European Journal of Radiology 176 (2024) 111509

Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

Role of artificial intelligence in brain tumour imaging

Ezekiel Chukwujindu [a,*], Hafsa Faiz [a], Sara AI-Douri [b], Khunsa Faiz [c], Alexandra De Sequeira [d]

[a] Independent Scholar, Canada
[b] University of Guelph, N1G 2W1, Canada
[c] McMaster University, Department of Radiology, L8S 4L8, Canada
[d] Royal College of Surgeons in Ireland, 123 St. Stephen's, Green D02 YN77, Ireland

ARTICLE INFO

Keywords:
Artificial intelligence
Neuroradiology
Brain Tumours
Virtual biopsy
Metastatic Tumours
Transformers

ABSTRACT

Artificial intelligence (AI) is a rapidly evolving field with many neuro-oncology applications. In this review, we discuss how AI can assist in brain tumour imaging, focusing on machine learning (ML) and deep learning (DL) techniques. We describe how AI can help in lesion detection, differential diagnosis, anatomic segmentation, molecular marker identification, prognostication, and pseudo-progression evaluation. We also cover AI applications in non-glioma brain tumours, such as brain metastasis, posterior fossa, and pituitary tumours. We highlight the challenges and limitations of AI implementation in radiology, such as data quality, standardization, and integration. Based on the findings in the aforementioned areas, we conclude that AI can potentially improve the diagnosis and treatment of brain tumours and provide a path towards personalized medicine and better patient outcomes.

1. Introduction

Central nervous system cancer is the tenth leading cause of death in men and women [1]. Brain tumour is not the primary cause of mortality, yet 40 % of all other cancer types can develop into brain cancer due to metastasis [1]. The diagnosis of brain tumors is predominantly based on neuroimaging findings, using techniques such as contrast-enhanced computed tomography (CT) or magnetic resonance imaging (MRI).

Among the central nervous system (CNS) neoplasms, the most common type is Glioma originating from glial cells [2].

Gliomas are heterogeneous groups of disease, with many different histotypes and molecular subtypes ranging from slow growing pilocytic astrocytoma to the aggressive glioblastoma multiforme (GBM). Given poor prognosis of patients with brain cancer at a higher stage, accurate grading is crucial for treatment and prognosis.

Common tests used for tumour diagnosis and grade estimation include neurological examination, imaging, biopsies, and biomarkers. Biopsies are the gold standard, but invasive and risky.

Techniques such as contrast-enhanced computed tomography (CT) or magnetic resonance imaging (MRI) are used for diagnosis of brain tumours. As they are non-invasive and accessible, many efforts have been made to increase the information from brain imaging.

Conventional MRI sequences, which include pre- and postcontrast T1-weighted imaging, T2-weighted imaging, and T2-weighted fluid-attenuated inversion recovery (FLAIR) sequences, help delineate tumour volume and morphologic characteristics. Unfortunately, contrast enhancement is nonspecific, and the detection of foci of tumour infiltration within the T2-weighted FLAIR signal intensity abnormality is nearly impossible with conventional sequences [3].

Advanced MRI methods, including diffusion-weighted imaging, diffusion tensor imaging, perfusion MRI and MR spectroscopy, are used clinically for grading gliomas and identifying regions of tumour infiltration. They are usually qualitative and vary across sites, units, and methods. With increasing incidence of brain tumours, [4] a non-invasive, automatic computer-aided tool that can diagnose and grade a tumour quickly is needed.

One of the ways in which tumours can be swiftly diagnosed is through artificial intelligence. Artificial intelligence (AI) is defined as machines performing tasks characteristic of human intelligence [5]. AI-based algorithms have been used in the healthcare field to improve diagnosis, predict outcomes, guide efforts in drug discovery and for rapid data processing in clinical research. Moreover, neuroimaging research in AI has grown exponentially. There have been several articles published on the use of AI in brain tumour imaging. Ce et al [6] provided

* Corresponding author.
E-mail addresses: ezekiel.chukwujindu@mail.mcgill.ca (E. Chukwujindu), hafsa.faiz@medportal.ca (H. Faiz), asarar9@gmail.com (S. AI-Douri), khunsa.faiz@medportal.ca (K. Faiz), alexandradesequ22@rcsi.com (A. De Sequeira).

https://doi.org/10.1016/j.ejrad.2024.111509
Received 24 January 2024; Received in revised form 29 April 2024; Accepted 13 May 2024
Available online 17 May 2024

a description of AI-based models and a narrative review of their applications in various studies concerning brain imaging. Zhu et al [7] reviewed the latest machine learning-based AI applications in the radiomic analysis of brain tumours, providing a perspective on challenges and future avenues. However, other aspects of AI, such as its usage in non-glioma evaluations, as well as a discussion on the promise of the use of transformers in neuro-oncology imaging, is an important area of research. Thus, this article aims to evaluate the various uses of AI-assisted tools in the diagnosis and treatment of brain tumours, with a unique focus on brain gliomas as well as non-glioma evaluations, and transformer-based networks in brain tumour imaging.

2. Machine learning

Machine learning (ML), falling under the umbrella of AI, incorporates algorithms and statistical models to make predictions about new data points [8,9]. In ML, computers learn automatically from data accumulation and improve with experience.

Deep learning (DL) is a subclass of ML that processes raw unstructured data using multi-layered artificial neural networks (ANN) [8,9]. It is currently the basis of most of the AI tools used for image interpretation.. It can extract features, analyze patterns, and classify information by learning multiple levels of lower and higher-order features. Lower-order features, for example, would include corners, edges, and other basic shapes. Higher-order features would include different gradations of image texture, more refined shapes, and image-specific patterns.

However, in cases where there is a lack of diversity within the training data or bias, the reliability of integrating such platforms into healthcare settings may be called into question [10]. This can potentially result in platforms failing to recognize rarer Brain tumours or those in earlier stages of progression.

3. Artificial intelligence in image analysis

One of the most common applications of AI is in the analysis of diagnostic imaging.

The process (Fig. 6) often commences with transforming the raw visual data into a format comprehensible to various deep learning models. This transformation is critical, whether the model in use is a Convolutional Neural Network (CNN), Vision Transformer (ViT), or another advanced architecture [11,12]. The core essence of these models lies in their ability to interpret the visual content of images through a structured analytical process, meticulously crafted to extract, highlight, and contextualize relevant features before making well-informed predictions. The process typically begins with preprocessing, where images are tailored to meet the models' specific requirements. Tasks such as resizing, normalizing pixel values, and data augmentation are essential to ensure compatibility and enhance the robustness of the models. The prepared images are then fed into the models either as pixel matrices or segmented patches, setting the stage for intricate feature analysis [12,13].

For feature extraction, CNNs utilize convolutional layers to apply filters that capture distinctive features within the images, such as edges or textures. This process is complemented by activation functions (e.g., ReLU) to introduce non-linearity, enabling the model to learn complex patterns. Pooling layers, particularly max pooling, further distill the feature representation by reducing dimensionality and emphasizing the most salient features. ViTs, in contrast, dissect the image into a series of patches and employ self-attention mechanisms. This allows the model to assess the relevance of each patch in relation to others, fostering a deep understanding of contextual relationships and crafting a comprehensive representation of the image [11,12,13,14].

DL models typically undergo further stages for contextualization. Both CNNs and ViTs synthesize the extracted features to construct a holistic understanding of the image. CNNs might employ additional convolutional and fully connected layers for this purpose, whereas ViTs leverage ongoing transformer blocks to refine the contextual representation of each patch within the overarching image narrative.

The decisive phase is where models apply the accumulated knowledge to predict outcomes. CNNs typically progress through fully connected layers to an output layer for this task, whereas ViTs reach a climax at the classification head, directly translating the transformer-encoded features into predictions [11,12,14]. A common mechanism employed here, particularly for classification tasks, is the softmax function, which converts the models' output into class probabilities. The culmination of this analytical process is the model's prediction, which could manifest as class labels (classification), bounding boxes (detection), or pixel-wise annotations (segmentation), depending on the task at hand [13].

Following the comprehensive synthesis of visual data through deep learning models, the evaluation of these models becomes paramount to ensure their efficacy and accuracy in clinical settings. The effectiveness of models in tasks such as classification, detection, and segmentation is measured using specific, relevant metrics that directly impact clinical outcomes. In classification, accuracy, sensitivity, specificity, precision, and recall provide a quantitative assessment of a model's diagnostic ability, derived from a confusion matrix that compares predicted outcomes with actual clinical diagnoses [13]. Similarly, the F1 score offers a balance between precision and sensitivity, reflecting the model's diagnostic reliability across different decision thresholds. For tasks involving detection and segmentation, the Intersection over Union (IOU) and the Dice or Jaccard coefficients are critical for evaluating the alignment between the model-generated outputs and the expert-defined ground truths [13]. These metrics ensure that the models not only perform well technically but also meet the practical demands of medical diagnostics, making them valuable tools in the advancement of radiology and patient care.

It is essential to understand the unique architectural innovations introduced over time to tackle specific challenges in deep neural network design. U-Net was specifically developed for medical image segmentation, introducing a symmetric expanding path to achieve precise localization crucial for biomedical applications [15]. GoogLeNet and its Inception successors were designed to optimize deep neural network efficiency, using inception modules to manage computational budget and model size effectively [15]. ResNet addresses the vanishing gradient problem by incorporating skip connections that allow for training deeper networks than previously possible [15]. DenseNet201 enhances feature propagation and reduces the number of parameters through its densely connected architecture [11]. Xception improves upon the Inception models by employing depthwise separable convolutions, optimizing both performance and computational efficiency. MobileNet, ideal for mobile and edge devices, uses similar convolutions to ensure model compactness without sacrificing accuracy [11]. Capsule Networks (CapsNet) tackle traditional CNN limitations in capturing spatial hierarchies and object relationships, enhancing the network's ability to recognize objects from various viewpoints and orientations. These innovations ensure that deep learning models not only achieve high technical performance but also cater to the nuanced demands of medical diagnostics, thereby advancing the field of radiology and improving patient care [11,15].

Clinicians and radiologists are already utilizing ML models for non-invasive diagnoses as well as for treatment planning of brain tumours, a process sometimes referred to as "virtual biopsy". This enables them to obtain crucial information about tumour characteristics based on imaging features, including infiltrating tumour margins, molecular markers, and prognosis–all of which are relevant for patient management, pre and post-treatment follow-up, and therapeutic decision making [16].

As such, AI tools not only save radiologists' time, but also provide a second opinion in the diagnosis.

4. Ai in brain gliomas

4.1. Lesion detection and grade Prediction

AI can improve diagnosis of small lesions [17] by using MRI, CT and PET scan data. (Table 3). Small lesions affect treatment choices and are very relevant for DL algorithms [18]. CAD tools need to be tuned to ensure accuracy and reduce overdiagnosis and overtreatment (Fig. 1).

Blanc-Durand et al [19] used dynamic 18F-FET PET images to detect brain lesions in glioma patients. They used a 3D U-Net CNN to classify lesion or non-lesion voxels from PET features. They got 0.9868 accuracy in training and 0.9856 in validation for lesion detection.

AI can also predict tumour grade from medical images, which matters for glioma treatment and prognosis. AI can spot subtle features that clinicians may overlook. Different ML and DL methods with transfer learning, have high accuracy for glioma grading. For example, Support Vector Machine (SVM) models [20] had 0.987 Area Under the Curve (AUC) for low grade glioma (LGG) LGG vs high grade glioma (HGG). Deep learning methods [21] like GoogLeNet did better than traditional methods, with 0.939 test AUC.

4.2. Anatomic segmentation and Volumetry

Segmentation refers to identifying the boundaries of an object in the image. Frequently, the object is an organ, a tissue, a pathologic lesion, or another structure used for diagnosis or management of a particular disease.

Accurate segmentation of gliomas on routine MRI is crucial for diagnosis, treatment planning and prognosis [22], (Fig. 2) Traditional approaches to segmentation rely on manual, semiautomated or fully automated delineation of the object of interest. Currently, it is accomplished manually which involves separating tumour tissues, such as edema and necrosis, from normal brain tissue such as gray matter, white matter, and cerebrospinal fluid. However, this process is time consuming and is highly dependent on the subjective decisions of individual radiologists. As a result, the use of computational tools in glioma segmentation has been of high interest for radiologists, (Table 4) in their pursuit towards objectivity and accuracy (Fig. 3) [23].

