



Review

Deep CNNs for glioma grading on conventional MRIs: Performance analysis, challenges, and future directions

Sonam Saluja*, Munesh Chandra Trivedi and Ashim Saha

Department of Computer Science and Engineering, National Institute of Technology Agartala, Tripura 799046, India

* **Correspondence:** Email: sonamcse.sch@nita.ac.in.

Abstract: The increasing global incidence of glioma tumors has raised significant healthcare concerns due to their high mortality rates. Traditionally, tumor diagnosis relies on visual analysis of medical imaging and invasive biopsies for precise grading. As an alternative, computer-assisted methods, particularly deep convolutional neural networks (DCNNs), have gained traction. This research paper explores the recent advancements in DCNNs for glioma grading using brain magnetic resonance images (MRIs) from 2015 to 2023. The study evaluated various DCNN architectures and their performance, revealing remarkable results with models such as hybrid and ensemble based DCNNs achieving accuracy levels of up to 98.91%. However, challenges persisted in the form of limited datasets, lack of external validation, and variations in grading formulations across diverse literature sources. Addressing these challenges through expanding datasets, conducting external validation, and standardizing grading formulations can enhance the performance and reliability of DCNNs in glioma grading, thereby advancing brain tumor classification and extending its applications to other neurological disorders.

Keywords: brain tumors; classification; deep learning; gliomas grading; MRIs

1. Introduction

The brain is an incredibly complex organ, and its function relies on the well-coordinated activity of diverse cell types. The brain is composed of two main types of cells: neurons and glial cells. Neurons, also known as nerve cells, are responsible for transmitting electrical and chemical signals in the brain,

enabling functions such as thinking, feeling, and movement. Glial cells, or neuroglia, support and modulate the activity of neurons. There are several types of glial cells, including astrocytes, oligodendrocytes, ependymal and microglia, each with specific functions. Astrocytes regulate neurotransmission, form the blood-brain barrier, and support and nourish nerve cells. Oligodendrocytes are involved in the production of myelin, which insulates nerve fibers for rapid signal transmission. Ependymal cells serve as a lining for the ventricles of the brain and the central canal of the spinal cord. They contribute significantly to the homeostasis of the central nervous system (CNS) by regulating fluid balance, providing structural support, and participating in neurogenesis. Microglia act as the primary form of immune defense in the CNS, protecting the brain and spinal cord from infection and injury. The roles of glial cells are diverse and essential for maintaining brain homeostasis, supporting neuronal function, and regulating the brain response to injury and disease. Neurons and glial cells work together to ensure the proper functioning of the brain. Understanding the functions of both cell types is crucial for comprehending the complexities of brain function and the impact of various brain disorders. Glioma is one of the primary brain cancers that originates in the glial cells of the brain. Approximately one-third of CNS cancers are gliomas. Gliomas are categorized according to their subgroups and a numerical grading system. According to the American Cancer Society, three subtypes of gliomas are astrocytomas, oligodendrogliomas, and ependymomas [1]. The grade of a tumor relates to the microscopic appearance of these subtype cancer cells. The 2021 world health organization (WHO) classification of cancers of the CNS categorizes glioma tumors into four categories depending on the progression of malignancy aggressiveness [2,3]. Grade I tumors grow slowly and are sometimes entirely resectable with surgery, but grade IV tumors are aggressive, rapidly growing, and challenging to treat. The most frequent primary tumors were astrocytomas (38.7%), with high grade gliomas (HGGs) (59.5%) making up the majority [4]. The clinical course for a particular patient considers the tumor location, potential symptoms, and the viability of alternative treatment techniques. Hence early detection of tumor cells is crucial for treating patients. Due to the availability of cutting-edge diagnostic and therapeutic tools, physicians can now effectively diagnose patients and administer treatment without endangering their health. One of the most reliable approaches to accomplishing this goal is using medical imaging. Using imaging technology, doctors can look for anomalies in a patient's bones and tissues without cutting them open. Patients with brain tumors benefit significantly from the use of healthcare imaging methods such as X-ray, magnetic resonance imaging (MRI), ultrasound, magnetic resonance spectroscopy, and computed tomography [5]. MRI is one of the most preferred noninvasive neuroimaging techniques to diagnose brain tumors as it provides high-contrast images, especially in the case of soft tissues [6].

