11. Wardlaw, J.M.; Mair, G.; von Kummer, R.; Williams, M.C.; Li, W.; Storkey, A.J.; Trucco, E.; Liebeskind, D.S.; Farrall, A.; Bath, P.M.; et al. Accuracy of automated computer-aided diagnosis for stroke imaging: A critical evaluation of current evidence. *Stroke* **2022**, *53*, 2393–2403. [\[CrossRef\]](#)
12. Soun, J.E.; Chow, D.S.; Nagamine, M.; Takhtawala, R.S.; Filippi, C.G.; Yu, W.; Chang, P.D. Artificial intelligence and acute stroke imaging. *AJNR Am. J. Neuroradiol.* **2021**, *42*, 2–11. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Mokli, Y.; Pfaff, J.; dos Santos, D.P.; Herweh, C.; Nagel, S. Computer-aided imaging analysis in acute ischemic stroke—Background and clinical applications. *Neurol. Res. Pract.* **2019**, *1*, 23. [\[CrossRef\]](#) [\[PubMed\]](#)

14. Murray, N.M.; Unberath, M.; Hager, G.D.; Hui, F.K. Artificial intelligence to diagnose ischemic stroke and identify large vessel occlusions: A systematic review. *J. Neurointerv. Surg.* **2020**, *12*, 156–164. [\[CrossRef\]](#)
15. Shafaat, O.; Bernstock, J.D.; Shafaat, A.; Yedavalli, V.S.; Elsayed, G.; Gupta, S.; Sotoudeh, E.; Sair, H.I.; Yousem, D.M.; Sotoudeh, H. Leveraging artificial intelligence in ischemic stroke imaging. *J. Neuroradiol.* **2022**, *49*, 343–351. [\[CrossRef\]](#)
16. Dawud, A.M.; Yurtkan, K.; Oztoprak, H. Application of deep learning in neuroradiology: Brain haemorrhage classification using transfer learning. *Comput. Intell. Neurosci.* **2019**, *2019*, 4629859. [\[CrossRef\]](#)
17. Chavva, I.R.; Crawford, A.L.; Mazurek, M.H.; Yuen, M.M.; Prabhat, A.M.; Payabvash, S.; Sze, G.; Falcone, G.J.; Matouk, C.C.; de Havenon, A.; et al. Deep learning applications for acute stroke management. *Ann. Neurol.* **2022**, *92*, 574–587. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Demeestere, J.; Wouters, A.; Christensen, S.; Lemmens, R.; Lansberg, M.G. Review of perfusion imaging in acute ischemic stroke: From time to tissue. *Stroke* **2020**, *51*, 1017–1024. [\[CrossRef\]](#)
19. Mokin, M.; Levy, E.I.; Saver, J.L.; Siddiqui, A.H.; Goyal, M.; Bonafé, A.; Cognard, C.; Jahan, R.; Albers, G.W.; SWIFT PRIME Investigators; et al. Predictive value of RAPID assessed perfusion thresholds on final infarct volume in SWIFT PRIME (solitaire with the intention for thrombectomy as primary endovascular treatment). *Stroke* **2017**, *48*, 932–938. [\[CrossRef\]](#)
20. Boned, S.; Padroni, M.; Rubiera, M.; Tomasello, A.; Coscojuela, P.; Romero, N.; Muchada, M.; Rodríguez-Luna, D.; Flores, A.; Rodríguez, N.; et al. Admission CT perfusion may overestimate initial infarct core: The ghost infarct core concept. *J. Neurointerv. Surg.* **2017**, *9*, 66–69. [\[CrossRef\]](#)
21. Hoving, J.W.; Marquering, H.A.; Majoie, C.B.L.M.; Yassi, N.; Sharma, G.; Liebeskind, D.S.; van der Lugt, A.; Roos, Y.B.; van Zwam, W.; van Oostenbrugge, R.J.; et al. Volumetric and spatial accuracy of computed tomography perfusion estimated ischemic core volume in patients with acute ischemic stroke. *Stroke* **2018**, *49*, 2368–2375. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Hakim, A.; Christensen, S.; Winzeck, S.; Lansberg, M.G.; Parsons, M.W.; Lucas, C.; Robben, D.; Wiest, R.; Reyes, M.; Zaharchuk, G. Predicting infarct core from computed tomography perfusion in acute ischemia with machine learning: Lessons from the ISLES challenge. *Stroke* **2021**, *52*, 2328–2337. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Soltanpour, M.; Greiner, R.; Boulanger, P.; Buck, B. Improvement of automatic ischemic stroke lesion segmentation in CT perfusion maps using a learned deep neural network. *Comput. Biol. Med.* **2021**, *137*, 104849. [\[CrossRef\]](#)
24. Clerigues, A.; Valverde, S.; Bernal, J.; Freixenet, J.; Oliver, A.; Llado, X. Acute ischemic stroke lesion core segmentation in CT perfusion images using fully convolutional neural networks. *Comput. Biol. Med.* **2019**, *115*, 103487. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention 2015*; Springer International Publishing: Cham, Switzerland, 2015; pp. 234–241.
26. Wang, X.; Fan, Y.; Zhang, N.; Li, J.; Duan, Y.; Yang, B. Performance of machine learning for tissue outcome prediction in acute ischemic stroke: A systematic review and meta-analysis. *Front. Neurol.* **2022**, *13*, 910259. [\[CrossRef\]](#)
27. Liew, S.; Lo, B.P.; Donnelly, M.R.; Zavaliangos-Petropulu, A.; Jeong, J.N.; Barisano, G.; Hutton, A.; Simon, J.P.; Juliano, J.M.; Suri, A.; et al. A large, curated, open-source stroke neuroimaging dataset to improve lesion segmentation algorithms. *Sci. Data* **2022**, *9*, 320–327. [\[CrossRef\]](#)
28. Amukotuwa, S.A.; Straka, M.; Smith, H.; Chandra, R.V.; Dehkharghani, S.; Fischbein, N.J.; Bammer, R. Automated detection of intracranial large vessel occlusions on computed tomography angiography: A single center experience. *Stroke* **2019**, *50*, 2790–2798. [\[CrossRef\]](#)
29. McLouth, J.; Elstrott, S.; Chaibi, Y.; Quenet, S.; Chang, P.D.; Chow, D.S.; Soun, J.E. Validation of a deep learning tool in the detection of intracranial hemorrhage and large vessel occlusion. *Front. Neurol.* **2021**, *12*, 656112. [\[CrossRef\]](#)
30. Olive-Gadea, M.; Crespo, C.; Granes, C.; Hernandez-Perez, M.; Pérez de la Ossa, N.; Laredo, C.; Urra, X.; Carlos Soler, J.; Soler, A.; Puyalto, P.; et al. Deep learning based software to identify large vessel occlusion on noncontrast computed tomography. *Stroke* **2020**, *51*, 3133–3137. [\[CrossRef\]](#)
31. Rava, R.A.; Peterson, B.A.; Seymour, S.E.; Snyder, K.V.; Mokin, M.; Waqas, M.; Hoi, Y.; Davies, J.M.; Levy, E.I.; Siddiqui, A.H.; et al. Validation of an artificial intelligence-driven large vessel occlusion detection algorithm for acute ischemic stroke patients. *Neuroradiol. J.* **2021**, *34*, 408–417. [\[CrossRef\]](#)