Wu et al [22] created computational models in brain glioma segmentation that require only 5–7 s for segmenting one case. By using DL models like CNN, they were able to segment subregions of brain glioma with high accuracy (Sørensen–Dice scores of 0.80, 0.83, and 0.91 for enhancing tumor, tumor core, and whole tumor, respectively), efficiency, reliability, and generalization ability. AI-assisted segmentation tools are based on DL and ML algorithms. Currently, DL-based models have a greater impact on brain tumour segmentation and classification tasks compared to ML-based models.

In 2022, Akinyelu et al conducted a survey comparing the most recently developed segmentation techniques based on ML, CNN, Capsule Networks (CapsNet), and ViT. These methods primarily contribute to identifying the grade of a brain tumour and developing the best treatment for it [24].

CNN algorithms, though representing the most often used DL algorithm, pose some drawbacks, such as the need for large quantities of training datasets and the inability to correctly identify inputs with different rotations and transformations.

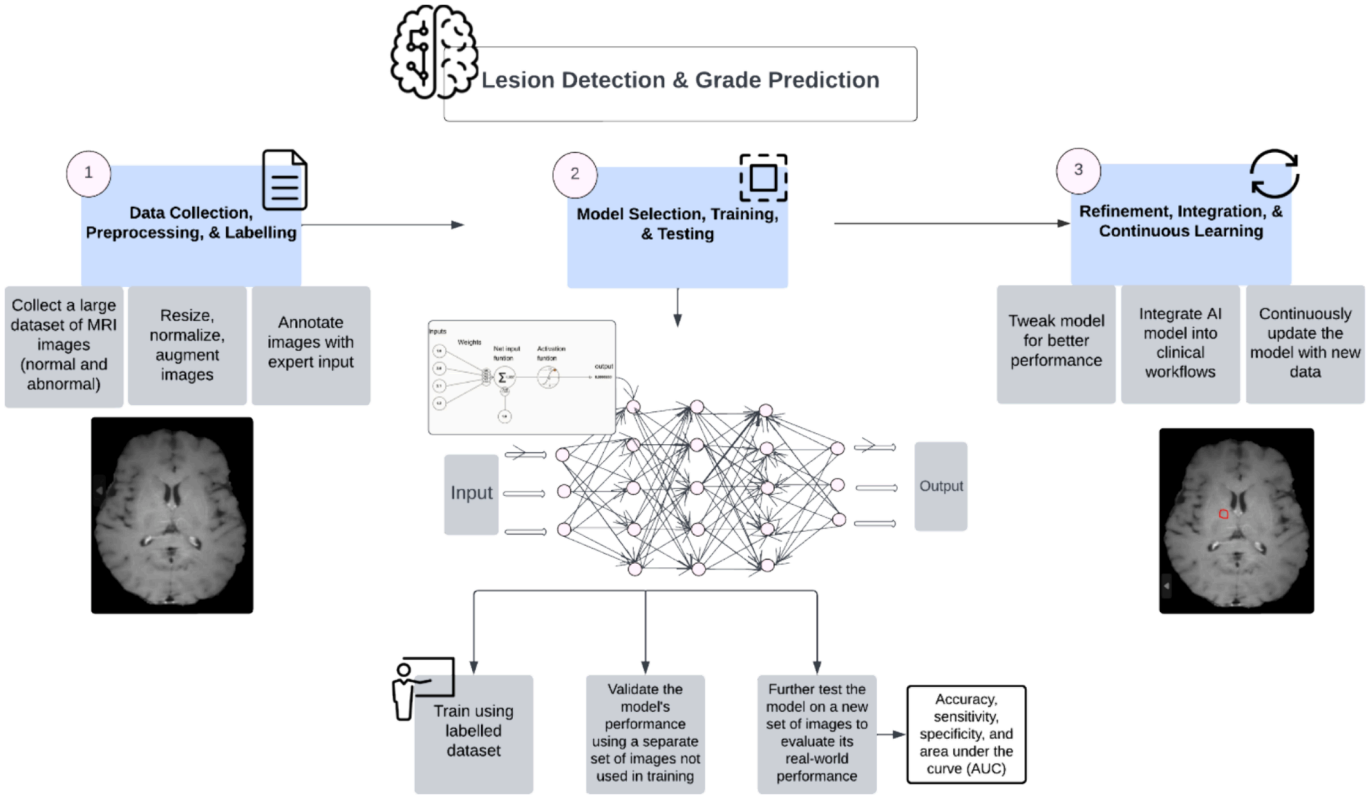

Fig. 1. Flowchart showing standard steps for lesion detection and grade prediction. 1. Collect a large dataset of MRI images, including both normal and abnormal (with lesions) images. Process the images to enhance quality & consistency. This can involve resizing, normalizing, and augmenting the images to improve the training process. Annotate the images with expert input, marking lesions accurately. This step is crucial for supervised learning models. 2. Choose an appropriate AI model for lesion detection (common choices include convolutional neural networks (CNNs) and deep learning models). Train the AI model using the labelled dataset; this involves feeding the images into the model, allowing it to learn differentiation between normal tissues and lesions. Validate the model's performance using a separate set of images not used in training–this step assesses the model's accuracy and generalizability. Further test the model on a new set of images to evaluate its real-world performance. Metrics like accuracy, sensitivity, specificity, and area under the curve (AUC) are often used. 3. Based on test results, refine and tweak the model for better performance. Integrate the AI model into clinical workflows, where it can assist radiologists in lesion detection on MRI images.

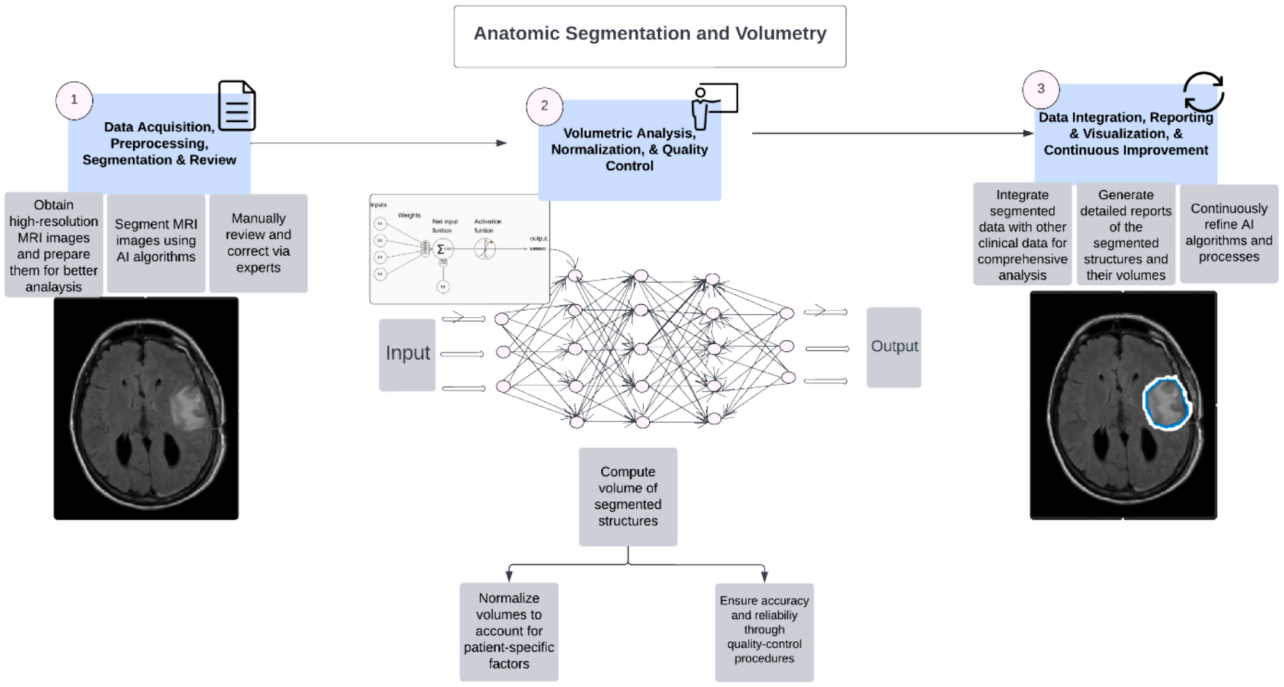

Fig. 2. Flowchart showing standard steps for Anatomic Segmentation and Volumetry. 1. Obtain high-resolution MRI images. The quality and resolution of these images is crucial for accurate segmentation and volumetry. Process the images, which includes noise reduction, contrast enhancement, and standardization of image dimensions. The goal is to prepare the images for better analysis and segmentation. Using AI algorithms, particularly those based on deep learning (CNNs), segment the MRI images to identify and delineate different anatomical structures. This involves classifying each pixel or voxel as belonging to a specific tissue or structure. Often, automated segmentation is followed by manual review and correction by experts, ensuring accuracy in delineating complex anatomical structures. 2. After segmentation, compute the volume of the segmented structures. This is particularly important in assessing organ sizes, tumour volumes, or changes in brain structures in neurological disorders. Normalize the volumes to account for patient-specific factors like overall head size or body size. This step is critical for comparative studies and clinical assessments. Ensure the accuracy and reliability of the segmentation and volumetry through quality control procedures, which might involve cross-verification with other imaging modalities or repeat analyses. 3. Integrate the segmented and volumetric data with other clinical data for comprehensive analysis, aiding in diagnosis, treatment planning, and monitoring disease progression. Generate detailed reports and visualizations of the segmented structures and their volumes for clinical use. Continuously refine the AI algorithms and processes based on feedback and evolving medical knowledge to improve accuracy and efficiency.

CapsNet have been proposed to address the limitations of CNN for tumour segmentation. CapsNets require smaller datasets for training compared to CNNs and consider the surrounding tissues of the tumour [24].

These models can be effectively used in clinical practice to aid neuro-oncologists or radiologists in obtaining quick and accurate segmentation.

4.3. Molecular marker Identification

The revised WHO classification of CNS tumours uses molecular parameters such as IDH genotype and 1p/19q codeletion to classify gliomas [25]. According to this classification system, both low-grade astroctyomas and oligodendrogliomas are classified by the presence of IDH 1 and IDH 2 mutations as well as loss of portions of chromosomes 1 and 19, known as 1p19q codeletion. Having IDH1 or IDH2 mutations is associated with improved survival, as these gliomas respond better to Temozolomide therapy [26]. In a study by Beiko et al, in both WHO grade III and IV gliomas, resection of non-enhancing tumor after total resection of enhancing component correlated with improvements in survival as opposed to IDH mutant [27]. These findings show the importance of genetic information for patient monitoring and individualized therapies.[12].

Currently, genomic profiling is usually done on tissue samples from enhancing tumour components, which may not reflect the whole tumour heterogeneity. [12] Biopsy samples often have low tumour content [28] and genetic testing can be expensive, limited, and time-consuming. Non-invasive imaging techniques that can provide genetic information may overcome these limitations. (Table 5).

The process of obtaining genomic information from brain imaging led to the establishment of a new field: radio genomics. CNNs applied to conventional MRI modalities have been used to differentiate IDH mutant gliomas from IDH wild-type tumours with 92 % accuracy, consistent with prior visual assessment and underlying pathophysiology that IDH wild-type tumours demonstrate more infiltrative, ill-defined margins [29].

Other than a single manuscript reporting slight frontal lobe predilection, there are no consistent MR imaging features that can reliably and accurately predict 1p19q codeletion tumours [30]. Chang et al used CNN to predict 1p19q codeletion status achieving an accuracy of 94 % [29].

Approximately 33 %–57 % of diffuse gliomas exhibit hypermethylation of the promoter of the *MGMT* gene [31]. *MGMT* promoter hypermethylation has been associated with better prognosis owing to improved sensitivity to alkylating agents (eg, temozolomide) [31]. Radiomic studies have identified distinct imaging signatures for this molecular marker. Studies have been able to predict *MGMT* methylation status with up to 88 % accuracy by combining texture features with ML methods.

Akkus et al used a multi-scale CNN on post-contrast T1- and T2-weighted MR images to predict 1p19q codeletion with 93 % accuracy. They identified enhancement, infiltrative margins, and left frontal lobe as associated features [32].