One of the distinguishing features of modern healthcare operations is that, as the number of patients increases, they generate enormous amounts of data on various interconnected procedures. In comparison to other aspects of healthcare, the generation of data from medical imaging is by far the most prolific, and this production rate is accelerating at an exponential rate. On the other hand, the data volume often exceeds traditional analysis capacity. This is a critical problem to address because proper data interpretation is one of the fundamental building blocks for complex systems like medical imaging. The second problem with human interpretation is that it is prone to inaccuracy for several reasons, such as being under stress, not having enough background, and not having enough experience. Therefore, the solution that makes the most sense is to employ artificial intelligence (AI). Applications that use machine learning / deep learning (ML/DL) may perform data analysis substantially more accurately and quickly, making it much simpler for medical professionals to handle information and

carefully evaluate test results [7]. In medical image analysis, the deep convolutional neural network (DCNN) has recently gained the most popularity. A CNN can automate and optimize image segmentation procedures by utilizing a wide range of classification and segmentation algorithms that extract as many relevant details as required from the data. DL allows images to be fed directly to CNNs, and important features can be learned automatically. Simple features within images are learned at shallow layers and deeper layers near the output layer are known to learn more complex high-order features [8]. The quality and quantity of the dataset with annotations substantially influence the DL algorithm performance. However, annotating a large number of medical images is problematic since annotation can be time-consuming and is knowledge-specific [9]. In the case of a limited training dataset, transfer learning (TL) is a promising approach. It improves a network that has been previously trained on a vast labeled dataset from some other field. Applying learnt information to the target dataset speeds up network convergence while reducing computational costs during training [10]. Although DL algorithms can analyze medical images with high accuracy, they have yet to replace the role of a human specialist due to various challenges, such as a lack of sufficient data for training, a data imbalance problem, and a lack of a connection between clinicians and researchers. This structured review aims to assess recent advances in the automatic identification and classification of glioma tumors using a DL framework. In this review, we look at recent advances in DCNN techniques for glioma tumor classification, current research limitations, and future research directions in this field.

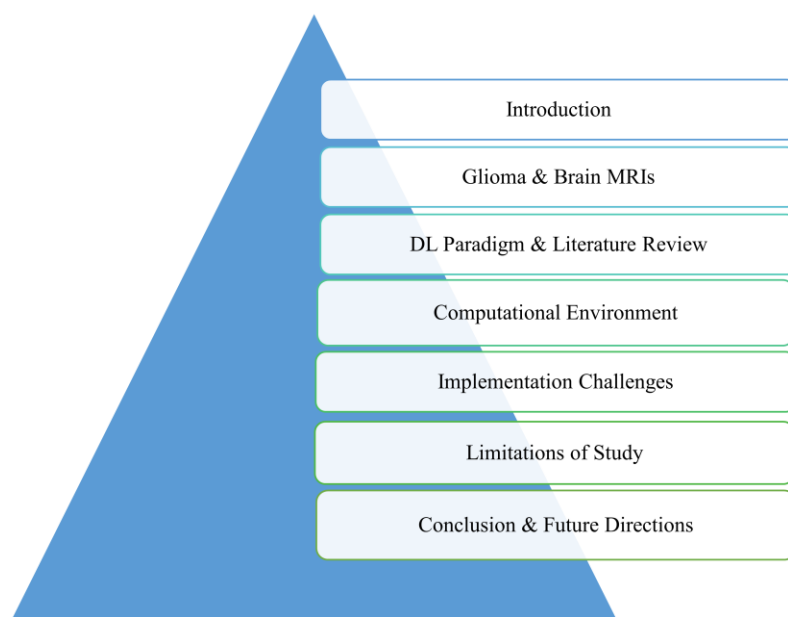


Figure 1. The flow of the conceptual framework of this research paper.