32. Rodrigues, G.; Barreira, C.M.; Bouslama, M.; Haussen, D.C.; Al-Bayati, A.; Pisani, L.; Liberato, B.; Bhatt, N.; Frankel, M.R.; Nogueira, R.G. Automated large artery occlusion detection in stroke: A single-center validation study of an artificial intelligence algorithm. *Cerebrovasc. Dis.* **2022**, *51*, 259–264. [\[CrossRef\]](#)
33. Schlossman, J.; Ro, D.; Salehi, S.; Chow, D.; Yu, W.; Chang, P.D.; Soun, J.E. Head-to-head comparison of commercial artificial intelligence solutions for detection of large vessel occlusion at a comprehensive stroke center. *Front. Neurol.* **2022**, *13*, 1026609. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Tolhuisen, M.L.; Ponomareva, E.; Boers, A.M.M.; Jansen, I.G.H.; Koopman, M.S.; Sales Barros, R.; Berkhemer, O.A.; van Zwam, W.H.; van der Lugt, A.; Majoie, C.B.L.M.; et al. A convolutional neural network for anterior intra-arterial thrombus detection and segmentation on non-contrast computed tomography of patients with acute ischemic stroke. *Appl. Sci.* **2020**, *10*, 4861. [\[CrossRef\]](#)
35. Weyland, C.S.; Papanagiotou, P.; Schmitt, N.; Joly, O.; Bellot, P.; Mokli, Y.; Ringleb, P.A.; Kastrup, A.; Möhlenbruch, M.A.; Bendszus, M.; et al. Hyperdense artery sign in patients with acute ischemic stroke-automated detection with artificial intelligence-driven software. *Front. Neurol.* **2022**, *13*, 807145. [\[CrossRef\]](#)
36. Yahav-Dovrat, A.; Saban, M.; Merhav, G.; Lankri, I.; Abergel, E.; Eran, A.; Tanne, D.; Nogueira, R.G.; Sivan-Hoffmann, R. Evaluation of artificial intelligence-powered identification of large-vessel occlusions in a comprehensive stroke center. *AJNR Am. J. Neuroradiol.* **2021**, *42*, 247–254. [\[CrossRef\]](#) [\[PubMed\]](#)

37. Becks, M.J.; Manniesing, R.; Vister, J.; Pegge, S.A.H.; Steens, S.C.A.; van Dijk, E.J.; Prokop, M.; Meijer, F.J.A. Brain CT perfusion improves intracranial vessel occlusion detection on CT angiography. *J. Neuroradiol.* **2019**, *46*, 124–129. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Stib, M.T.; Vasquez, J.; Dong, M.P.; Kim, Y.H.; Subzwari, S.S.; Triedman, H.J.; Wang, A.; Wang, H.C.; Yao, A.D.; Jayaraman, M.; et al. Detecting large vessel occlusion at multiphase CT angiography by using a deep convolutional neural network. *Radiology* **2020**, *297*, 640–649. [\[CrossRef\]](#)
39. Barber, P.A.; Demchuk, A.M.; Zhang, J.; Buchan, A.M. Validity and reliability of a quantitative computed tomography score in predicting outcome of hyperacute stroke before thrombolytic therapy. ASPECTS study group. Alberta stroke programme early CT score. *Lancet* **2000**, *355*, 1670–1674. [\[CrossRef\]](#)
40. Farzin, B.; Fahed, R.; Guilbert, F.; Poppe, A.Y.; Daneault, N.; Durocher, A.P.; Lanthier, S.; Boudjani, H.; Khoury, N.N.; Roy, D.; et al. Early CT changes in patients admitted for thrombectomy: Intrarater and interrater agreement. *Neurology* **2016**, *87*, 249–256. [\[CrossRef\]](#)
41. Chen, W.; Wu, J.; Wei, R.; Wu, S.; Xia, C.; Wang, D.; Liu, D.; Zheng, L.; Zou, T.; Li, R.; et al. Improving the diagnosis of acute ischemic stroke on non-contrast CT using deep learning: A multicenter study. *Insights Imaging* **2022**, *13*, 184. [\[CrossRef\]](#)
42. Naganuma, M.; Tachibana, A.; Fuchigami, T.; Akahori, S.; Okumura, S.; Yi, K.; Matsuo, Y.; Ikeno, K.; Yonehara, T. Alberta stroke program early CT score calculation using the deep learning-based brain hemisphere comparison algorithm. *J. Stroke Cerebrovasc. Dis.* **2021**, *30*, 105791. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Hoelter, P.; Muehlen, I.; Goelitz, P.; Beuscher, V.; Schwab, S.; Doerfler, A. Automated ASPECT scoring in acute ischemic stroke: Comparison of three software tools. *Neuroradiology* **2020**, *62*, 1231–1238. [\[CrossRef\]](#) [\[PubMed\]](#)
44. Maegerlein, C.; Fischer, J.; Monch, S.; Berndt, M.; Wunderlich, S.; Seifert, C.L.; Lehm, M.; Boeckh-Behrens, T.; Zimmer, C.; Friedrich, B. Automated calculation of the Alberta stroke program early CT score: Feasibility and reliability. *Radiology* **2019**, *291*, 141–148. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Goebel, J.; Stenzel, E.; Guberina, N.; Wanke, I.; Koehrmann, M.; Kleinschnitz, C.; Umutlu, L.; Forsting, M.; Moenninghoff, C.; Radbruch, A. Automated ASPECT rating: Comparison between the frontier ASPECT score software and the Brainomix software. *Neuroradiology* **2018**, *60*, 1267–1272. [\[CrossRef\]](#)
46. Cao, Z.; Xu, J.; Song, B.; Chen, L.; Sun, T.; He, Y.; Wei, Y.; Niu, G.; Zhang, Y.; Feng, Q.; et al. Deep learning derived automated ASPECTS on non-contrast CT scans of acute ischemic stroke patients. *Hum. Brain Mapp.* **2022**, *43*, 3023–3036. [\[CrossRef\]](#)
47. Nagel, S.; Sinha, D.; Day, D.; Reith, W.; Chapot, R.; Papanagiotou, P.; Warburton, E.A.; Guylar, P.; Tysoe, S.; Fassbender, K.; et al. E-ASPECTS software is non-inferior to neuroradiologists in applying the ASPECT score to computed tomography scans of acute ischemic stroke patients. *Int. J. Stroke* **2017**, *12*, 615–622. [\[CrossRef\]](#)
48. Albers, G.W.; Wald, M.J.; Mlynash, M.; Endres, J.; Bammer, R.; Straka, M.; Maier, A.; Hinson, H.E.; Sheth, K.N.; Taylor Kimberly, W.; et al. Automated calculation of Alberta stroke program early CT score: Validation in patients with large hemispheric infarct. *Stroke* **2019**, *50*, 3277–3279. [\[CrossRef\]](#)