Using radiomics analysis on preoperative MRI, Meng et al predicted ATRX status in 123 gliomas (grades II–IV) with high accuracy (0.79), sensitivity (0.73), and specificity (0.86). The AUC for ATRX mutation

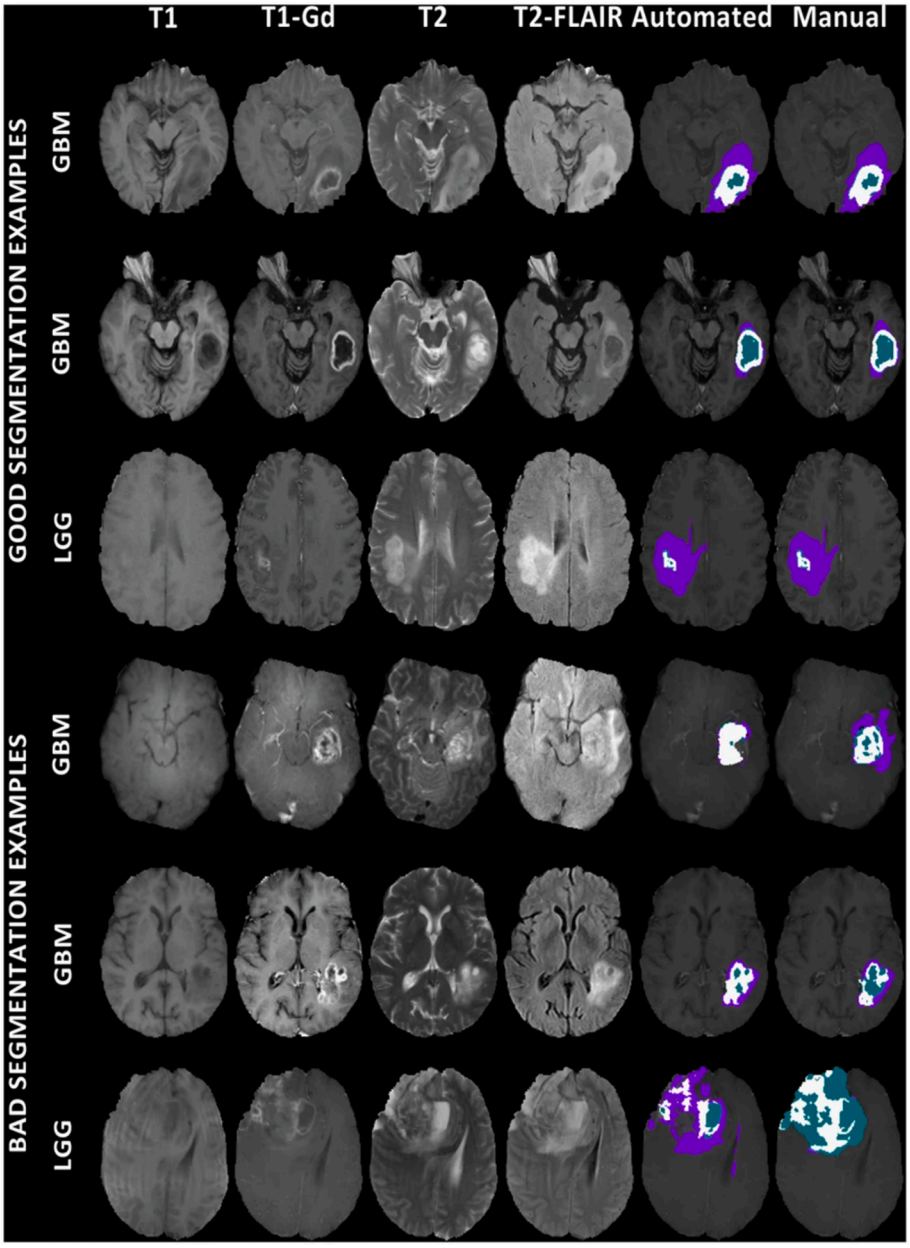

Fig. 3. Single slice multimodal MRI scans illustrating GLISTRboost segmentation examples.
Reproduced from Bakas, S. et al. Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features. Sci. Data 4:170117 (2017), licensed under CC BY 4.0. Good segmentations are represented in the first three row, while bad segmentations are represented in the last three rows.[13]

was 0.84 (95 % CI: 0.63–0.91) [33].

Habould et al developed a fully automated tool that uses radiomics to predict the molecular status and grade of brain gliomas, based on clinical and laboratory data. The tool performed very well in distinguishing low-grade from high-grade gliomas, as well as in detecting ATRX and IDH1/2 mutations (Fig. 4). However, it was less accurate in predicting 1p19q and MGMT status [34].

Thus, AI methods are useful in noninvasively distinguishing different central nervous system (CNS) malignancies prior to biopsy at levels comparable with trained neuroradiologists and can accurately classify genetic mutations of low and high-grade gliomas.

4.4. Prognostication

Basic imaging metrics, including maximal dimension and enhancing volume, have been used for prognostication. Additional studies have shown that in patients with GBM, both non-enhancing tumours and areas of infiltration are good predictors of overall survival (OS); poor OS is correlated with higher regional cerebral blood volume (rCBV) and EGFRvIII amplification (a marker of neo-angiogenesis) [35]. Various ML approaches, including SVM classifiers, have been utilized in grading and evaluating the prognosis of gliomas. In a study with 105 high-grade glioma patients, Macyszyn et al demonstrated that their SVM model could classify patients' survival into short or long-term categories with an accuracy range of 82–88 % [36]. The most predictive features in this model included tumour volumes, angiogenesis (enhancing tumour volume), peritumoural infiltration, cell density (trace diffusion values), and distance to the ventricles.

AI-assisted tools can help diagnose brain tumours, devise follow-ups and management plans, and predict post-operative complications [37].

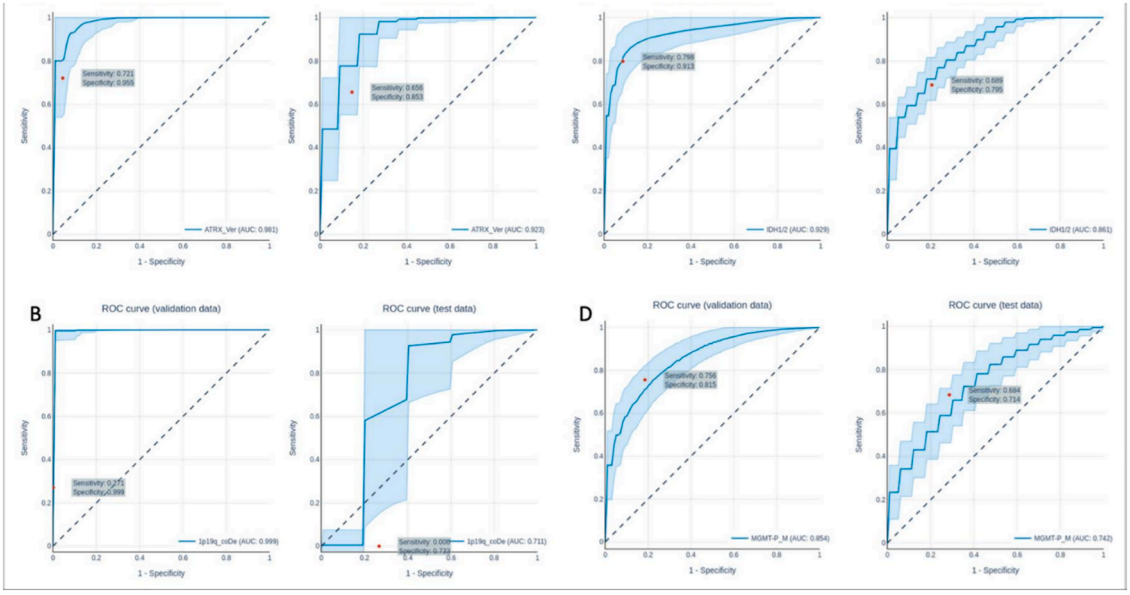

Fig. 4. ROC curves of the prediction of the ATRX expression loss. .
Reproduced from Haubold, J. et al. Fully Automated MR Based Virtual Biopsy of Cerebral Gliomas.[24] Cancers (Basel). 2021 Dec 8;13(24):6186. licensed under CC BY 4.0. (A), the 1p19q co-deletion (B), the IDH1/IDH2 mutation (C) and the MGMT-status (D) in the validation data set (left) and the test data set (right).[24]

A recent review by Williams et al discussed several studies on the potential of integrating AI to predict the development of common post-operative complications in brain tumor patients [37]. These complications include venous thromboembolism, falls, hypoglycemia, adverse drug events, and pressure ulcers [38,39].

Glioma patients also face the challenge of distinguishing tumor recurrence from radiation necrosis, which can affect their treatment options. Current MRI methods are not very reliable for this task, and artificial intelligence (AI) still struggles to capture tumor complexity.

Few studies have explored this problem so far. However, AI can help measure tumor response to treatment by analyzing tumor volume and predicting outcomes. AI can also help identify imaging features that are linked to tumor immune activity, which is important for evaluating immunotherapies in GBM. These features are derived from different types of MRI images and T-cell gene expression data. Using MRI images of different types (T1-weighted post-contrast and T2-FLAIR) and gene expression data of T-cell markers (CD3D/E/G) from GBM patients, Narang et al. found six imaging features that are related to the activity of CD3 T-cells within GBM tumours [40].

4.5. Pseudo-progression

Psuedo-progression is defined as new contrast enhancement that occurs due to radio-chemotherapy, and subsides independent of any alterations in treatment.[24] The differentiation of pseudo-progression from true progression remains a crucial diagnostic dilemma. Macdonald criteria, the first and most widely used criteria to assess treatment response in patients with high-grade glioma (HGG), considers all enhancing lesions as progression without consideration for treatment-related processes such as inflammation or necrosis. The Response Assessment in Neuro-Oncology (RANO) criteria recognizes true disease progression within 12 weeks post-treatment only with pathological confirmation or if new lesions have appeared outside the radiation field; otherwise, pseudo-progression is instead considered as a possible diagnosis.

A recent metanalysis showed that up to 36 % of cases of pseudo-progression were underdiagnosed using RANO criteria.

Studies have successfully incorporated ML algorithms to predict pseudo-progression [41]. Results from one study showed 89.91 % sensitivity and 93.72 % specificity of an optimized one-class SVM(OC-SVM) classifier for pseudo-progression, as shown in the Table 1.25.

Techniques such as DL and ML CNN-LSTM (long short-term memory), are being used on patients with GBM from different institutions, to differentiate PsP from TP [42].

However, the lack of “ground truth” or histopathologically proven cases, and an insufficient number of well-curated, annotated MR images of PsP cases, may explain the relative absence of CNN manuscripts devoted to prediction of PsP, which remains a critical, unmet need in neuro-oncology.

5. Ai applications in non-glioma evaluation

Non-glioma brain tumours are a diverse group of brain tumours that can originate from different cell types and locations within the central nervous system. They include metastatic tumours, meningiomas, pituitary tumours, ependymomas, medulloblastomas, hemangioblastomas, and others. These tumours have different histopathological features, clinical manifestations, prognoses, and treatment options. Therefore, accurate and reliable methods for brain tumour classification are essential for improving patient care and outcomes.

5.1. Metastasis

BM is the most common type of intracranial tumour in adults [43]. It occurs when cancer cells from a primary site spread to the brain through the bloodstream or lymphatic system. BM can originate from various primary cancers [43], such as lung cancer, breast cancer, melanoma,

Table 1
Results of Obtained Optimal Classifier One-Class Support Vector Machine (OC-SVM) for Pseudoprogression.

Characteristics	Maximum AUROC	Sensitivity (TP)	Specificity (FP)
Gamma = 5 ν = 0.06 Threshold = -0.12	0.9439	89.91 %	93.72 %

Note: AUROC = Area Under Receiver Operating Characteristic Curve; gamma = width of radius base function; ν = fraction of outliers in training sample.

renal cell carcinoma, or colorectal cancer.

AI techniques have been applied to BM differentiation from primary brain tumours. Using preoperative MRI data, Bae et al [44] developed a radiomic model to differentiate between GBM and brain metastasis using preoperative MRI data from 166 patients in the training cohort and 82 patients in the validation cohort. The model used 265 radiomic features from T2-weighted and contrast-enhanced T1-weighted MRI sequences, and various machine learning classifiers, such as adaptive boosting (AdaBoost), SVM, and linear discriminant analysis (LDA). The deep neural network model achieved an AUC of 0.956 in the validation cohort. Swinburne et al. [45] used SVM and Multi-Layer Perceptron (MLP) models to differentiate between intracranial GBM, Central Nervous System Lymphomas, and BM using preoperative brain MRI. The MLP model trained on fractional plasma volume (Vp) values from the non-enhancing T2 signal hyperintense region (NET2) surrounding the enhancing tumor component differentiated the three tumor classes with a moderate accuracy of 54 %.