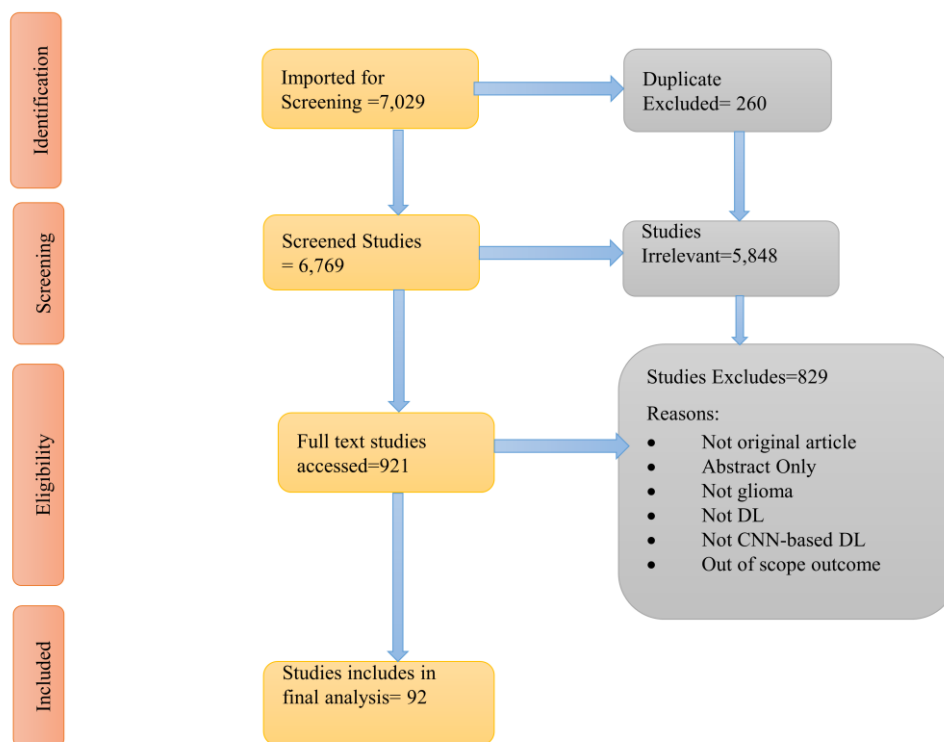


Figure 2. Overview of screening and selection process.

This comprehensive review aims to provide researchers with the most up-to-date information in the brain MRI image classification field, including the advantages and disadvantages of existing DL techniques and algorithms. Figure 1 depicts the conceptual framework presented in this review paper. In total, 7029 records were retrieved through the search process. After a comprehensive assessment, 921 full-text articles were meticulously examined. Among these, 829 articles were deemed irrelevant and subsequently excluded, resulting in 92 studies that were considered for further analysis. The entire study screening and selection process is visually represented in Figure 2. For data collection, a meticulous approach was adopted wherein key data elements such as study purpose, methodology, model performance, and risk of bias were extracted and summarized for each of the 108 included studies. To ensure a comprehensive coverage of relevant literature, a variety of online scientific research repositories were consulted. This included well-regarded sources such as IEEE Xplore, Medline, Google Scholar, ScienceDirect, and ResearchGate. Notably, the search was refined to cover articles published between 2015 and 2023 to ensure relevance to the selected time frame. The search strategy employed a robust combination of domain-specific and methodological search terms, totaling 180 distinct combinations.

The structure of the review is as follows: Sections 2 and 3 contain detailed information regarding the glioma tumor grades, MR imaging, and available imaging databases for tumor classification. In Section 4, we delve into the DL paradigm in imaging and discuss the evolution of techniques utilized

by DCNN architectures and in medical imaging. Moving to Section 5, we outline the fundamental stages of DCNN approaches for classifying glioma tumors and present an overview of pertinent primary studies, datasets, and computational methods utilized for developing glioma classification models, along with their respective performance evaluations. Section 6 is dedicated to discussing the implementation challenges associated with the studied architecture. Finally, Sections 7 and 8 encompass the limitations of this study and our concluding remarks. In Section 9, we offer recommendations for enhancing future research in this domain.

2. Glioma grading

Glioma is an umbrella term for primary brain tumors that are categorized based on their putative cell of origin. The WHO classification is the international standard for glioma diagnosis [3]. According to histology criteria [11], glioma tumors are classified into four categories based on the degree of aggressiveness. The histological features that contribute to each glioma grade include cellularity (cell number), mitotic activity (cell division rate), pleomorphism (cell size/shape variation), necrosis (dead tissue presence), and vascularity (blood vessel density) and endothelial proliferation (increased blood vessel growth). Knowing the type of glioma before surgery or other therapies is crucial for clinical planning and decision-making. Figure 3 describes glioma grades and their characteristics according to WHO [3]. The 5-year survival statistics for each glioma grade are detailed in Table 1.