49. Guberina, N.; Dietrich, U.; Radbruch, A.; Goebel, J.; Deuschl, C.; Ringelstein, A.; Köhrmann, M.; Kleinschnitz, C.; Forsting, M.; Mönninghoff, C. Detection of early infarction signs with machine learning-based diagnosis by means of the Alberta stroke program early CT score (ASPECTS) in the clinical routine. *Neuroradiology* **2018**, *60*, 889–901. [\[CrossRef\]](#)
50. Hemphill, J.C., 3rd; Greenberg, S.M.; Anderson, C.S.; Becker, K.; Bendok, B.R.; Cushman, M.; Fung, G.L.; Goldstein, J.N.; Macdonald, R.L.; Mitchell, P.H.; et al. Guidelines for the management of spontaneous intracerebral hemorrhage: A guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke* **2015**, *46*, 2032–2060. [\[CrossRef\]](#)
51. Goldstein, J.N.; Gilson, A.J. Critical care management of acute intracerebral hemorrhage. *Curr. Treat. Options Neurol.* **2011**, *13*, 204–216. [\[CrossRef\]](#)
52. Krishnamurthi, R.V.; Feigin, V.L.; Forouzanfar, M.H.; Mensah, G.A.; Connor, M.; Bennett, D.A.; Moran, A.E.; Sacco, R.L.; Anderson, L.M.; Truelsen, T.; et al. Global and regional burden of first-ever ischaemic and haemorrhagic stroke during 1990–2010: Findings from the global burden of disease study 2010. *Lancet Glob. Health* **2013**, *1*, 259. [\[CrossRef\]](#) [\[PubMed\]](#)
53. Heit, J.J.; Coelho, H.; Lima, F.O.; Granja, M.; Aghaebrahim, A.; Hanel, R.; Kwok, K.; Haerian, H.; Cereda, C.W.; Venkatasubramanian, C.; et al. Automated cerebral hemorrhage detection using RAPID. *AJNR Am. J. Neuroradiol.* **2021**, *42*, 273–278. [\[CrossRef\]](#) [\[PubMed\]](#)
54. Colasurdo, M.; Leibushor, N.; Robledo, A.; Vasandani, V.; Luna, Z.A.; Rao, A.S.; Garcia, R.; Srinivasan, V.M.; Sheth, S.A.; Avni, N.; et al. Automated detection and analysis of subdural hematomas using a machine learning algorithm. *J. Neurosurg.* **2022**, *138*, 1077–1084. [\[CrossRef\]](#) [\[PubMed\]](#)
55. Seyam, M.; Weikert, T.; Sauter, A.; Brehm, A.; Psychogios, M.; Blackham, K.A. Utilization of artificial intelligence-based intracranial hemorrhage detection on emergent noncontrast CT images in clinical workflow. *Radiol. Artif. Intell.* **2022**, *4*, e210168. [\[CrossRef\]](#)
56. Matsoukas, S.; Scaggiante, J.; Schuldt, B.R.; Smith, C.J.; Chennareddy, S.; Kalagara, R.; Majidi, S.; Bederson, J.B.; Fifi, J.T.; Mocco, J.; et al. Accuracy of artificial intelligence for the detection of intracranial hemorrhage and chronic cerebral microbleeds: A systematic review and pooled analysis. *Radiol. Medica* **2022**, *127*, 1106–1123. [\[CrossRef\]](#)
57. RSNA Intracranial Hemorrhage Detection. Available online: <https://www.kaggle.com/competitions/rsna-intracranial-hemorrhage-detection/overview> (accessed on 26 January 2023).

58. Flanders, A.E.; Prevedello, L.M.; Shih, G.; Halabi, S.S.; Kalpathy-Cramer, J.; Ball, R.; Mongan, J.T.; Stein, A.; Kitamura, F.C.; Lungren, M.P.; et al. Construction of a machine learning dataset through collaboration: The RSNA 2019 brain CT hemorrhage challenge. *Radiol. Artif. Intell.* **2020**, *2*, e190211. [\[CrossRef\]](#)
59. Wang, X.; Shen, T.; Yang, S.; Lan, J.; Xu, Y.; Wang, M.; Zhang, J.; Han, X. A deep learning algorithm for automatic detection and classification of acute intracranial hemorrhages in head CT scans. *Neuroimage Clin.* **2021**, *32*, 102785. [\[CrossRef\]](#)
60. Zhao, X.; Chen, K.; Wu, G.; Zhang, G.; Zhou, X.; Lv, C.; Wu, S.; Chen, Y.; Xie, G.; Yao, Z. Deep learning shows good reliability for automatic segmentation and volume measurement of brain hemorrhage, intraventricular extension, and peripheral edema. *Eur. Radiol.* **2021**, *31*, 5012–5020. [\[CrossRef\]](#)
61. Patel, A.; Schreuder, F.H.B.M.; Klijn, C.J.M.; Prokop, M.; Ginneken, B.V.; Marquering, H.A.; Roos, Y.B.W.E.M.; Baharoglu, M.I.; Meijer, F.J.A.; Manniesing, R. Intracerebral haemorrhage segmentation in non-contrast CT. *Sci. Rep.* **2019**, *9*, 17858. [\[CrossRef\]](#)
62. Islam, M.; Sanghani, P.; See, A.A.Q.; James, M.L.; King, N.K.K.; Ren, H. ICHNet: Intracerebral hemorrhage (ICH) segmentation using deep learning. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 456–463.
63. Kok, Y.E.; Pszczolkowski, S.; Law, Z.K.; Ali, A.; Krishnan, K.; Bath, P.M.; Sprigg, N.; Dineen, R.A.; French, A.P. Semantic segmentation of spontaneous intracerebral hemorrhage, intraventricular hemorrhage, and associated edema on CT images using deep learning. *Radiol. Artif. Intell.* **2022**, *4*, e220096. [\[CrossRef\]](#)
64. Porter, K.R.; McCarthy, B.J.; Freels, S.; Kim, Y.; Davis, F.G. Prevalence estimates for primary brain tumors in the United States by age, gender, behavior, and histology. *Neuro Oncol.* **2010**, *12*, 520–527. [\[CrossRef\]](#) [\[PubMed\]](#)
65. Ostrom, Q.T.; Patil, N.; Cioffi, G.; Waite, K.; Kruchko, C.; Barnholtz-Sloan, J.S. CBTRUS statistical report: Primary brain and other central nervous system tumors diagnosed in the United States in 2013–2017. *Neuro Oncol.* **2020**, *22* (Suppl. S2), iv1–iv96. [\[CrossRef\]](#) [\[PubMed\]](#)
66. Fan, Y.; Zhang, X.; Gao, C.; Jiang, S.; Wu, H.; Liu, Z.; Dou, T. Burden and trends of brain and central nervous system cancer from 1990 to 2019 at the global, regional, and country levels. *Arch. Public Health* **2022**, *80*, 209. [\[CrossRef\]](#) [\[PubMed\]](#)
67. Liu, Q.; Tong, X.; Wang, J. Management of brain metastases: History and the present. *Chin. Neurosurg. J.* **2019**, *5*, 1. [\[CrossRef\]](#)
68. Sperduto, P.W.; Mesko, S.; Li, J.; Cagney, D.; Aizer, A.; Lin, N.U.; Nesbit, E.; Kruser, T.J.; Chan, J.; Braunstein, S.; et al. Survival in patients with brain metastases: Summary report on the updated diagnosis-specific graded prognostic assessment and definition of the eligibility quotient. *J. Clin. Oncol.* **2020**, *38*, 3773–3784. [\[CrossRef\]](#)