AI techniques have also been applied on determining the primary tumour source for BM. Ortiz-Ramón et al [46] used 43 texture-based features and Random Forest (RF) classifier to differentiate between BM derived from lung cancer, breast cancer, and melanoma using MRI-based texture Analysis. The RF classifier demonstrated that lung cancer BM could be differentiated from breast cancer and melanoma BM with high accuracy (AUC $>$ 0.9) using a few features of the optimal dataset. The study also found that breast cancer and melanoma BM were poorly classified (AUC $=$ 0.607) using the same features, indicating that RF classifiers are not suitable for discriminating between breast and melanoma BM.

Detection and segmentation of even the smallest brain metastasis is crucial for treatment planning. Various AI techniques have been applied to this task [1747] (Fig. 5), most using contrast-enhanced three-dimensional (3D) gradient echo (GRE) imaging datasets. Considering that black blood imaging technique [48] can suppress blood vessel signals, enabling clearer delineation and better detection of small brain metastases [49], Won Park et al [18] used a U-net-based DL model to segment brain metastases on black-blood (BB) and gradient-echo (GRE) MRI sequences from 188 patients in the training set and 45 patients in the test set. The combined BB-GRE model achieved high sensitivity (93 %) and precision (85 %) for detecting brain metastases, especially for small lesions ($<$3 mm). The model also showed high Dice coefficient (0.822) for segmenting brain metastases.

The clinical presentation of brain metastases (BM) may reflect the primary tumour site. Previous studies have mainly relied on radiomic features extracted from conventional MRI images. To evaluate the added value of clinical features for the AI model, Han et al. [50] conducted a retrospective study of glioblastoma multiforme (GBM) and metastatic brain tumors (MET) from two institutions, with data collected from January 2010 to December 2017 for MET and from January 2014 to December 2015 for GBM. They extracted 841 radiomic features from MRI scans and performed feature selection methods. They then built four radiomics models using different algorithms (logistic regression, support vector machine, decision tree, and random forest). They also constructed clinical-radiological models that incorporated patient age, sex, tumor size, edema ratio, and location. Furthermore, they trained combination models that integrated clinical, radiological, and radiomic variables. They applied these three types of models to distinguish between GBM and MET, and between MET-lung and MET-other. The results indicated that the combination models achieved the highest performance in both classification tasks. The combination GBM model had an AUC of 0.774 in the external validation, while the clinical-radiologic model had an AUC of 0.674. The combination MET model had an AUC of 0.833 in the external validation, while the clinical-radiologic model had an AUC of 0.759. The study demonstrated that the combination of radiomic features and clinical-radiological factors can enhance the accuracy of differentiating between GBM and MET, and between MET-lung and MET-other.

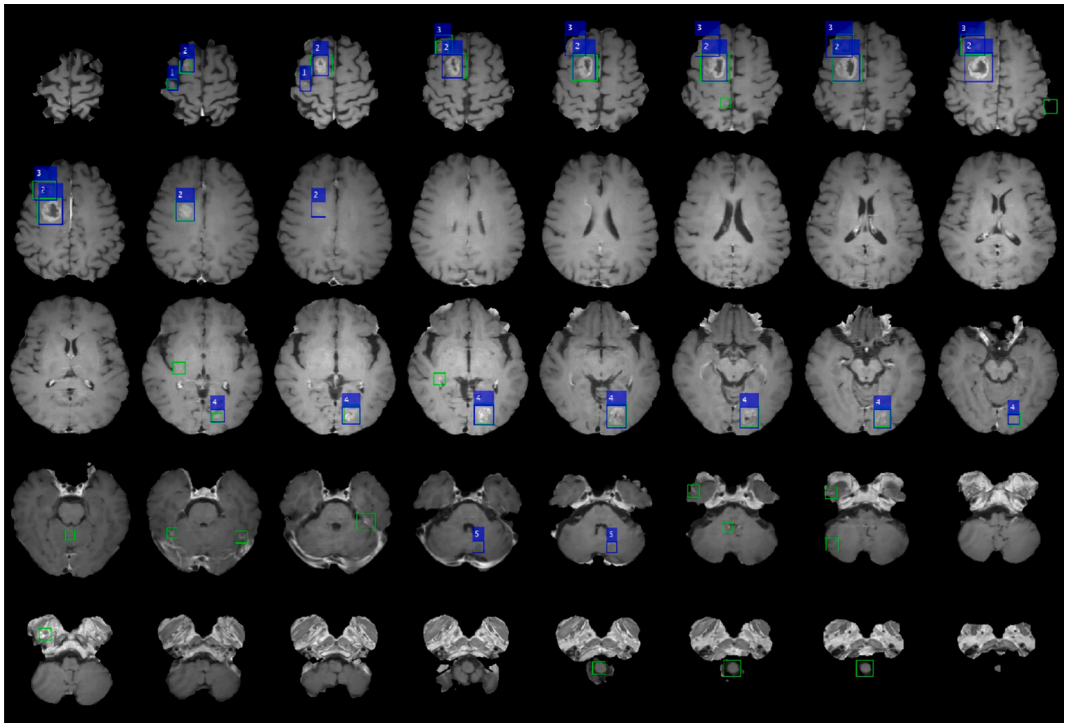

Fig. 5. Example of brain metastasis detection.)
Reproduced from Oh, JH., Lee, K.M., Kim, HG. et al. Deep learning-based detection algorithm for brain metastases on black blood imaging. Sci Rep 12, 19503 (2022). [37] licensed under CC BY 4.0. The numbers 1–5 are displayed at each lesion. Blue boxes show the label that the lesion placed more than two adjacent slices and green boxes show the prediction result by deep learning algorithm. This figure was generated by MATLAB (MathWorks, R2020b, Natick, MA, USA)[37]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.

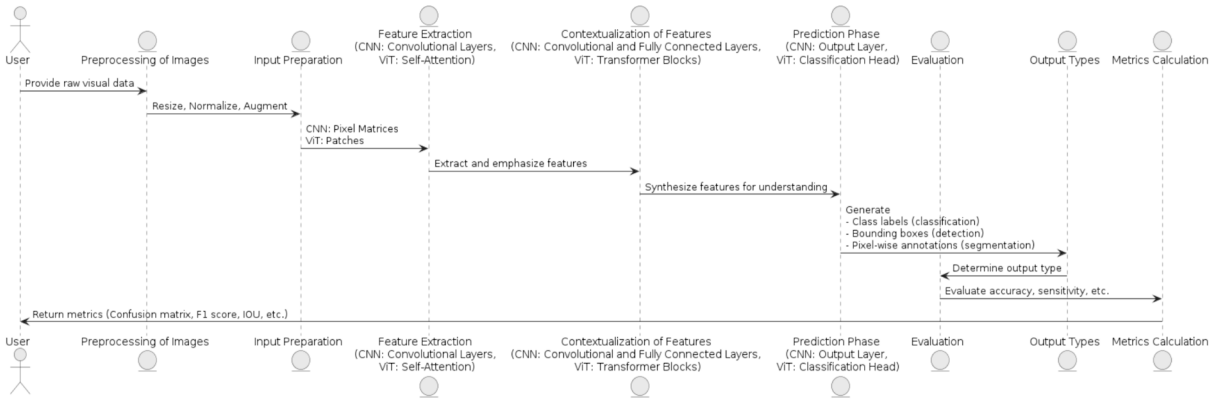

Fig. 6. Sequential Workflow of AI in Diagnostic Imaging. Step-by-step process of using artificial intelligence, specifically through Convolutional Neural Networks (CNN) and Vision Transformers (ViT), for the analysis of diagnostic images. It details the stages from preprocessing raw data to evaluating model performance with specific metrics for classification, detection, and segmentation tasks.

DNN is superior to ML and human readers for BM detection using MRI imaging. Cho et al [51] reviewed 16 studies that applied ML and DNN to brain metastasis detection using MR imaging. They found that DNN outperformed ML and human readers in most metrics, especially in reducing false positives ($P < 0.001$) while maintaining high sensitivity. Bae et al [51] compared various ML models and DNN with radiomic features from CE or PT regions of the brain tumours. The DNN achieved the highest performance with an AUC of 0.956, compared to AUC of 0.890 for ML and 0.904 for human experts.

5.2. Posterior fossa tumours

Posterior Fossa Tumours (PFT) are tumours in the posterior fossa region of the brain, which has the cerebellum, brainstem, and fourth ventricle. PFTs account for about 15 % of all intracranial tumours in adults [52] and about 60 % of all intracranial tumours in children [53]. PFTs have various types, such as pilocytic astrocytoma, medulloblastoma, ependymoma, ATRT, and hemangioblastoma. They differ in biology, symptoms, prognosis, and treatment.

As the treatment modalities are different for the various PFTs, differentiating between these tumour types are important. Payabvash et al. [54] used ADC histogram metrics and CART decision tree models and random forest machine learning algorithms to differentiate between different PFT types, such as pilocytic astrocytoma, medulloblastoma, ependymoma, ATRT, and hemangioblastoma. The study included 200 patients who underwent surgical pathology and MRI with ADC map, T2-weighted, FLAIR, and post contrast T1-weighted sequences. The study used classification and regression tree (CART) decision tree models and random forest machine learning algorithms to differentiate between tumor types. The CART model achieved accurate classification rates ranging from 30 % to 90 % for 7 histopathologies, while the random forest models showed high performance in multiclass differentiation, with an averaged AUC of 0.961 in training datasets and 0.873 in validation datasets. This study demonstrated the potential of Machine learning algorithms for PFTs. Another study by Quon et al. [55] developed a deep learning model to detect and classify four types of pediatric posterior fossa tumors: diffuse midline glioma of the pons, medulloblastoma, pilocytic astrocytoma, and ependymoma, using T2-weighted MRIs. The model was based on the modified ResNeXt-50-32x4d architecture and used transfer learning and data augmentation techniques to improve its performance. The results showed 0.99 AUROC value for tumor detection and 0.92 accuracy and 0.80 F1 score for tumor classification on a multi-institutional dataset of 739 scans. The model outperformed two of the four radiologists in tumor classification and was comparable with the other two. A systematic review of ML algorithms developed to classify and diagnose pediatric PFTs by Yearley et al. [56] showed that the algorithms exhibited variable performance based on sample size, classifier(s) used, and the individual tumor types being investigated. The median reported AUC was 0.87, and the most popular algorithms were support vector machine and random forest.

Some AI techniques have been used for PFT treatment. Post-operative hydrocephalus is a well-known complication after resection of PFT and can lead to long-term CSF diversion and prolonged hospitalization. Bray et al [57] used an ANN model with 17 variables to predict hydrocephalus after PFT surgery needing permanent CSF diversion. The model has high accuracy (0.902 AUC), calibration, and usefulness in both validation cohorts, and beats other classifiers. The model can find high-risk patients and reduce hospital time and cost.

A systematic review of ML algorithms developed to classify and diagnose pediatric PFTs by Yearley et al [56] showed that the algorithms exhibited variable performance based on sample size, classifier(s) used, and the individual tumour types being investigated. The median reported AUC was 0.87, and the most popular algorithms were support vector machine and random forest.

5.3. Pituitary tumours

Pituitary tumours are tumours that arise from the pituitary gland, a small endocrine organ located at the base of the brain that secretes various hormones that regulate various bodily functions.

Some patients with pituitary adenoma, especially those with acromegaly [58], would benefit from surgery [59]. Tumour consistency is one of the factors that affects the surgical treatment of pituitary tumours [59]. Pre-operative knowledge of tumour consistency is relevant to surgical planning, risks and outcome, as fibrous tumours tend to be larger in size and invade neighbouring structure and cavernous sinus [59]. Fan et al. [60] developed a radiomics model to predict the tumor consistency of pituitary macroadenomas in patients with acromegaly using conventional MRI sequences. The model combines a radiomics signature derived from four selected features and clinical characteristics such as Knosp grade. The model shows high accuracy (0.88 in internal validation and 0.86 in external validation), calibration, and clinical usefulness in both internal and external validation cohorts, and outperforms the use of clinical features alone (0.72 in internal validation and 0.68 in external validation). Also, Zeynalova et al. [61] evaluated the potential value of machine learning (ML)-based histogram analysis on conventional T2-weighted MRI for predicting consistency of pituitary macroadenomas (PMA) and compared it with that of signal intensity ratio (SIR) evaluation. The model combines six selected texture features and surgical and histopathological findings as reference standard. The model shows high accuracy (0.71 AUC value), calibration, and clinical usefulness in a 10-fold cross-validation protocol and outperforms the use of SIR evaluation (0.55 AUC value). These models can help to preoperatively evaluate tumour consistency in PMA and plan proper surgical

strategy.