Grade I gliomas are slow-growing astrocytomas made up of pilocytic cells that do not spread to other organs of the body. They exhibit minimal mitotic activity and lack necrosis. These tumors are the safest because they grow slowly, have clear borders, and have the best chance of survival. So, they can be removed surgically and cured with a low chance of returning [12]. Most of the time, these gliomas are found in children and young adults. Grade I gliomas are also called low grade gliomas (LGGs). Grade II gliomas are harmless and more common in adults. These gliomas tend to spread into nearby healthy tissue and have fuzzy edges. Because of this, it is hard to get rid of them with surgery. They show increased cellular atypia (abnormalities) and mitotic activity compared to Grade I, with rare focal necrosis permissible. Depending on the location and size, chemotherapy and radiation can be used as treatments [13]. Since the outlook is better than grades III-IV, they are in the LGG group. Grade III are types of glioma tumors that are cancerous. They display marked cellular atypia, frequent mitoses, and widespread necrosis, indicating their malignant nature. They are also called anaplastic gliomas. The word “anaplastic” is used to describe glioma brain tumor cells that divide quickly. Some cases of astrocytoma or oligodendroglioma transform into the aggressive form. It is harder to deal with than LGGs [14], also known as HGGs. Grade III tumors tend to spread rapidly and are likely to become grade IV tumors. Grade IV are the most malignant glioma tumors and have the lowest survival rate. They exhibit extreme cellular atypia, brisk mitoses, extensive necrosis, and microvascular proliferation (new blood vessels), hinting at their invasive potential. Primary glioblastomas grow quickly, while LGGs can turn into secondary glioblastomas. They often happen to older people and rarely to children [15]. The 2021 WHO classification emphasizes a layered approach that integrates molecular markers such as isocitrate dehydrogenase (IDH) mutation and 1p/19q codeletion alongside traditional histopathological features for a more accurate diagnosis and prognosis [3].

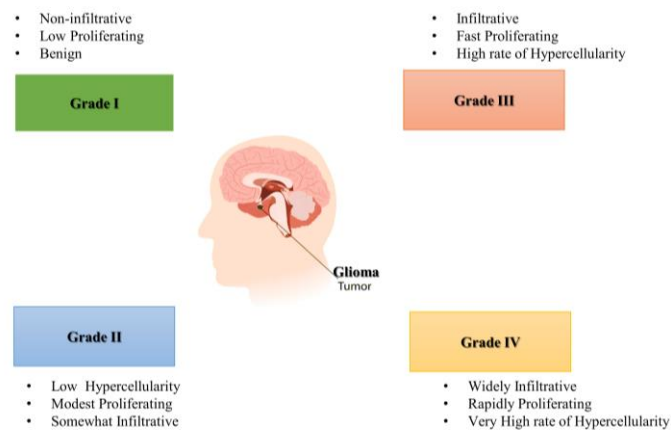


Figure 3. Glioma grades and their characteristics according to WHO [3].

Table 1. Classification of gliomas into WHO grades, characteristics, and 5-year survival rate [16,17]

Glioma Grades	Grade Type	Glioma Type	Characteristics	Prognosis (5 Year Survival Rate in Adults)
I	LGG	pilocytic astrocytomas [18]	<ul style="list-style-type: none"> • slow growing, • well-defined borders, • good prognosis. 	~95%
II	LGG	diffuse astrocytomas oligodendroglioma [19]	<ul style="list-style-type: none"> • slow growing, • invades neighbouring tissue, • good prognosis. 	~48–80%
III	HGG	anaplastic astrocytoma anaplastic oligodendroglioma [20]	<ul style="list-style-type: none"> • tumor cells do not have a uniform appearance, • fastest growing, • invades neighbouring tissue, • poor prognosis. 	~34–62%
IV	HGG	Glioblastoma [21]	<ul style="list-style-type: none"> • composed of numerous different cell types, • fastest growing, • more than half of all gliomas are glioblastoma, • can occur or as a result of a lower grade astrocytoma or oligodendroglioma, • poor prognosis. 	~11%