69. Shur, J.D.; Doran, S.J.; Kumar, S.; Ap Dafydd, D.; Downey, K.; O'Connor, J.P.B.; Papanikolaou, N.; Messiou, C.; Koh, D.M.; Orton, M.R. Radiomics in oncology: A practical guide. *Radiographics* **2021**, *41*, 1717–1732. [\[CrossRef\]](#)
70. Ozkara, B.B.; Chen, M.M.; Federau, C.; Karabacak, M.; Briere, T.M.; Li, J.; Wintermark, M. Deep learning for detecting brain metastases on MRI: A systematic review and meta-analysis. *Cancers* **2023**, *15*, 334. [\[CrossRef\]](#)
71. Cho, S.J.; Sunwoo, L.; Baik, S.H.; Bae, Y.J.; Choi, B.S.; Kim, J.H. Brain metastasis detection using machine learning: A systematic review and meta-analysis. *Neuro Oncol.* **2021**, *23*, 214–225. [\[CrossRef\]](#)
72. Kikuchi, Y.; Togao, O.; Kikuchi, K.; Momosaka, D.; Obara, M.; Van Cauteren, M.; Fischer, A.; Ishigami, K.; Hiwatashi, A. A deep convolutional neural network-based automatic detection of brain metastases with and without blood vessel suppression. *Eur. Radiol.* **2022**, *5*, 2998–3005. [\[CrossRef\]](#) [\[PubMed\]](#)
73. Grovik, E.; Yi, D.; Iv, M.; Tong, E.; Rubin, D.; Zaharchuk, G. Deep learning enables automatic detection and segmentation of brain metastases on multisequence MRI. *J. Magn. Reson. Imaging* **2020**, *51*, 175–182. [\[CrossRef\]](#)
74. Yin, S.; Luo, X.; Yang, Y.; Shao, Y.; Ma, L.; Lin, C.; Yang, Q.; Wang, D.; Luo, Y.; Mai, Z.; et al. Development and validation of a deep-learning model for detecting brain metastases on 3D post-contrast MRI: A multi-center multi-reader evaluation study. *Neuro Oncol.* **2022**, *24*, 1559–1570. [\[CrossRef\]](#)
75. Huang, Y.; Bert, C.; Sommer, P.; Frey, B.; Gaip, U.; Distel, L.V.; Weissmann, T.; Uder, M.; Schmidt, M.A.; Dörfler, A.; et al. Deep learning for brain metastasis detection and segmentation in longitudinal MRI data. *Med. Phys.* **2022**, *49*, 5773–5786. [\[CrossRef\]](#)
76. Badrigilan, S.; Nabavi, S.; Abin, A.A.; Rostampour, N.; Abedi, I.; Shirvani, A.; Ebrahimi Moghaddam, M. Deep learning approaches for automated classification and segmentation of head and neck cancers and brain tumors in magnetic resonance images: A meta-analysis study. *Int. J. Comput. Assist. Radiol. Surg.* **2021**, *16*, 529–542. [\[CrossRef\]](#)
77. Das, S.; Nayak, G.K.; Saba, L.; Kalra, M.; Suri, J.S.; Saxena, S. An artificial intelligence framework and its bias for brain tumor segmentation: A narrative review. *Comput. Biol. Med.* **2022**, *143*, 105273. [\[CrossRef\]](#) [\[PubMed\]](#)
78. van Kempen, E.J.; Post, M.; Mannil, M.; Witkam, R.L.; Ter Laan, M.; Patel, A.; Meijer, F.J.A.; Henssen, D. Performance of machine learning algorithms for glioma segmentation of brain MRI: A systematic literature review and meta-analysis. *Eur. Radiol.* **2021**, *31*, 9638–9653. [\[CrossRef\]](#) [\[PubMed\]](#)
79. Shaikat, Z.; Farooq, Q.U.A.; Tu, S.; Xiao, C.; Ali, S. A state-of-the-art technique to perform cloud-based semantic segmentation using deep learning 3D U-net architecture. *BMC Bioinform.* **2022**, *23*, 251–259. [\[CrossRef\]](#)
80. Sidibe, I.; Tensaouti, F.; Roques, M.; Cohen-Jonathan-Moyal, E.; Laprie, A. Pseudoprogression in glioblastoma: Role of metabolic and functional MRI-systematic review. *Biomedicine* **2022**, *10*, 285. [\[CrossRef\]](#) [\[PubMed\]](#)
81. Sun, Y.; Yan, L.; Han, Y.; Nan, H.Y.; Xiao, G.; Tian, Q.; Pu, W.H.; Li, Z.Y.; Wei, X.C.; Wang, W.; et al. Differentiation of pseudoprogression from true progression in glioblastoma patients after standard treatment: A machine learning strategy combined with radiomics features from T(1)-weighted contrast-enhanced imaging. *BMC Med. Imaging* **2021**, *21*, 17. [\[CrossRef\]](#)

82. Rathore, S.; Akbari, H.; Doshi, J.; Shukla, G.; Rozycki, M.; Bilello, M.; Lustig, R.; Davatzikos, C. Radiomic signature of infiltration in peritumoral edema predicts subsequent recurrence in glioblastoma: Implications for personalized radiotherapy planning. *J. Med. Imaging* **2018**, *5*, 021219. [\[CrossRef\]](#)
83. Chang, P.; Grinband, J.; Weinberg, B.D.; Bardis, M.; Khy, M.; Cadena, G.; Su, M.Y.; Cha, S.; Filippi, C.G.; Bota, D.; et al. Deep-learning convolutional neural networks accurately classify genetic mutations in gliomas. *AJNR Am. J. Neuroradiol.* **2018**, *39*, 1201–1207. [\[CrossRef\]](#)
84. Wardlaw, J.M.; Valdes Hernandez, M.C.; Munoz-Maniega, S. What are white matter hyperintensities made of? relevance to vascular cognitive impairment. *J. Am. Heart Assoc.* **2015**, *4*, 001140. [\[CrossRef\]](#) [\[PubMed\]](#)
85. Guerrero, R.; Qin, C.; Oktay, O.; Bowles, C.; Chen, L.; Joules, R.; Wolz, R.; Valdés-Hernández, M.C.; Dickie, D.A.; Wardlaw, J.; et al. White matter hyperintensity and stroke lesion segmentation and differentiation using convolutional neural networks. *Neuroimage Clin.* **2017**, *17*, 918–934. [\[CrossRef\]](#) [\[PubMed\]](#)
86. Debette, S.; Markus, H.S. The clinical importance of white matter hyperintensities on brain magnetic resonance imaging: Systematic review and meta-analysis. *BMJ* **2010**, *341*, c3666. [\[CrossRef\]](#) [\[PubMed\]](#)
87. Hasan, T.F.; Barrett, K.M.; Brott, T.G.; Badi, M.K.; Lesser, E.R.; Hodge, D.O.; Meschia, J.F. Severity of white matter hyperintensities and effects on all-cause mortality in the Mayo Clinic Florida familial cerebrovascular diseases registry. *Mayo Clin. Proc.* **2019**, *94*, 408–416. [\[CrossRef\]](#)
88. Habes, M.; Erus, G.; Toledo, J.B.; Zhang, T.; Bryan, N.; Launer, L.J.; Rosseel, Y.; Janowitz, D.; Doshi, J.; Van der Auwera, S.; et al. White matter hyperintensities and imaging patterns of brain ageing in the general population. *Brain* **2016**, *139*, 1164–1179. [\[CrossRef\]](#)