AI techniques has also been used to predict treatment response and cavernous sinus invasion in pituitary adenomas. This is important because invasive pituitary adenomas are hard to treat, hence individualized treatment plan would benefit these patients the most. Fang et al [62] used a pre-trained Resnet50 CNN model with transfer learning and data augmentation to diagnose cavernous sinus invasion (CSI) in pituitary adenoma (PA) using T1-enhanced MRI. The model had 0.98 AUC-ROC, 0.95 sensitivity and 0.96 specificity for CSI prediction, and beat the Knosp grading system, which had 0.84 AUC-ROC, 0.96 sensitivity and 0.46 specificity. Niu et al [63] used a radiomics method with contrast-enhanced T1 (CE-T1) and T2 MR images to predict CS invasion by PA. The method extracts imaging features and uses an SVM model with a radiomics signature and clinical characteristics. The model had 0.899 AUC in the training set and 0.871 AUC in the test set for the radiomics nomogram and beat the clinico-radiological model and the single-image models. These models can help to preoperatively evaluate tumour invasiveness and plan proper surgical strategy.

AI techniques have also been used to differentiate pituitary tumours from other Primary CNS tumours. Ucuzal et al [64] and Haq et al [65] developed DL models to classify brain tumours into three types: glioma, meningioma, and pituitary, from MRI images. Ucuzal et al [64] used Auto-Keras to train and test on a dataset of 3064 MR images from 233 patients. Their web-based software reported 98 % and above performance on various metrics for classifying the tumour types. The software is publicly accessible at [https://biostatapps.inonu.edu.tr/BTSY/]. Haq et al [65] used the ResNet-50 architecture and transfer learning and data augmentation techniques to improve their model. Their results reported 99.89 % accuracy on augmented brain tumour MR images data set, and out-performed other CNN models such as VGG-16, Inception V3, DenseNet201, Xception, and MobileNet.

Although several studies conducted on ML and treatment response for pituitary tumour reported promising results, [66,67,68,69,70] Rech et al [66] systematically reviewed 20 studies on ML models for pituitary surgery outcomes, such as endocrine remission, tumour recurrence, and complications. They used the TRIPOD statement to evaluate the reporting quality. The results showed that the studies had low adherence to TRIPOD criteria (median 65 %) and high performance (median AUC 0.84). The review concluded that ML applications in pituitary surgery are still nascent and need more validation and transparency.

6. Transformer-based neural networks in brain tumour imaging

Transformers, a groundbreaking architecture initially revolutionizing natural language processing, as seen in models like ChatGPT, have expanded well beyond their original purpose. Their remarkable success in understanding and generating human language has paved the way for their application in medicine, specifically neuro-oncology imaging

Table 2
Recent Studies on the transformer-based neural network in brain tumour imaging.

Author and Year	Sample/Data Size	Imaging Modality/ Sequences and Clinical Data	AI Model	Task	Main Results	Limitations
Baur et al., 2023	834 training, 208 validation, 209 test samples	T1, T1Gd, T2, T2FLAIR MRI	CKD-TransBTS	Brain tumor segmentation	Sensitivity: 0.869–0.878, Specificity: 0.921–0.935	Under-segmentation of necrosis region in some cases
Fink et al., 2022	10,455 patients	CT, MRI, US exams	BERT (NLP model)	Assessment of oncologic outcomes	BERT F1 score: 0.70 (95 % CI: 0.68–0.73)	Not as effective as radiologists in curating outcomes, similar performance to medical students.
Anaya-Isaza et al., 2023	233 subjects (BTD dataset)	T1-Gd	Various (InceptionResNetV2, DenseNet121, etc.)	Tumor classification	F1 scores: 68.50 % – 95.39 %	InceptionResNetV2 had significantly lower accuracy in meningioma detection. Variation in the data sets makes it difficult to determine detection accuracy and only 2D images databases were used.
Anaya-Isaza et al., 2023	253 subjects (MRI-D dataset)	T1WI	Various (InceptionResNetV2, DenseNet121, etc.)	Tumor detection	F1 scores improved by 3.4 % from scratch, 1 % improvement for data augmentation	Variation in the data sets makes it difficult to determine detection accuracy and only 2D images databases were used.
Anaya-Isaza et al., 2023	110 subjects (TCGA-LGG dataset)	T1WI, T1-Gd, FLAIR sequences	Various (InceptionResNetV2, Cross-Transformer, etc.)	Tumor detection	All networks achieved scores over 90 %	Variation in the data sets makes it difficult to determine detection accuracy and only 2D images databases were used.
Wu et al., 2022	493 glioma patients	T2-weighted MRI, clinical data (gender, age, grade)	Swin Transformer and ResNet	Predicting IDH mutation status in glioma tumours	Swin Transformer outperformed ResNet; best results achieved with 1.5 × Tumor Bbox input for Swin Transformer	Only T2 images were used. Only one external dataset was used, potential generalization and interpretability issues.
Lyu et al., 2022	1582 MRI images from 1,399 patients	Contrast-enhanced T1-weighted MRI, fast spoiled gradient echo brain MRI exams	Deep learning	Identifying primary organ site of metastatic brain tumours and	Deep learning allowed for accurate diagnosis of primary organ site of brain metastases with AUC of 0.878	Models had limited data storage and dataset lacked diversity for model training. Limited imaged of the primary cancer site used, lack of diversity in tumour classes and histological subtypes, model unable to diagnose patients with cancer of unknown primary origin.
Pinaya et al. 2022	15,000 radiologically normal images	FLAIR images, brain MR datasets with small vessel disease, demyelinating lesions, and tumors	Vector Quantized Variational Autoencoder (VQ-VAE) with autoregressive transformers	Unsupervised anomaly detection and segmentation in brain imaging	Superior anomaly detection on 2D and 3D data; effective in segmenting real-world lesions in different datasets	Different resolution images may provide inconsistent results in transformer performance.

Table 3
Studies on Lesion Detection.

Author and Year	Dataset	AI Model	Tumour type	Main Findings	Limitations
Blanc-Durand et al., 2018	F-FET PET	3D U-Net CNN	Glioma	Achieved 100 % sensitivity and specificity in lesion detection with no false positives reported across 11 validation cases.	Limited dataset size of 37 patients may impact generalizability; potential overfitting due to small sample size despite data augmentation.
Park et al., 2021	MRI	3D U-Net	Brain Metastases	The combined 3D black-blood and 3D gradient echo model achieved 93.1 % sensitivity in detecting brain metastases.	Limited by a single-center, retrospective study design, and small dataset size which may impact generalizability.
Chen et al., 2023	MRI	Two-level Histogram-based Morphometry (HBM), SVM	Gliomas	Achieved AUC of 0.921 for glioma detection using structural MRI data.	Limited dataset size (99 for detection, 134 for grading) may affect generalizability.
Gokila Brindha et al., 2021	MRI	ANN and CNN	Brain Tumor	The CNN model achieved a testing accuracy of 89 % and a validation accuracy of 94 %.	The paper does not discuss the performance on different tumor types or stages, which might affect its applicability to varied clinical cases.
Isselmou Abd El Kader et al., 2021	MRI	Deep Wavelet Auto-Encoder	Brain Tumor	The model achieved an accuracy of 99.3 % in detecting brain tumors, demonstrating high efficiency.	The study is limited by lack of validation on an independent external dataset.
Yakub Bhanothu et al., 2020	MRI	Faster R-CNN	Glioma, Meningioma, Pituitary	The model achieved mean average precision (mAP) of 77.60 % for tumor detection.	Limited dataset size and variation in tumor appearance might affect the model's generalizability.

(Table 2), offering promising advancements in brain tumor diagnosis, classification and treatment. Transformers' ability to efficiently process and analyze complex data makes them a valuable tool in medical imaging, especially for the diagnosis of brain tumours.

6.1. Transformer efficiency in MRI-Based brain tumour segmentation

The efficient segmentation of brain tumours using MRI is crucial for enhancing diagnosis and treatment. Transformer-based algorithms have emerged as a powerful tool to significantly improving both the accuracy and efficiency of diagnosis over traditional methods. Lan et al [71], 2023 highlight the potential of transformer-based algorithms in neuro-oncology, such as SPP-U-Net and RMTF-Net for MRI-based brain tumour segmentation. The Swin Transformer model showed promise in predicting molecular expressions of gliomas, particularly that of IDH mutation status. It surpassed conventional CNN-based methods, achieving an overall area under the receiver operating characteristic curve (AUC) of 0.878 (95 % CI: 0.873–0.883).

Anaya-Isaza et al [72], 2023 conducted a comparative analysis of various AI techniques for brain tumour classification and detection. They demonstrated that transfer learning and data augmentation improved accuracy up to 6 % and proposed a model that was significantly more timely efficient compared to other models.

Furthermore, Pinaya et al [73], 2022 developed a method for unsupervised anomaly detection and segmentation in brain imaging using transformers, where various anomalies were assessed and models were able to detect both image-wise and pixel/voxel-wise anomalies. Models achieved an impressive area under the receiver operating characteristic curve (AUROC) of 1.000 against far out-of-distribution (OOD) classes and 0.921 against near OOD classes in synthetic datasets, showcasing superior performance compared to other models.

6.2. Clinical Knowledge-Driven hybrid transformer for enhanced segmentation

Integrating clinical knowledge into AI models for brain tumour segmentation enhances precision medicine and treatment planning. Lin et al [74], 2023 developed the CKD-TransBTS model, a clinical knowledge-driven hybrid transformer, that demonstrated significant advancements in brain tumour segmentation. The model achieved superior Dice scores for enhanced tumour (ET), tumour core (TC), and whole tumour (WT), with the HD95 metric for ET notably lower at 5.93 mm compared to the next best model at 9.01 mm. Incorporation of clinical knowledge into this model proved a significant improvement in precise and efficient segmentation of brain tumours using MRI.

6.3. Leveraging deep NLP for structured oncology reports

Deep NLP models like BERT can efficiently extract and analyze complex medical data from structured oncology reports, aiding in accurate tumour response classification. Fink et al [75], 2022 developed a deep NLP model using bidirectional encoder representations from transformers (BERT) to classify tumour response categories based on structured oncology reports. Analyzing data from 10,455 patients, the BERT model achieved a notable F1 score of 0.70 (95 % CI: 0.68–0.73) on 802 free-text oncology reports, outperforming the reference linear support vector classifier (F1 0.63; 95 % CI: 0.61–0.66) and performing similarly to medical students but not as well as radiologists (F1, 0.73; 95 % CI: 0.72, 0.75).

6.4. Predicting molecular characteristics without refined tumour segmentation

Predicting tumours' molecular characteristics without detailed segmentation is crucial for personalized cancer treatment. Wu et al [76], 2022 aimed to predict IDH mutation status from MRI scans without refined tumor segmentation. The Swin Transformer models they developed achieved an average AUC of 0.965 and an ACC of 92.3 % in the internal test, and 0.842 AUC with 76.6 % ACC in the external test, outperforming the ResNet models. Best results were achieved when using a 1.0 × tumor bounding box input strategy. Highlighting the potential of transformer models in medical imaging for predicting complex molecular characteristics like IDH mutation status.

6.5. Non-Invasive classification of brain metastases

Non-invasively classifying brain metastases into primary organ sites is vital for determining effective treatment strategies. Lyu et al [77], 2022 developed a deep-learning approach to classify brain metastases into primary organ sites using whole-brain MRI data. The model achieved an overall AUC of 0.878 (95 % CI: 0.873–0.883) in tenfold cross-validation, indicating high accuracy in diagnosing brain metastases into categories like lung, breast, and melanoma. Binary classification experiments showed even higher accuracy, particularly for lung and breast categories with an AUC of 0.959.

Although transformers demonstrate immense promise in neuro-oncology there are some challenges hindering their integration into clinical practice. Limited availability of well-annotated, diverse datasets for training impacts the models' generalization abilities [72]. The complexity of transformers also introduces interpretability issues which is essential for gaining the models trust in its ability to accurately diagnose and classify brain tumours [77]. Integrating transformers into

Table 4
Studies on Brain tumour Segmentation.