3. MRI

The use of imaging technology is essential for treating intracranial tumors. In recent years, numerous medical imaging tools have been developed to aid clinicians in diagnosing the character and location of the disease. MRI has become the benchmark for diagnosing and monitoring brain malignancies, and its uses continue to expand [22]. Improved neuro-oncological imaging not only enhances the detection of various lesions in the CNS but also permits the formulation of a more nuanced treatment approach. Both structural and functional MRI were found to have significant correlations with disease stage and prognosis in cancer patients. In recent years, MRI has received a lot of interest and appreciation because it is noninvasive and provides the finest contrast in cellular structure [23–25]. An MR scanner can capture many images of the subject under investigation from multiple viewpoints, with varying contrast and physical properties; this is known as multiple modality imaging [26]. Brain malignancies are often diagnosed using four MRI imaging sequences: T1W (T1 weighted), T2W (T2 weighted), T1Wc (T1 weighted post contrast), FLAIR (fluid-attenuated inversion recovery). Figure 4 shows sample MR sequences from the BraTS (brain tumor segmentation) challenge 2018 dataset. These sequences provide complementary information about the morphology and physiology of gliomas and enable a comprehensive assessment of the tumor, highlighting features like vascularity, edema, and infiltration patterns. In most segmentation approaches, T2 MRI is utilized. Due to the complicated structure and anatomy of the human brain, the radiologist uses all four MRI techniques to diagnose and classify the type of brain tumor. T1W scans can differentiate between healthy and diseased tissues. T2W scans delineate edematous areas. T1Wc pictures are utilized to locate the tumor boundary. FLAIR imaging can differentiate between edematous and cerebrospinal fluid-filled regions. The changes in the images produced by different MR modalities can be used to establish a contrast between the edema tissue, neoplastic tissue, necrosis tissue, and the unaffected brain, thus forming a tumor border. In addition, functional MRI (fMRI) and diffusion tensor imaging (DTI) are sometimes used to evaluate the alterations in brain function and connectivity induced by gliomas. These techniques provide insights into neuronal activity and the integrity of white matter, which are crucial for identifying eloquent brain areas and assessing the response to treatment. When compared to conventional imaging methods, such as computed tomography (CT) and positron emission tomography (PET), MRI offers superior soft tissue contrast, which enables the precise delineation of gliomas and their surrounding structures. This high contrast resolution is instrumental in facilitating accurate tumor localization and characterization, which are critical for diagnosis and treatment planning. Moreover, the noninvasive nature of MRI and the absence of ionizing radiation make it safe for repeated imaging, a feature that is particularly advantageous for pediatric and vulnerable populations. The use of gadolinium contrast in MRI can highlight areas with a disrupted blood-brain barrier, often indicative of tumor presence and activity. This further enhances the diagnostic accuracy and helps assess tumor aggressiveness. Although modalities such as PET and CT scans have their strengths, such as speed and affordability for initial evaluation or bone involvement and the ability to assess tumor metabolism respectively, they primarily excel in gross anatomical visualization. MRI offers exceptional detail, diverse imaging sequences, functional insights, and safety. This makes MRI an invaluable tool for comprehensive diagnosis, treatment planning, and monitoring of gliomas [27]. Table 2 summarizes the sources of the MRI datasets utilized in this review.

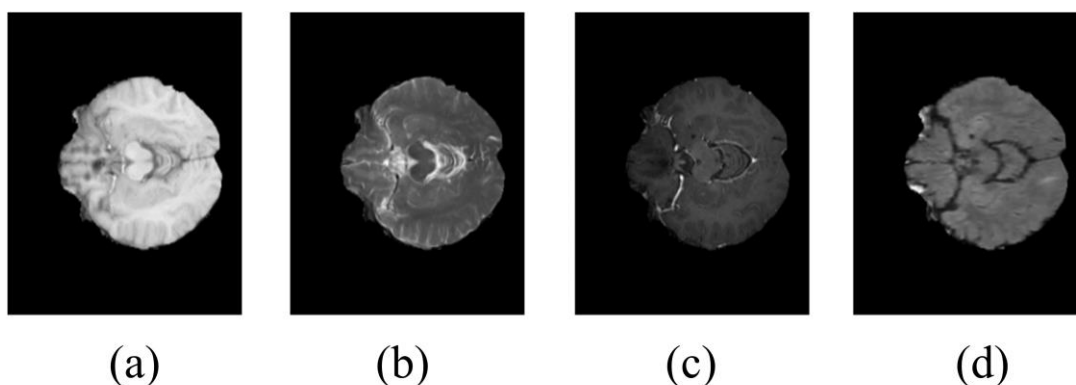


Figure 4. Examples of MR sequences from the BRATS 2018 HGG samples of a single subject. (a) T1W (b) T2W (c)T1Wc (d)FLAIR.

Table 2. An overview of commonly utilized publicly available datasets for brain tumor analysis.