89. Wardlaw, J.M.; Smith, E.E.; Biessels, G.J.; Cordonnier, C.; Fazekas, F.; Frayne, R.; Lindley, R.I.; O'Brien, J.T.; Barkhof, F.; Benavente, O.R.; et al. Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. *Lancet Neurol.* **2013**, *12*, 822–838. [\[CrossRef\]](#) [\[PubMed\]](#)
90. Giorgio, A.; De Stefano, N. Clinical use of brain volumetry. *J. Magn. Reson. Imaging* **2013**, *37*, 1–14. [\[CrossRef\]](#)
91. Driscoll, I.; Davatzikos, C.; An, Y.; Wu, X.; Shen, D.; Kraut, M.; Resnick, S.M. Longitudinal pattern of regional brain volume change differentiates normal aging from MCI. *Neurology* **2009**, *72*, 1906–1913. [\[CrossRef\]](#)
92. De Bresser, J.; Tiehuis, A.M.; Van Den Berg, E.; Reijmer, Y.D.; Jongen, C.; Kappelle, L.J.; Mali, W.P.; Viergever, M.A.; Biessels, G.J.; Utrecht Diabetic Encephalopathy Study Group. Progression of cerebral atrophy and white matter hyperintensities in patients with type 2 diabetes. *Diabetes Care* **2010**, *33*, 1309–1314. [\[CrossRef\]](#)
93. De Bresser, J.; Reijmer, Y.D.; Van Den Berg, E.; Breedijk, M.A.; Kappelle, L.J.; Viergever, M.A.; Biessels, G.J.; Utrecht Diabetic Encephalopathy Study Group. Microvascular determinants of cognitive decline and brain volume change in elderly patients with type 2 diabetes. *Dement. Geriatr. Cogn. Disord.* **2010**, *30*, 381–386. [\[CrossRef\]](#)
94. Ikram, M.A.; Vrooman, H.A.; Vernooij, M.W.; van der Lijn, F.; Hofman, A.; van der Lugt, A.; Niessen, W.J.; Breteler, M.M. Brain tissue volumes in the general elderly population: The rotterdam scan study. *Neurobiol. Aging* **2008**, *29*, 882–890. [\[CrossRef\]](#)
95. de Groot, J.C.; De Leeuw, F.; Oudkerk, M.; Hofman, A.; Jolles, J.; Breteler, M. Cerebral white matter lesions and subjective cognitive dysfunction: The rotterdam scan study. *Neurology* **2001**, *56*, 1539–1545. [\[CrossRef\]](#)
96. Moeskops, P.; de Bresser, J.; Kuijf, H.J.; Mendrik, A.M.; Biessels, G.J.; Pluim, J.P.W.; Išgum, I. Evaluation of a deep learning approach for the segmentation of brain tissues and white matter hyperintensities of presumed vascular origin in MRI. *NeuroImage Clin.* **2018**, *17*, 251–262. [\[CrossRef\]](#)
97. Omoumi, P.; Ducarouge, A.; Tournier, A.; Harvey, H.; Kahn, C.E., Jr.; Louvet-de Verchère, F.; Pinto Dos Santos, D.; Kober, T.; Richiardi, J. To buy or not to buy—Evaluating commercial AI solutions in radiology (the ECLAIR guidelines). *Eur. Radiol.* **2021**, *31*, 3786–3796. [\[CrossRef\]](#) [\[PubMed\]](#)
98. Tran, P.; Thoprakarn, U.; Gourieux, E.; Dos Santos, C.L.; Cavedo, E.; Guizard, N.; Cotton, F.; Krolak-Salmon, P.; Delmaire, C.; Heidelberg, D.; et al. Automatic segmentation of white matter hyperintensities: Validation and comparison with state-of-the-art methods on both multiple sclerosis and elderly subjects. *NeuroImage Clin.* **2022**, *33*, 102940. [\[CrossRef\]](#) [\[PubMed\]](#)
99. Gibson, E.; Gao, F.; Black, S.E.; Lobaugh, N.J. Automatic segmentation of white matter hyperintensities in the elderly using FLAIR images at 3T. *J. Magn. Reson. Imaging* **2010**, *31*, 1311–1322. [\[CrossRef\]](#)
100. Yaakub, S.N.; Heckemann, R.A.; Keller, S.S.; McGinnity, C.J.; Weber, B.; Hammers, A. On brain atlas choice and automatic segmentation methods: A comparison of MAPER & FreeSurfer using three atlas databases. *Sci. Rep.* **2020**, *10*, 2837. [\[PubMed\]](#)
101. Kelly, B.S.; Judge, C.; Bollard, S.M.; Clifford, S.M.; Healy, G.M.; Aziz, A.; Mathur, P.; Islam, S.; Yeom, K.W.; Lawlor, A.; et al. Correction to: Radiology Artificial Intelligence: A systematic review and evaluation of methods (RAISE). *Eur. Radiol.* **2022**, *32*, 8054. [\[CrossRef\]](#)
102. Li, H.; Jiang, G.; Zhang, J.; Wang, R.; Wang, Z.; Zheng, W.S.; Menze, B. Fully convolutional network ensembles for white matter hyperintensities segmentation in MR images. *Neuroimage* **2018**, *183*, 650–665. [\[CrossRef\]](#)
103. Kuijf, H.J.; Biesbroek, J.M.; De Bresser, J.; Heinen, R.; Andermatt, S.; Bento, M.; Berseth, M.; Belyaev, M.; Cardoso, M.J.; Casamitjana, A.; et al. Standardized assessment of automatic segmentation of white matter hyperintensities and results of the WMH segmentation challenge. *IEEE Trans. Med. Imaging* **2019**, *38*, 2556–2568. [\[CrossRef\]](#)
104. Collins, D.L.; Zijdenbos, A.P.; Kollokian, V.; Sled, J.G.; Kabani, N.J.; Holmes, C.J.; Evans, A.C. Design and construction of a realistic digital brain phantom. *IEEE Trans. Med. Imaging* **1998**, *17*, 463–468. [\[CrossRef\]](#) [\[PubMed\]](#)

105. Mendrik, A.M.; Vincken, K.L.; Kuijf, H.J.; Breeuwer, M.; Bouvy, W.H.; de Bresser, J.; Alansary, A.; de Bruijne, M.; Carass, A.; El-Baz, A.; et al. MRBrainS challenge: Online evaluation framework for brain image segmentation in 3T MRI scans. *Comput. Intell. Neurosci.* **2015**, *2015*, 813696. [\[CrossRef\]](#) [\[PubMed\]](#)
106. Menze, B.H.; Jakab, A.; Bauer, S.; Kalpathy-Cramer, J.; Farahani, K.; Kirby, J.; Burren, Y.; Porz, N.; Slotboom, J.; Wiest, R.; et al. The multimodal brain tumor image segmentation benchmark (BRATS). *TMI* **2015**, *34*, 1993–2024. [\[CrossRef\]](#) [\[PubMed\]](#)
107. Styner, M.; Lee, J.; Chin, B.; Chin, M.; Commowick, O.; Tran, H.; Markovic-Plese, S.; Jewells, V.; Warfield, S. 3D segmentation in the clinic: A grand challenge II: MS lesion segmentation. *MIDAS J.* **2008**. [\[CrossRef\]](#)
108. Commowick, O.; Istace, A.; Kain, M.; Laurent, B.; Leray, F.; Simon, M.; Pop, S.C.; Girard, P.; Améli, R.; Ferré, J.C.; et al. Objective evaluation of multiple sclerosis lesion segmentation using a data management and processing infrastructure. *Sci. Rep.* **2018**, *8*, 13650. [\[CrossRef\]](#)
109. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.