Author and Year	Dataset	AI Model	Tumour type	Main Findings	Limitations
Roelant S. Eijgelaar, et al., 2020	BraTS dataset plus a clinical dataset from six hospitals during 2012–2013, including 751 patients in total with glioblastoma.	DeepMedic	Glioblastoma	– A model trained only on the BraTS dataset had a median Dice score of 0.81 but performed poorly on clinical data (0.49). – Sparsified training improved segmentation performance on both complete and incomplete datasets, achieving Dice scores over 0.8 with site-specific data, comparable to models trained on all data.	– Variability in MRI acquisition protocols and the quality of clinical images affected the model's performance.
Dongwei Liu, et al., 2022	Brain Tumor Segmentation (BraTS) Challenge 2020 and 2021	SGEResU-Net	Glioma	Achieved DSC values of 83.31 %, 91.64 %, and 86.85 % for enhancing tumor, whole tumor, and tumor core, respectively. Hausdorff distances (95 %) of 19.278, 5.945, and 7.567.	Specific limitations not detailed in the summary.
Yuqi Han, Lingling Zhang, et al., 2021	350 patients from two institutions	Fisher decision tree, reliefF random forest	Glioblastoma Multiforme and Metastasis	– Fisher_DT and reliefF_RF showed best performances for GBM vs MET and MET-lung vs MET-other.	Specific limitations not detailed; focus on radiological without much clinical integration.
Yae Won Park, Yohan Jun, et al., 2021	188 patients with brain metastases for training; test set of 45 patients with and 49 without metastases.	3D U-net using 3D BB and 3D GRE images	Brain metastases	The combined 3D BB and 3D GRE model significantly improved segmentation performance with a Dice coefficient of 0.822.	Single-center, retrospective study, which may limit generalizability to broader clinical settings.
Shaocheng Wu, Hongyang Li, et al., 2020	2018 Multimodal Brain Tumor Segmentation Challenge (BraTS) dataset	Two-dimensional U-Net models	Gliomas	Achieved mean Sørensen–Dice scores of 0.80, 0.84, and 0.91 for enhancing tumor, tumor core, and whole tumor respectively.	Single modality (MRI) and the retrospective nature of the study.
Spyridon Bakas, et al., 2017	TCGA glioma MRI collections (243 pre-operative scans)	GLISTRboost	Gliomas	Dice Similarity Coefficient: 0.86; Sensitivity: 85 %; Specificity: 99 %; Hausdorff Distance: 3.2 mm. The method integrates a generative-discriminative model improving segmentation accuracy for glioma sub-regions.	Manual corrections were necessary for accuracy.
Agus Subhan Akbar et al. 2022	BraTS 2018, 2019, 2020, 2021	Single Level UNet3D with MRAB	Brain tumor	Dice scores: ET 77.71 %, TC 79.77 %, WT 89.59 % for BraTS 2018; improvements in segmentation accuracy and performance	Not specified in the provided tex
Pranjal Agrawal et al. 2022	BraTS 2020	3D-UNet	Brain tumor	Achieved segmentation accuracy with a Dice coefficient of 1.28, sensitivity of 0.97, specificity of 0.99, and precision of 0.88.	High computational cost, requires high-end GPUs, and challenges in processing 3D data.
Yan Hu & Yong Xia. 2017	BraTS 2017	3D deep neural network, U-Net	Brain tumor	Achieved Dice similarity coefficients of 0.81 for whole tumor, 0.69 for core tumor, and 0.55 for enhancing tumor	High computational complexity
Ramin Ranjbarzadeh et al. 2021	BRATS 2018	Cascade CNN with Distance-Wise Attention	Brain tumor	Dice scores: whole tumor 0.9203, enhancing tumor 0.9113, tumor core 0.8726; high accuracy and reduced computational cost	Requires high-end GPUs for efficient training; complexity in handling 3D data
Sindhu Devunooru et al. 2019	No specific dataset mentioned	VoxResNet and various CNN models	Brain tumor	Utilized various deep learning models with features like 3D residual networks, enhancing image contrast and accuracy, though specific metrics are not detailed	Limited testing on large datasets, high computational requirements, and need for extensive validation
T. Ruba et al. 2022	BRATS 2015	3D U-Net, TLN, LSIS based ITSN	Brain tumor	Achieved Dice scores of 0.9064 for the complete tumor, 0.8425 for the core, and 0.8153 for the enhanced regions using the LSIS operator for feature extraction and a cascaded CNN architecture.	High computational cost, limited to high-grade gliomas, may not generalize well outside BRATS dataset.
Zeeshan Shaukat et al. 2022	BRATS Dataset	Cloud-based 3D U-Net	Brain tumor	Achieved an average Dice score of 95 %, demonstrating high accuracy in glioma segmentation.	Requires high-end GPUs and cloud infrastructure for efficient training and accessibility.

existing clinical workflows, ensuring robustness across diverse data sources, and meeting real-time processing demands as well as ensuring data security, patient-privacy, predications bias further complicate their implementation. The use of only some imaging modalities to test the efficacy of various transformer models further calls into question the model's generalizability and adaptability [77]. Overcoming these challenges is imperative for the seamless integration of transformers into clinical practice, ultimately enhancing neuro-oncology diagnostics and treatment strategies.

6.6. Transformers perform better than CNN on corrupted images but not on clean images

The performance metrics of ViTs and CNNs reveal distinct advantages for each architecture. According to Oh et al, ViTs demonstrate superior robustness in handling corrupted medical images compared to CNNs. For example, ViTs maintain a higher average AUROC under various levels of image corruption, registering minimal performance drop from 0.878 in clean conditions to 0.811 in corrupted settings. In contrast, CNNs experienced a more significant drop from 0.872 to 0.794 when faced with corrupted images [78].

This robustness of ViTs is crucial in scenarios where medical images

Table 5
Studies on Brain Tumour Molecular Marker Identification.

Author and Year	Dataset	AI Model	Tumour type	Main Findings	Limitations
Banerjee et al., 2020	TCGA-GBM, TCGA-LGG, MICCAI BraTS 2017	ConvNets (PatchNet, SliceNet, VolumeNet)	Gliomas	VolumeNet achieved the highest LOPO test accuracy of 97.19 % and holdout test accuracy of 95 % for LGG/HGG classification.	Limited by training on specific datasets, which may not generalize across broader clinical populations.
Xiong et al., 2016	84 patients with oligodendroglial tumours	DTI and cMRI	Oligodendroglial tumours	Conventional MRI and DTI values correlated with IDH1/2 mutations, showing that tumours with mutations often had higher minimal ADC and lower maximal FA values, indicative of lower cell density. DTI values were significant for IDH mutation assessment (maximal FA: $P = 0.009$, minimal ADC: $P = 0.001$), but not for 1p/19q genotyping.	The study was retrospective and limited to a single institution. The small sample size and lack of prospective validation might limit generalizability.
Narang et al., 2017	79 GB patients from TCGA, 69 from MD Anderson	Predictive model based on MRI texture features	Glioblastoma	Model achieved an accuracy of 97.1 % (AUC of 0.993) in the training set and 76.5 % (AUC of 0.847) in the test set. This indicates a significant relationship between MRI-derived textural features and CD3 T-cell infiltration.	Limited by the retrospective nature and potential variation in imaging protocols across the dataset.
Meng et al., 2022	123 patients with gliomas, WHO grades II–IV	SVM with LASSO	Gliomas	Achieved AUCs of 0.93 (training) and 0.84 (validation) for predicting ATRX status. The model showed sensitivity of 91 %, specificity of 82 %, and accuracy of 88 % in the training set, and sensitivity of 73 %, specificity of 86 %, and accuracy of 79 % in the validation set.	Retrospective design, single-center data, modest sample size. Potential generalizability issues.
Karami et al., 2023	146 adult-type gliomas	ResNet with multi-shell dMRI and cMRI	Adult-type gliomas	Achieved an accuracy of 81 % ± 5 % for IDH mutation status prediction and 60 % ± 5 % for predicting three molecular subtypes. The combination of cMRI and dMRI inputs showed the best performance compared to each modality used alone.	Limited by retrospective single-center data, modest sample size, and the potential lack of generalizability to other settings or MRI protocols.
Haubold et al., 2021	217 patients with cerebral gliomas	DeepMedic (CNN)	Cerebral gliomas	Achieved AUCs of 0.981/0.885 for differentiating low-grade from high-grade gliomas. Best results for ATRX expression loss prediction with AUCs of 0.979/0.923.	The study utilized a retrospective design and a single MRI protocol, which may limit generalizability. Fully automated segmentation may still require validation in diverse clinical settings.
Akkus et al., 2017	159 LGG patients, 477 image slices	Multi-scale CNN	Low-grade gliomas	Achieved sensitivity of 93.3 %, specificity of 82.22 %, and accuracy of 87.7 % in predicting 1p/19q status from T1C and T2 weighted MR images.	Limited sample size, and the study's external validity might be restricted due to the uniform scanning protocol used. Overfitting addressed only through data augmentation.
Hollon et al., 2023	Multicenter, International (153 patients)	DeepGlioma (CNN with Transformer architecture and genetic embedding)	Diffuse gliomas	Achieved mean molecular classification accuracy of 93.3 % ± 1.6 %. F1 scores were 96.3 % for IDH, 96.6 % for 1p19q co-deletion, and 94.7 % for ATRX.	Limited external testing cohort mainly in the United States and Europe, may not generalize globally. Model interpretability is a challenge.
Gopal S. Tandel et al., 2020	REMBRANDT	Convolutional Neural Network (CNN) with AlexNet transfer learning	Multiclass brain tumours (including Astrocytoma, Oligodendroglioma, GBM)	The CNN-based deep learning model showed superior performance compared to traditional ML models across five multiclass tumour datasets. Achieved mean accuracies of 100 %, 95.97 %, 96.65 %, 87.14 %, and 93.74 % with mean AUCs of 0.99.	The study relied on predefined datasets and primarily utilized T2-weighted MRI images, which may not generalize across diverse clinical settings or imaging conditions.
Beig et al., 2020	TCIA, Ivy-GAP, Cleveland Clinic	LASSO Cox Regression	Glioblastoma	The radiomic risk score (RRS) from Gd-T1w MRI predicted progression-free survival with a concordance index of 0.81 on training and 0.84 on the test set. Radiogenomic analysis linked RRS features to biological processes like cell differentiation and angiogenesis.	Limited radiogenomic analysis to one cohort due to data availability. Potential batch effects in RNA-sequencing data.
Ortiz-Ramón	67 untreated brain metastases from 38 cancer patients	Random Forest	Brain Metastases	3D texture features quantized with 32 Gy-levels had the highest AUC of 0.873 for classifying metastases	Limited dataset size, restricted to three primary tumor types, and the

(*continued on next page*)

Table 5 (*continued*)

Author and Year	Dataset	AI Model	Tumour type	Main Findings	Limitations
et al., 2018				from lung, breast, and melanoma origins.	use of a single imaging modality and machine.
Lao et al., 2017	75 patients from TCGA, 37 patients from Sun Yat-Sen University Cancer Center	Deep Learning (CNN with transfer learning)	Glioblastoma Multiforme	The deep learning-based radiomics model achieved a C-index of 0.710 in validation. Six deep features were crucial for survival prediction.	Limited data set, retrospective study, reliance on transfer learning.
Han et al., 2021	Retrospective dataset	Logistic Regression, Support Vector Machine, Decision Tree, Random Forest	GBM and MET	Best classifiers: fisher_DT (GBM vs MET, AUC: 0.696) and reliefF_RF (MET-lung vs MET-other, AUC: 0.759). Combination models outperformed clinical models.	Study retrospective, variable MET origins. Small case numbers for MET subgroups from different primary origins.
Chang et al., 2018	The Cancer Imaging Archives	Convolutional Neural Networks (CNN)	Gliomas	The CNN achieved high accuracy for classifying genetic mutations: IDH1 mutation status at 94 %, 1p/19q codeletion at 92 %, and MGMT promotor methylation status at 83 %. Distinct imaging features were identified for each mutation.	The study is limited by its retrospective design and the use of a relatively small, heterogeneous dataset from multiple sites. Generalization to unseen datasets remains untested.

may suffer from various artifacts or quality issues, such as in the complex imaging environment of neuro-oncology, where precise anomaly detection is critical despite potential visual obstructions or distortions. Conversely, CNNs continue to excel in tasks where the integrity of local features and textures is maintained, highlighting their utility in extracting detailed textural information critical for specific diagnostic processes. Although this study focused primarily on chest and skin images, extending this analysis to brain tumor imaging could provide deeper insights into the comparative effectiveness of these architectures in neuro-oncology, potentially guiding more tailored model selection and hybrid approaches in clinical practice.

7. Challenges

Implementing AI in radiology presents several challenges, including the need for high-quality, ground-truth data, seamless integration into existing user workflows, and the development of methods that are generalizable, interpretable, and robust across different settings and population groups [79].

Data Quality and Diversity: Large, well-annotated, and diverse datasets are essential to minimize measurement errors [80] and enhance algorithm performance across various sites, parameters, and populations. However, assembling such datasets is time-consuming and expensive. Collaborative data sharing and harmonization efforts, along with the use of standardized imaging protocols and synthetic data generation through advanced deep learning techniques like generative adversarial networks, are crucial for improving data quality and diversity.