110. Liu, L.; Chen, S.; Zhu, X.; Zhao, X.; Wu, F.; Wang, J. Deep convolutional neural network for accurate segmentation and quantification of white matter hyperintensities. *Neurocomputing* **2020**, *384*, 231–242. [\[CrossRef\]](#)
111. Elizar, E.; Zulkifley, M.A.; Muharar, R.; Zaman, M.H.M.; Mustaza, S.M. A review on multiscale-deep-learning applications. *Sensors* **2022**, *22*, 7384. [\[CrossRef\]](#)
112. Walton, C.; King, R.; Rechtman, L.; Kaye, W.; Leray, E.; Marrie, R.A.; Robertson, N.; La Rocca, N.; Uitdehaag, B.; van der Mei, I.; et al. Rising prevalence of multiple sclerosis worldwide: Insights from the atlas of MS, third edition. *Mult. Scler.* **2020**, *26*, 1816–1821. [\[CrossRef\]](#)
113. Carass, A.; Roy, S.; Jog, A.; Cuzzocreo, J.L.; Magrath, E.; Gherman, A.; Button, J.; Nguyen, J.; Prados, F.; Sudre, C.H.; et al. Longitudinal multiple sclerosis lesion segmentation: Resource and challenge. *NeuroImage* **2017**, *148*, 77–102. [\[CrossRef\]](#)
114. Commowick, O.; Cervenansky, F.; Cotton, F.; Dojat, M. MSSEG-2 Challenge Proceedings: Multiple Sclerosis New Lesions Segmentation Challenge Using a Data Management and Processing Infrastructure. In Proceedings of the MICCAI 2021—24th International Conference on Medical Image Computing and Computer Assisted Intervention, Strasbourg, France, 27 September–1 October 2021; p. 126. Available online: <https://hal.inria.fr/hal-03358968> (accessed on 15 May 2023).
115. Hudomiet, P.; Hurd, M.D.; Rohwedder, S. Dementia prevalence in the United States in 2000 and 2012: Estimates based on a nationally representative study. *J. Gerontol. B Psychol. Sci. Soc. Sci.* **2018**, *73* (Suppl. S1), S10–S19. [\[CrossRef\]](#)
116. Gale, S.A.; Acar, D.; Daffner, K.R. Dementia. *Am. J. Med.* **2018**, *131*, 1161–1169. [\[CrossRef\]](#)
117. Jack, C.R.J.; Wiste, H.J.; Vemuri, P.; Weigand, S.D.; Senjem, M.L.; Zeng, G.; Bernstein, M.A.; Gunter, J.L.; Pankratz, V.S.; Aisen, P.S.; et al. Brain beta-amyloid measures and magnetic resonance imaging atrophy both predict time-to-progression from mild cognitive impairment to Alzheimer’s disease. *Brain* **2010**, *133*, 3336–3348. [\[CrossRef\]](#) [\[PubMed\]](#)
118. Lang, L.; Clifford, A.; Wei, L.; Zhang, D.; Leung, D.; Augustine, G.; Danat, I.M.; Zhou, W.; Copeland, J.R.; Anstey, K.J.; et al. Prevalence and determinants of undetected dementia in the community: A systematic literature review and a meta-analysis. *BMJ Open* **2017**, *7*, e011146. [\[CrossRef\]](#) [\[PubMed\]](#)
119. Alzheimer’s Association. 2022 Alzheimer’s disease facts and figures. *Alzheimers Dement.* **2022**, *18*, 700–789. [\[CrossRef\]](#)
120. Prince, M.B.R.; Ferri, C. *World Alzheimer Report 2011: The Benefits of Early Diagnosis and Intervention*; Institute of Psychiatry, King’s College: London, UK, 2011.
121. Kleiman, M.J.; Plewes, A.D.; Owora, A.; Grout, R.W.; Dexter, P.R.; Fowler, N.R.; Galvin, J.E.; Miled, Z.B.; Boustani, M. Digital detection of dementia (D(3)): A study protocol for a pragmatic cluster-randomized trial examining the application of patient-reported outcomes and passive clinical decision support systems. *Trials* **2022**, *23*, 868. [\[CrossRef\]](#) [\[PubMed\]](#)
122. Bron, E.E.; Klein, S.; Reinke, A.; Papma, J.M.; Maier-Hein, L.; Alexander, D.C.; Oxtoby, N.P. Ten years of image analysis and machine learning competitions in dementia. *Neuroimage* **2022**, *253*, 119083. [\[CrossRef\]](#)
123. Allen, G.I.; Amoroso, N.; Anghel, C.; Balagurusamy, V.; Bare, C.J.; Beaton, D.; Bellotti, R.; Bennett, D.A.; Boehme, K.L.; Boutros, P.C.; et al. Crowdsourced estimation of cognitive decline and resilience in Alzheimer’s disease. *Alzheimers Dement.* **2016**, *12*, 645–653. [\[CrossRef\]](#) [\[PubMed\]](#)
124. Marinescu, R.V.; Oxtoby, N.P.; Young, A.L.; Bron, E.E.; Toga, A.W.; Weiner, M.W.; Barkhof, F.; Fox, N.C.; Eshaghi, A.; Toni, T.; et al. The Alzheimer’s disease prediction of longitudinal evolution (TADPOLE) challenge: Results after 1 year follow up. *J. Mach. Learn. Biomed. Imaging* **2021**, *1*, 1–60. [\[CrossRef\]](#)
125. US Preventive Services Task Force; Owens, D.K.; Davidson, K.W.; Krist, A.H.; Barry, M.J.; Cabana, M.; Caughey, A.B.; Doubeni, C.A.; Epling, J.W., Jr.; Kubik, M.; et al. Screening for cognitive impairment in older adults: US Preventive Services Task Force recommendation statement. *JAMA* **2020**, *323*, 757–763. [\[CrossRef\]](#)
126. Fisch, L.; Leenings, R.; Winter, N.R.; Dannlowski, U.; Gaser, C.; Cole, J.H.; Hahn, T. Editorial: Predicting chronological age from structural neuroimaging: The predictive analytics competition 2019. *Front. Psychiatry* **2021**, *12*, 710932. [\[CrossRef\]](#)
127. Gong, W.; Beckmann, C.F.; Vedaldi, A.; Smith, S.M.; Peng, H. Optimising a simple fully convolutional network for accurate brain age prediction in the PAC 2019 challenge. *Front. Psychiatry* **2021**, *12*, 627996. [\[CrossRef\]](#) [\[PubMed\]](#)
128. Bron, E.E.; Smits, M.; van der Flier, W.M.; Vrenken, H.; Barkhof, F.; Scheltens, P.; Papma, J.M.; Steketee, R.M.; Méndez Orellana, C.; Meijboom, R.; et al. Standardized evaluation of algorithms for computer-aided diagnosis of dementia based on structural MRI: The CADDementia challenge. *Neuroimage* **2015**, *111*, 562–579. [\[CrossRef\]](#) [\[PubMed\]](#)

129. Folego, G.; Weiler, M.; Casseb, R.F.; Pires, R.; Rocha, A. Alzheimer's disease detection through whole-brain 3D-CNN MRI. *Front. Bioeng. Biotechnol.* **2020**, *8*, 534592. [\[CrossRef\]](#) [\[PubMed\]](#)
130. Sorensen, L.; Igel, C.; Pai, A.; Balas, I.; Anker, C.; Lillholm, M.; Nielsen, M.; Alzheimer's Disease Neuroimaging Initiative and the Australian Imaging Biomarkers and Lifestyle flagship study of ageing. Differential diagnosis of mild cognitive impairment and Alzheimer's disease using structural MRI cortical thickness, hippocampal shape, hippocampal texture, and volumetry. *Neuroimage Clin.* **2016**, *13*, 470–482. [\[CrossRef\]](#)
131. Vieira, S.; Pinaya, W.H.L.; Mechelli, A. Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: Methods and applications. *Neurosci. Biobehav. Rev.* **2017**, *74 Pt A*, 58–75. [\[CrossRef\]](#)