Integration into Clinical Workflows: Effective integration of AI into radiological workflows is crucial for enhancing diagnostic accuracy. This involves designing user-friendly AI tools that complement the radiologist's expertise and facilitate a collaborative approach to diagnosis. Ensuring that AI tools are intuitive and add value to the radiologist's daily tasks without adding undue complexity is a significant challenge.

Explainability and Trust: AI systems must not only perform well but also be interpretable by their users. Radiologists need to understand how AI tools arrive at their conclusions to trust and effectively use these tools in clinical practice. Developing AI models that provide transparent, understandable outputs is essential for their acceptance and effective use.

Benchmarking and Standardization: There is a lack of clear, targeted "use cases" or tasks for benchmarking AI algorithms in neuro-oncologic imaging and radiomics, making it difficult to assess and compare performance consistently. The American College of Radiology Data Science Institute aims to address this by providing standardization and benchmarking tools and datasets.

Robustness and Generalizability: AI tools must be robust to changes in imaging settings, equipment, and population demographics to maintain their accuracy and reliability. Developing algorithms that can adapt to these variations without a loss in performance is challenging but necessary for widespread clinical adoption.

Education and Adoption: As AI tools become more integrated into clinical practice, it is imperative for radiologists to become proficient in their use. This involves not only training on the technical aspects of AI but also understanding its limitations, ethical considerations, and potential impacts on patient care.

8. Conclusion

This review discusses the use of AI in brain tumour imaging. The development of CAD tools can improve diagnostic accuracy in detecting small metastatic brain lesions, allowing for early and accurate treatment planning, particularly for stereotactic radiosurgery.

AI-driven extraction of imaging features that are not visible to the human eye is transforming radiological image analysis and reporting from a qualitative interpretation to an objective, quantifiable, and reproducible task.

Segmentation is crucial for surgery or radiation therapy planning, lesion monitoring, and even the development of radiomics-based tools. However, manual segmentation is time-consuming, so researchers have developed semi-automated or fully automated AI-based tools to assist radiologists in their daily practice. These tools provide objective measurements of tumor burden and growth patterns. Differential diagnosis of primary brain neoplasms can be difficult, particularly for PCNSL and HGG. Non-invasive AI-based techniques for accurate diagnosis can revolutionize the approach to brain disorders, avoiding invasive biopsies and allowing for the most appropriate treatment to begin.

The "virtual biopsy" is showing promising results in differential diagnosis and non-invasive characterization of tumor histotypes, allowing for increasingly personalized therapeutic plans. The better a lesion is characterized, the better the chances clinicians have of identifying effective therapies and predicting complications, recurrences, and progression.

Transformer models in neuro-oncology have demonstrated immense potential. From enhancing the accuracy and efficiency of MRI-based brain tumour segmentation to predicting complex molecular characteristics without detailed segmentation, these models are reshaping the field of neuro-oncology. The use of deep NLP models such as BERT for structured oncology reports and non-invasive classification of brain metastases further solidifies the significance of Transformers in neuro-oncology. Although CNN models are still the best performing, especially on clean image datasets, as transformer-based models continue to

evolve, their ability to process complex data holds great promise.

All these AI applications aim to achieve personalized medicine, improved patient outcomes, and increased survival. The future development and widespread adoption of these tools will benefit clinicians and patients alike, resulting in a personalized medical approach.

9. Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used ChatGPT in order to Analyze reviewed articles and to retrieve results findings. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

CRediT authorship contribution statement

Ezekiel Chukwujindu: Writing – review & editing, Writing – original draft, Formal analysis, Data curation, Conceptualization. **Hafsa Faiz:** Writing – review & editing, Writing – original draft, Formal analysis, Data curation, Conceptualization. **Sara AI-Douri:** Writing – review & editing, Writing – original draft, Formal analysis, Data curation, Conceptualization. **Khunsa Faiz:** Writing – review & editing, Writing – original draft, Formal analysis, Data curation, Conceptualization. **Alexandra De Sequeira:** Writing – review & editing, Writing – original draft, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] G.S. Tandel, A. Balestrieri, T. Jujaray, N.N. Khanna, L. Saba, J.S. Suri, Multiclass magnetic resonance imaging brain tumor classification using artificial intelligence paradigm, Comput Biol Med 122 (Jul. 2020) 103804, https://doi.org/10.1016/j.compbiomed.2020.103804.

[2] P. Wesseling, D. Capper, WHO 2016 Classification of gliomas, Neuropathol Appl Neurobiol 44 (2) (Feb. 2018) 139–150, https://doi.org/10.1111/nan.12432.

[3] S.J. Price, et al., Improved delineation of glioma margins and regions of infiltration with the use of diffusion tensor imaging: an image-guided biopsy study, AJNR Am J Neuroradiol 27 (9) (Oct. 2006) 1969–1974.

[4] N. Grech, T. Dalli, S. Mizzi, L. Meilak, N. Calleja, A. Zrinzo, Rising incidence of glioblastoma multiforme in a well-defined population, Cureus (May 2020), https://doi.org/10.7759/cureus.8195.

[5] S. Aneja, E. Chang, A. Omuro, Applications of artificial intelligence in neuro-oncology, Curr Opin Neurol 32 (6) (Dec. 2019) 850–856, https://doi.org/10.1097/WCO.0000000000000761.

[6] M. Cè, et al., Artificial intelligence in brain tumor imaging: a step toward personalized medicine, Current Oncology vol. 30, no. 3. MDPI (Mar. 01, 2023.) 2673–2701, https://doi.org/10.3390/curroncol30030203.

[7] M. Zhu, et al., Artificial intelligence in the radiomic analysis of glioblastomas: A review, taxonomy, and perspective, Front Oncol 12 (Aug. 2022), https://doi.org/10.3389/fonc.2022.924245.

[8] M. Rowe, An introduction to machine learning for clinicians, Academic Medicine 94 (10) (Oct. 2019) 1433–1436, https://doi.org/10.1097/ACM.0000000000002792.

[9] I.H. Sarker, Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions, SN Comput Sci 2 (6) (Nov. 2021) 420, https://doi.org/10.1007/s42979-021-00815-1.

[10] G. Sun, Y.-H. Zhou, AI in healthcare: navigating opportunities and challenges in digital communication, Front Digit Health 5 (Dec. 2023), https://doi.org/10.3389/fdgth.2023.1291132.

[11] A.W. Salehi, et al., A Study of CNN and transfer learning in medical imaging: advantages, challenges, future scope, Sustainability 15 (7) (Mar. 2023) 5930, https://doi.org/10.3390/su15075930.

[12] K. Al-hammuri, F. Gebali, A. Kanan, I.T. Chelvan, Vision transformer architecture and applications in digital health: a tutorial and survey, Vis Comput Ind Biomed Art 6 (1) (Jul. 2023) 14, https://doi.org/10.1186/s42492-023-00140-9.

[13] E. Montagnon, et al., Deep learning workflow in radiology: a primer, Insights Imaging 11 (1) (Dec. 2020) 22, https://doi.org/10.1186/s13244-019-0832-5.

[14] D.R. Sarvamangala, R.V. Kulkarni, Convolutional neural networks in medical image understanding: a survey, Evol Intell 15 (1) (Mar. 2022) 1–22, https://doi.org/10.1007/s12065-020-00540-3.

[15] Y. Yang, et al., A comparative analysis of eleven neural networks architectures for small datasets of lung images of COVID-19 patients toward improved clinical decisions, Comput Biol Med 139 (Dec. 2021) 104887, https://doi.org/10.1016/j.compbiomed.2021.104887.

[16] S. Vagvala, J.P. Guenette, C. Jaimes, R.Y. Huang, Imaging diagnosis and treatment selection for brain tumors in the era of molecular therapeutics, Cancer Imaging 22 (1) (19, Apr. 2022,) pp, https://doi.org/10.1186/s40644-022-00455-5.

[17] Z. Zhou, et al., Computer-aided detection of brain metastases in T1-weighted MRI for stereotactic radiosurgery using deep learning single-shot detectors, Radiology 295 (2) (May 2020) 407–415, https://doi.org/10.1148/radiol.2020191479.

[18] Y. Won Park *et al.*, "Robust performance of deep learning for automatic detection and segmentation of brain metastases using three-dimensional black-blood and three-dimensional gradient echo imaging", doi: 10.1007/s00330-021-07783-3/ Published.

[19] P. Blanc-Durand, A. Van Der Gucht, N. Schaefer, E. Itti, J.O. Prior, Automatic lesion detection and segmentation of 18F-FET PET in gliomas: A full 3D U-Net convolutional neural network study, PLoS One 13 (4) (Apr. 2018) e0195798.

[20] Q. Tian, et al., Radiomics strategy for glioma grading using texture features from multiparametric MRI, Journal of Magnetic Resonance Imaging 48 (6) (Dec. 2018) 1518–1528, https://doi.org/10.1002/jmri.26010.

[21] Y. Yang, et al., Glioma grading on conventional MR images: a deep learning study with transfer learning, Front Neurosci 12 (Nov. 2018), https://doi.org/10.3389/fnins.2018.00804.

[22] S. Wu, H. Li, D. Quang, Y. Guan, Three-Plane–assembled Deep Learning Segmentation of Gliomas, Radiol Artif Intell 2 (2) (Mar. 2020) e190011.

[23] S. Bakas, et al., Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features, Sci Data 4 (1) (Sep. 2017) 170117, https://doi.org/10.1038/sdata.2017.117.

[24] A.A. Akinyelu, F. Zaccagna, J.T. Grist, M. Castelli, L. Rundo, Brain tumor diagnosis using machine learning, convolutional neural networks, capsule neural networks and vision transformers, applied to MRI: a survey, J Imaging 8 (8) (Jul. 2022), https://doi.org/10.3390/jimaging8080205.

[25] D.N. Louis, et al., The 2016 world health organization classification of tumors of the central nervous system: a summary, Acta Neuropathol 131 (6) (Jun. 2016) 803–820, https://doi.org/10.1007/s00401-016-1545-1.

[26] A. Zlochower, D.S. Chow, P. Chang, D. Khatri, J.A. Boockvar, C.G. Filippi, Deep learning AI applications in the imaging of glioma, Topics in Magnetic Resonance Imaging 29 (2) (Apr. 2020) 115, https://doi.org/10.1097/RMR.0000000000000237.

[27] J. Beiko, et al., IDH1 mutant malignant astrocytomas are more amenable to surgical resection and have a survival benefit associated with maximal surgical resection, Neuro Oncol 16 (1) (Jan. 2014) 81–91, https://doi.org/10.1093/neuonc/not159.

[28] Cancer Genome Atlas Research Network, Comprehensive genomic characterization defines human glioblastoma genes and core pathways, Nature 455 (7216) (Oct. 2008) 1061–1068, https://doi.org/10.1038/nature07385.

[29] P. Chang, et al., Deep-learning convolutional neural networks accurately classify genetic mutations in gliomas, American Journal of Neuroradiology 39 (7) (Jul. 2018) 1201–1207, https://doi.org/10.3174/ajnr.A5667.

[30] J. Xiong, et al., Combination of diffusion tensor imaging and conventional MRI correlates with isocitrate dehydrogenase 1/2 mutations but not 1p/19q genotyping in oligodendroglial tumours, Eur Radiol 26 (6) (Jun. 2016) 1705–1715, https://doi.org/10.1007/s00330-015-4025-4.

[31] A.L. Rivera, et al., MGMT promoter methylation is predictive of response to radiotherapy and prognostic in the absence of adjuvant alkylating chemotherapy for glioblastoma, Neuro Oncol 12 (2) (Feb. 2010) 116–121, https://doi.org/10.1093/neuonc/nop020.

[32] Z. Akkus, et al., Predicting deletion of chromosomal arms 1p/19q in low-grade gliomas from MR images using machine intelligence, J Digit Imaging 30 (4) (Aug. 2017) 469–476, https://doi.org/10.1007/s10278-017-9984-3.

[33] L. Meng, R. Zhang, L. Fa, L. Zhang, L. Wang, and G. Shao, "ATRX status in patients with gliomas: Radiomics analysis.," *Medicine*, vol. 101, no. 37, p. e30189, Sep. 2022, doi: 10.1097/MD.0000000000030189.

[34] J. Haubold, et al., Fully automated MR based virtual biopsy of cerebral gliomas, Cancers (basel) 13 (24) (Dec. 2021) 6186, https://doi.org/10.3390/cancers13246186.

[35] R. Jain, et al., Genomic mapping and survival prediction in glioblastoma: molecular subclassification strengthened by hemodynamic imaging biomarkers, Radiology 267 (1) (Apr. 2013) 212–220, https://doi.org/10.1148/radiol.12120846.

[36] L. Macyszyn, et al., Imaging patterns predict patient survival and molecular subtype in glioblastoma via machine learning techniques, Neuro Oncol 18 (3) (Mar. 2016) 417–425, https://doi.org/10.1093/neuonc/nov127.

[37] S. Williams *et al.*, "Artificial Intelligence in Brain Tumour Surgery-An Emerging Paradigm.," *Cancers (Basel)*, vol. 13, no. 19, Oct. 2021, doi: 10.3390/cancers13195010.

[38] R.-S. Hsiao, Z. Mi, B.-R. Yang, L.-J. Kau, M.A. Bitew, T.-Y. Li, Body posture recognition and turning recording system for the care of bed bound patients, Technol Health Care 24 (Suppl 1) (2015) S307–S312, https://doi.org/10.3233/THC-151088.

[39] J. Howcroft, J. Kofman, E.D. Lemaire, Prospective fall-risk prediction models for older adults based on wearable sensors, IEEE Trans Neural Syst Rehabil Eng 25 (10) (Oct. 2017) 1812–1820, https://doi.org/10.1109/TNSRE.2017.2687100.

[40] S. Narang, et al., Tumor image-derived texture features are associated with CD3 T-cell infiltration status in glioblastoma, Oncotarget 8 (60) (Nov. 2017) 101244–101254, https://doi.org/10.18632/oncotarget.20643.

[41] X. Hu, K.K. Wong, G.S. Young, L. Guo, S.T. Wong, Support vector machine multiparametric MRI identification of pseudoprogression from tumor recurrence in patients with resected glioblastoma, Journal of Magnetic Resonance Imaging 33 (2) (Feb. 2011) 296–305, https://doi.org/10.1002/jmri.22432.
[42] B.-S. Jang, S.H. Jeon, I.H. Kim, I.A. Kim, Prediction of pseudoprogression versus progression using machine learning algorithm in glioblastoma, Sci Rep 8 (1) (Aug. 2018) 12516, https://doi.org/10.1038/s41598-018-31007-2.
[43] A. Saha, S. Ghosh, C. Roy, K. Choudhury, B. Chakrabarty, R. Sarkar, Demographic and clinical profile of patients with brain metastases: A retrospective study, Asian J Neurosurg 8 (03) (Sep. 2013) 157–161, https://doi.org/10.4103/1793-5482.121688.
[44] S. Bae, et al., Robust performance of deep learning for distinguishing glioblastoma from single brain metastasis using radiomic features: model development and validation, Sci Rep 10 (1) (2020) Dec, https://doi.org/10.1038/s41598-020-68980-6.
[45] N.C. Swinburne, et al., Machine learning for semiautomated classification of glioblastoma, brain metastasis and central nervous system lymphoma using magnetic resonance advanced imaging, Ann Transl Med 7 (11) (Jun. 2019) 232, https://doi.org/10.21037/atm.2018.08.05.
[46] R. Ortiz-Ramón, A. Larroza, S. Ruiz-España, E. Arana, D. Moratal, Classifying brain metastases by their primary site of origin using a radiomics approach based on texture analysis: a feasibility study, Eur Radiol 28 (11) (Nov. 2018) 4514–4523, https://doi.org/10.1007/s00330-018-5463-6.
[47] J.-H. Oh, K.M. Lee, H.-G. Kim, J.T. Yoon, E.J. Kim, Deep learning-based detection algorithm for brain metastases on black blood imaging, Sci Rep 12 (1) (Nov. 2022) 19503, https://doi.org/10.1038/s41598-022-23687-8.
[48] T.J. Kaufmann, et al., Consensus recommendations for a standardized brain tumor imaging protocol for clinical trials in brain metastases, Neuro Oncol 22 (6) (Jun. 2020) 757–772, https://doi.org/10.1093/neuonc/noaa030.
[49] J. Park, E.Y. Kim, Contrast-enhanced, three-dimensional, whole-brain, black-blood imaging: Application to small brain metastases, Magn Reson Med 63 (3) (Mar. 2010) 553–561, https://doi.org/10.1002/mrm.22261.
[50] Y. Han, et al., Differentiation between glioblastoma multiforme and metastasis from the lungs and other sites using combined clinical/routine MRI radiomics, Front Cell Dev Biol 9 (Aug. 2021), https://doi.org/10.3389/fcell.2021.710461.
[51] S. J. Cho, L. Sunwoo, S. H. Baik, Y. J. Bae, B. S. Choi, and J. H. Kim, "Brain metastasis detection using machine learning: a systematic review and meta-analysis," *Neuro-Oncology*, vol. 23, no. 2. Oxford University Press, pp. 214–225, Feb. 01, 2021. doi: 10.1093/neuonc/noaa232.
[52] A. Bose, U. Prasad, A. Kumar, M. Kumari, S.K. Suman, D.K. Sinha, Characterizing various posterior fossa tumors in children and adults with diffusion-weighted imaging and spectroscopy, Cureus (May 2023), https://doi.org/10.7759/cureus.39144.
[53] F. Nejat, M. El Khashab, J.T. Rutka, Initial management of childhood brain tumors: neurosurgical considerations, J Child Neurol 23 (10) (Oct. 2008) 1136–1148, https://doi.org/10.1177/0883073808321768.
[54] S. Payabvash, M. Aboian, T. Tihan, S. Cha, Machine learning decision tree models for differentiation of posterior fossa tumors using diffusion histogram analysis and structural MRI findings, Front Oncol 10 (Feb. 2020), https://doi.org/10.3389/fonc.2020.00071.
[55] J.L. Quon, et al., Deep learning for pediatric posterior fossa tumor detection and classification: a multi-institutional study, American Journal of Neuroradiology (Aug. 2020), https://doi.org/10.3174/ajnr.A6704.
[56] A.G. Yearley, et al., Machine learning in the classification of pediatric posterior fossa tumors: a systematic review, Cancers (basel) 14 (22) (Nov. 2022) 5608, https://doi.org/10.3390/cancers14225608.
[57] D. P. Bray *et al.*, "Artificial neural networks predict the need for permanent cerebrospinal fluid diversion following posterior fossa tumor resection," *Neurooncol Adv*, vol. 4, no. 1, Jan. 2022, doi: 10.1093/noajnl/vdac145.
[58] C. Dai, B. Sun, R. Wang, J. Kang, The application of artificial intelligence and machine learning in pituitary adenomas, Front Oncol 11 (Dec. 2021), https://doi.org/10.3389/fonc.2021.784819.
[59] A. Acitores Cancela, V. Rodríguez Berrocal, H. Pian, J. S. Martínez San Millán, J. J. Díez, and P. Iglesias, "Clinical relevance of tumor consistency in pituitary adenoma," *Hormones*, vol. 20, no. 3, pp. 463–473, Sep. 2021, doi: 10.1007/s42000-021-00302-5.
[60] Y. Fan, et al., Preoperative noninvasive radiomics approach predicts tumor consistency in patients with acromegaly: development and multicenter prospective validation, Front Endocrinol (lausanne) 10 (Jun. 2019), https://doi.org/10.3389/fendo.2019.00403.
[61] A. Zeynalova, et al., Preoperative evaluation of tumour consistency in pituitary macroadenomas: a machine learning-based histogram analysis on conventional T2-weighted MRI, Neuroradiology 61 (7) (Jul. 2019) 767–774, https://doi.org/10.1007/s00234-019-02211-2.
[62] Y. Fang, et al., Application of convolutional neural network in the diagnosis of cavernous sinus invasion in pituitary adenoma, Front Oncol 12 (Apr. 2022), https://doi.org/10.3389/fonc.2022.835047.
[63] J. Niu, et al., Preoperative prediction of cavernous sinus invasion by pituitary adenomas using a radiomics method based on magnetic resonance images, Eur Radiol 29 (3) (Mar. 2019) 1625–1634, https://doi.org/10.1007/s00330-018-5725-3.
[64] H. Ucuzal, S. YASAR, C. Colak, Classification of brain tumor types by deep learning with convolutional neural network on magnetic resonance images using a developed web-based interface, in: In 2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), IEEE, Oct. 2019, pp. 1–5, https://doi.org/10.1109/ISMSIT.2019.8932761.
[65] A. ul Haq, J.P. Li, S. Khan, M.A. Alshara, R.M. Alotaibi, C.B. Mawuli, DACBT: deep learning approach for classification of brain tumors using MRI data in IoT healthcare environment, Sci Rep 12 (1) (Dec. 2022), https://doi.org/10.1038/s41598-022-19465-1.
[66] M.M. Rech, et al., Machine learning models to forecast outcomes of pituitary surgery: a systematic review in quality of reporting and current evidence, Brain Sci 13 (3) (Mar. 2023), https://doi.org/10.3390/brainsci13030495.
[67] N. Qiao, et al., Machine learning in predicting early remission in patients after surgical treatment of acromegaly: a multicenter study, Pituitary 24 (1) (Feb. 2021) 53–61, https://doi.org/10.1007/s11102-020-01086-4.
[68] Y. Fan, S. Jiang, M. Hua, S. Feng, M. Feng, R. Wang, Machine learning-based radiomics predicts radiotherapeutic response in patients with acromegaly, Front Endocrinol (lausanne) 10 (Aug. 2019), https://doi.org/10.3389/fendo.2019.00588.
[69] Y. Zhang, et al., Radiomics approach for prediction of recurrence in non-functioning pituitary macroadenomas, Front Oncol 10 (Dec. 2020), https://doi.org/10.3389/fonc.2020.590083.
[70] L.F. Machado, P.C.L. Elias, A.C. Moreira, A.C. dos Santos, L.O. Murta Junior, MRI radiomics for the prediction of recurrence in patients with clinically non-functioning pituitary macroadenomas, Comput Biol Med 124 (Sep. 2020) 103966, https://doi.org/10.1016/j.compbiomed.2020.103966.
[71] Y. Lan, S. Zou, B. Qin, X. Zhu, Potential roles of transformers in brain tumor diagnosis and treatment, Brain-X 1 (2) (Jun. 2023), https://doi.org/10.1002/brx2.23.
[72] A. Anaya-Isaza, L. Mera-Jiménez, L. Verdugo-Alejo, L. Sarasti, Optimizing MRI-based brain tumor classification and detection using AI: A comparative analysis of neural networks, transfer learning, data augmentation, and the cross-transformer network, Eur J Radiol Open 10 (Jan. 2023), https://doi.org/10.1016/j.ejro.2023.100484.
[73] W.H.L. Pinaya, et al., Unsupervised brain imaging 3D anomaly detection and segmentation with transformers, Med Image Anal 79 (Jul. 2022), https://doi.org/10.1016/j.media.2022.102475.
[74] J. Lin, et al., CKD-TransBTS: clinical knowledge-driven hybrid transformer with modality-correlated cross-attention for brain tumor segmentation, IEEE Trans Med Imaging 42 (8) (Aug. 2023) 2451–2461, https://doi.org/10.1109/TMI.2023.3250474.
[75] M.A. Fink, et al., "Deep learning–based assessment of oncologic outcomes from natural language processing of structured radiology reports", radiol, Artif Intell 4 (5) (Sep. 2022), https://doi.org/10.1148/ryai.220055.
[76] J. Wu, Q. Xu, Y. Shen, W. Chen, K. Xu, X.R. Qi, Swin Transformer Improves the IDH Mutation Status Prediction of Gliomas Free of MRI-Based Tumor Segmentation, J Clin Med 11 (15) (Aug. 2022), https://doi.org/10.3390/jcm11154625.
[77] Q. Lyu, et al., A transformer-based deep-learning approach for classifying brain metastases into primary organ sites using clinical whole-brain MRI images, Patterns 3 (11) (Nov. 2022), https://doi.org/10.1016/j.patter.2022.100613.
[78] S. Oh, N. Kim, J. Ryu, Analyzing to discover origins of CNNs and ViT architectures in medical images, Sci Rep 14 (1) (Apr. 2024) 8755, https://doi.org/10.1038/s41598-024-58382-3.
[79] J. D. Rudie, A. M. Rauschecker, R. N. Bryan, C. Davatzikos, and S. Mohan, "Emerging Applications of Artificial Intelligence in Neuro-Oncology," *Radiology*, vol. 290, no. 3. NLM (Medline), pp. 607–618, Mar. 01, 2019. doi: 10.1148/radiol.2018181928.
[80] E. Loken, A. Gelman, Measurement error and the replication crisis, Science 355 (6325) (Feb. 2017) 584–585, https://doi.org/10.1126/science.aal3618.