132. Marcisz, A.; Alzheimer's Disease Neuroimaging Initiative; Polanska, J. Can T1-weighted magnetic resonance imaging significantly improve mini-mental state examination-based distinguishing between mild cognitive impairment and early-stage Alzheimer's disease? *J. Alzheimers Dis.* **2023**, *92*, 941–957. [\[CrossRef\]](#)
133. Beckett, L.A.; Harvey, D.J.; Gamst, A.; Donohue, M.; Kornak, J.; Zhang, H.; Kuo, J.H.; Alzheimer's Disease Neuroimaging Initiative. The Alzheimer's Disease Neuroimaging Initiative: Annual change in biomarkers and clinical outcomes. *Alzheimers Dement.* **2010**, *6*, 257–264. [\[CrossRef\]](#)
134. Cash, D.M.; Frost, C.; Iheme, L.O.; Ünay, D.; Kandemir, M.; Fripp, J.; Salvado, O.; Bourgeat, P.; Reuter, M.; Fischl, B.; et al. Assessing atrophy measurement techniques in dementia: Results from the MIRIAD atrophy challenge. *Neuroimage* **2015**, *123*, 149–164. [\[CrossRef\]](#)
135. Suk, H.; Lee, S.; Shen, D.; Alzheimer's Disease Neuroimaging Initiative. Latent feature representation with stacked auto-encoder for AD/MCI diagnosis. *Brain Struct. Funct.* **2015**, *220*, 841–859. [\[CrossRef\]](#)
136. Hoy, D.; March, L.; Brooks, P.; Blyth, F.; Woolf, A.; Bain, C.; Williams, G.; Smith, E.; Vos, T.; Barendregt, J.; et al. The global burden of low back pain: Estimates from the global burden of disease 2010 study. *Ann. Rheum. Dis.* **2014**, *73*, 968–974. [\[CrossRef\]](#)
137. Hoy, D.; Brooks, P.; Blyth, F.; Buchbinder, R. The epidemiology of low back pain. *Best Pract. Res. Clin. Rheumatol.* **2010**, *24*, 769–781. [\[CrossRef\]](#) [\[PubMed\]](#)
138. Cui, Y.; Zhu, J.; Duan, Z.; Liao, Z.; Wang, S.; Liu, W. Artificial intelligence in spinal imaging: Current status and future directions. *Int. J. Environ. Res. Public Health* **2022**, *19*, 11708. [\[CrossRef\]](#) [\[PubMed\]](#)
139. Nouh, M.R. Imaging of the spine: Where do we stand? *World J. Radiol.* **2019**, *11*, 55–61. [\[CrossRef\]](#) [\[PubMed\]](#)
140. Sekuboyina, A.; Hussein, M.E.; Bayat, A.; Löffler, M.; Liebl, H.; Li, H.; Tetteh, G.; Kukačka, J.; Payer, C.; Štern, D.; et al. VerSe: A vertebrae labelling and segmentation benchmark for multi-detector CT images. *Med. Image Anal.* **2021**, *73*, 102166. [\[CrossRef\]](#) [\[PubMed\]](#)
141. Li, X.; Dou, Q.; Chen, H.; Fu, C.W.; Qi, X.; Belavý, D.L.; Ambrecht, G.; Felsenberg, D.; Zheng, G.; Heng, P.A. 3D multi-scale FCN with random modality voxel dropout learning for intervertebral disc localization and segmentation from multi-modality MR images. *Med. Image Anal.* **2018**, *45*, 41–54. [\[CrossRef\]](#)
142. Kaka, H.; Zhang, E.; Khan, N. Artificial intelligence and deep learning in neuroradiology: Exploring the new frontier. *Can. Assoc. Radiol. J.* **2021**, *72*, 35–44. [\[CrossRef\]](#)
143. Altini, N.; De Giosa, G.; Fragasso, N.; Coscia, C.; Sibilano, E.; Prencipe, B.; Hussain, S.M.; Brunetti, A.; Buongiorno, D.; Guerriero, A.; et al. Segmentation and identification of vertebrae in CT scans using CNN, k-means clustering and k-NN. *Informatics* **2021**, *8*, 40. [\[CrossRef\]](#)
144. Jamaludin, A.; Lootus, M.; Kadir, T.; Zisserman, A.; Urban, J.; Battié, M.C.; Fairbank, J.; McCall, I.; Genodisc Consortium. ISSLS PRIZE IN BIOENGINEERING SCIENCE 2017: Automation of reading of radiological features from magnetic resonance images (MRIs) of the lumbar spine without human intervention is comparable with an expert radiologist. *Eur. Spine J.* **2017**, *26*, 1374–1383. [\[CrossRef\]](#) [\[PubMed\]](#)
145. Gorelik, N.; Gyftopoulos, S. Applications of artificial intelligence in musculoskeletal imaging: From the request to the report. *Can. Assoc. Radiol. J.* **2021**, *72*, 45–59. [\[CrossRef\]](#)
146. Hallinan, J.T.P.D.; Zhu, L.; Yang, K.; Makmur, A.; Algazwi, D.A.R.; Thian, Y.L.; Lau, S.; Choo, Y.S.; Eide, S.E.; Yap, Q.V.; et al. Deep learning model for automated detection and classification of central canal, lateral recess, and neural foraminal stenosis at lumbar spine MRI. *Radiology* **2021**, *300*, 130–138. [\[CrossRef\]](#)
147. Burns, J.E.; Yao, J.; Summers, R.M. Vertebral body compression fractures and bone density: Automated detection and classification on CT images. *Radiology* **2017**, *284*, 788–797. [\[CrossRef\]](#) [\[PubMed\]](#)
148. Burns, J.E.; Yao, J.; Summers, R.M. Artificial intelligence in musculoskeletal imaging: A paradigm shift. *J. Bone Miner. Res.* **2020**, *35*, 28–35. [\[CrossRef\]](#) [\[PubMed\]](#)
149. Murata, K.; Endo, K.; Aihara, T.; Suzuki, H.; Sawaji, Y.; Matsuoka, Y.; Nishimura, H.; Takamatsu, T.; Konishi, T.; Maekawa, A.; et al. Artificial intelligence for the detection of vertebral fractures on plain spinal radiography. *Sci. Rep.* **2020**, *10*, 20031. [\[CrossRef\]](#)
150. Martín-Noguerol, T.; Oñate Miranda, M.; Amrhein, T.J.; Paulano-Godino, F.; Xiberta, P.; Vilanova, J.C.; Luna, A. The role of artificial intelligence in the assessment of the spine and spinal cord. *Eur. J. Radiol.* **2023**, *161*, 110726. [\[CrossRef\]](#)
151. Wang, L.; Xu, Q.; Leung, S.; Chung, J.; Chen, B.; Li, S. Accurate automated Cobb angles estimation using multi-view extrapolation net. *Med. Image Anal.* **2019**, *58*, 101542. [\[CrossRef\]](#)
152. Wang, L.; Xie, C.; Lin, Y.; Zhou, H.Y.; Chen, K.; Cheng, D.; Dubost, F.; Collery, B.; Khanal, B.; Khanal, B.; et al. Evaluation and comparison of accurate automated spinal curvature estimation algorithms with spinal anterior-posterior X-ray images: The AASCE2019 challenge. *Med. Image Anal.* **2021**, *72*, 102115. [\[CrossRef\]](#) [\[PubMed\]](#)

153. Zhang, J.; Li, H.; Lv, L.; Zhang, Y. Computer-aided cobb measurement based on automatic detection of vertebral slopes using deep neural network. *Int. J. Biomed. Imaging* **2017**, *2017*, 9083916. [\[CrossRef\]](#)
154. Pham, N.; Ju, C.; Kong, T.; Mukherji, S.K. Artificial intelligence in head and neck imaging. *Semin. Ultrasound CT MRI* **2022**, *43*, 170–175. [\[CrossRef\]](#) [\[PubMed\]](#)
155. Chen, M.; Chen, G.; Wang, D.; Zhang, J.; Di, R.; Li, F.; Zhou, Z.; Piao, S.; Li, Y.; Dai, Y. Deep learning-based segmentation of cerebral aneurysms in 3D TOF-MRA using coarse-to-fine framework. *arXiv* **2021**, arXiv:2110.13432.
156. DeJohn, C.R.; Grant, S.R.; Seshadri, M. Application of machine learning methods to improve the performance of ultrasound in head and neck oncology: A literature review. *Cancers* **2022**, *14*, 665. [\[CrossRef\]](#)
157. Santer, M.; Kloppenburg, M.; Gottfried, T.; Runge, A.; Schmutzhard, J.; Vorbach, S.M.; Mangesius, J.; Riedl, D.; Mangesius, S.; Widmann, G.; et al. Current applications of artificial intelligence to classify cervical lymph nodes in patients with head and neck squamous cell Carcinoma—A systematic review. *Cancers* **2022**, *14*, 5397. [\[CrossRef\]](#) [\[PubMed\]](#)
158. Volpe, S.; Pepa, M.; Zaffaroni, M.; Bellerba, F.; Santamaria, R.; Marvaso, G.; Isaksson, L.J.; Gandini, S.; Starzyńska, A.; Leonardi, M.C.; et al. Machine learning for head and neck cancer: A safe bet?—A clinically oriented systematic review for the radiation oncologist. *Front. Oncol.* **2021**, *11*, 772663. [\[CrossRef\]](#) [\[PubMed\]](#)
159. Deng, W.; Luo, L.; Lin, X.; Fang, T.; Liu, D.; Dan, G.; Chen, H. Head and neck cancer tumor segmentation using support vector machine in dynamic contrast-enhanced MRI. *Contrast Media Mol. Imaging* **2017**, *2017*, 8612519. [\[CrossRef\]](#) [\[PubMed\]](#)
160. Huang, W.; Chan, K.L.; Zhou, J. Region-based nasopharyngeal carcinoma lesion segmentation from MRI using clustering- and classification-based methods with learning. *J. Digit. Imaging* **2013**, *26*, 472–482. [\[CrossRef\]](#) [\[PubMed\]](#)
161. Al Ajmi, E.; Forghani, B.; Reinhold, C.; Bayat, M.; Forghani, R. Spectral multi-energy CT texture analysis with machine learning for tissue classification: An investigation using classification of benign parotid tumours as a testing paradigm. *Eur. Radiol.* **2018**, *28*, 2604–2611. [\[CrossRef\]](#) [\[PubMed\]](#)
162. Ranjbar, S.; Ning, S.; Zwart, C.M.; Wood, C.P.; Weindling, S.M.; Wu, T.; Mitchell, J.R.; Li, J.; Hoxworth, J.M. Computed tomography-based texture analysis to determine human papillomavirus status of oropharyngeal squamous cell carcinoma. *J. Comput. Assist. Tomogr.* **2018**, *42*, 299–305. [\[CrossRef\]](#) [\[PubMed\]](#)
163. Ramkumar, S.; Ranjbar, S.; Ning, S.; Lal, D.; Zwart, C.M.; Wood, C.P.; Weindling, S.M.; Wu, T.; Mitchell, J.R.; Li, J.; et al. MRI-based texture analysis to differentiate sinonasal squamous cell carcinoma from inverted papilloma. *AJNR Am. J. Neuroradiol.* **2017**, *38*, 1019–1025. [\[CrossRef\]](#)
164. Fujima, N.; Homma, A.; Harada, T.; Shimizu, Y.; Tha, K.K.; Kano, S.; Mizumachi, T.; Li, R.; Kudo, K.; Shirato, H. The utility of MRI histogram and texture analysis for the prediction of histological diagnosis in head and neck malignancies. *Cancer Imaging* **2019**, *19*, 5–9. [\[CrossRef\]](#)
165. Mahmood, H.; Shaban, M.; Rajpoot, N.; Khurram, S.A. Artificial intelligence-based methods in head and neck cancer diagnosis: An overview. *Br. J. Cancer* **2021**, *124*, 1934–1940. [\[CrossRef\]](#)
166. Chinnery, T.; Arifin, A.; Tay, K.Y.; Leung, A.; Nichols, A.C.; Palma, D.A.; Mattonen, S.A.; Lang, P. Utilizing artificial intelligence for head and neck cancer outcomes prediction from imaging. *Can. Assoc. Radiol. J.* **2021**, *72*, 73–85. [\[CrossRef\]](#)
167. Fu, F.; Wei, J.; Zhang, M.; Yu, F.; Xiao, Y.; Rong, D.; Shan, Y.; Li, Y.; Zhao, C.; Liao, F.; et al. Rapid vessel segmentation and reconstruction of head and neck angiograms using 3D convolutional neural network. *Nat. Commun.* **2020**, *11*, 4829. [\[CrossRef\]](#) [\[PubMed\]](#)
168. Park, A.; Chute, C.; Rajpurkar, P.; Lou, J.; Ball, R.L.; Shpanskaya, K.; Jabarkheel, R.; Kim, L.H.; McKenna, E.; Tseng, J.; et al. Deep Learning–Assisted diagnosis of cerebral aneurysms using the HeadXNet model. *JAMA Netw. Open* **2019**, *2*, e195600. [\[CrossRef\]](#) [\[PubMed\]](#)
169. Yang, H.; Cho, K.; Kim, J.; Kim, J.H.; Kim, Y.B.; Oh, J.H. Rupture risk prediction of cerebral aneurysms using a novel convolutional neural network-based deep learning model. *J. NeuroInterv. Surg.* **2022**, *15*, 200–204. [\[CrossRef\]](#) [\[PubMed\]](#)
170. Ryu, J.Y.; Hong, H.K.; Cho, H.G.; Lee, J.S.; Yoo, B.C.; Choi, M.H.; Chung, H.Y. Deep learning for the automatic segmentation of extracranial venous malformations of the head and neck from MR images using 3D U-net. *J. Clin. Med.* **2022**, *11*, 5593. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.