Name	Modalities	Size (No. of Patients)	Sources
TCGA-GBM	T1W, T1Wc, T2W, FLAIR	199	[28–30]
TCGA-LGG	T1W, T1Wc, T2W, FLAIR	299	[31–33]
REMBRANDT	T1W, T2W, FLAIR, DWI	112	[34,35]
BraTS	T1W, T1Wc, T2W, FLAIR	2019: 335 (259 HGG, 76 LGG); 2018: 284 (209 HGG, 75 LGG); 2017: 285 (210 HGG, 75 LGG)	[36–38]
ClinicalTrials.gov	T1W, T1Wc, T2W, FLAIR	113 (52 LGG, 61 HGG)	[39]
Radiopaedia	T1W, T1Wc, T2W, FLAIR	121 (36 Grade I,32 Grade II ,25 Grade III,28 Grade IV)	[40]

4. Deep learning paradigm

While there has been progress in glioma treatment, it is far from sufficient. Before initiating therapy for gliomas, it is critical to determine the tumor stage accurately. The complex and diverse nature of gliomas, characterized by their multidimensional and heterogeneous features, necessitates the development of advanced, automated systems for accurate diagnosis. This urgent need stems from the inherent risks associated with traditional surgical methods like biopsies, especially for tumors located in critical brain regions. Automated systems, such as computer-aided diagnosis (CAD) and AI

algorithms, offer promising solutions by enhancing both tumor localization and classification precision. They can assist in glioma detection, grading, segmentation, and even knowledge discovery, leveraging extracted features to predict tumor characteristics. This provides invaluable insights to clinicians, guiding treatment decisions and optimizing patient outcome. Furthermore, automation streamlines the diagnostic process, reducing the burden on healthcare professionals and potentially expediting treatment initiation for glioma patients. In the past decade, ML has seen substantial expansion in its applications to the field of neuro-oncology, with the diagnosis of glioma tumors using MRI, emerging as a prominent focus of interest. Several authors have used traditional ML approaches, which entail a sequence of steps beginning with preprocessing, continuing with feature extraction and feature selection, and concluding with applying a classification algorithm to offer a result [41]. Several approaches were used to extract the features, including discrete wavelet transform, gray level co-occurrence matrix, histogram of oriented gradients, genetic algorithm, and zernike moments. Particle swarm optimization and principal component analysis have been used by several authors in this discipline to help them decide which features to include. The most extensively utilized classification strategy for classification was SVM (support vector machine), which multiple authors adopted. Other authors use random forest, adaboost technique, instance-based k-nearest with log and Gaussian weight kernels, extreme learning machine, and sequential minimal optimization as categorization strategies [42]. However, the quality of the classification process in ML studies largely depends on manually created features discovered by feature extraction techniques, which is a time-consuming and error-prone process. There are limitations to employing these manually created features, as they cannot be changed during model training, and it is uncertain if they are the most effective attributes for classification. Additionally, these features require rigorous validation and often exhibit limited generalizability, struggling to adapt to new patient populations or imaging protocols. This significantly hinders their applicability across diverse datasets and clinical scenarios [43,44]. Moreover, traditional ML architectures often encounter difficulties in integrating and effectively leveraging multimodality data such as MRI, PET, and genetic information, due to the complex relationships existing between these modalities. These challenges are particularly pronounced in glioma classification. Glioma datasets often vary in terms of imaging modalities, acquisition parameters, and tumor phenotypes, making it challenging for manually engineered features to adapt to such variability. Consequently, the performance of traditional ML models relying on manually created features may degrade when applied to new datasets or clinical scenarios. DL with its ability to automatically learn features directly from data, offers a promising solution to these challenges. By eliminating the need for manual feature engineering, DL models can capture more subtle and complex patterns in the data, potentially leading to improved glioma classification performance. Furthermore, DL models can be designed to effectively integrate multimodal data, thereby fully exploiting the complementary information provided by each modality.

DL is a subfield of ML. Here the processes of selecting features from images and classifying them are carried out concurrently by a single algorithm and learning does not require the participation of humans during the training process. Feature extraction is accomplished by a multilayer, nonlinear processing architecture. As we proceed deeper into the network, data abstraction is aided by the fact that each layer output serves as the input to the layer below it [45]. The usage of CNNs in various image processing problems is becoming increasingly common due to their prominence as a DL technique. CNNs ability to discern patterns has made it popular, especially in the image processing field. A CNN generally has three layers stacked on top of one another. The convolutional layer is

responsible for extracting features from images. It delivers visual knowledge of the dataset images to the network and addresses the use of learnable kernels. Each kernel is typically convolved across the spatial dimensions of the input by the convolutional layer to produce a feature map as an output. The pooling layer is responsible for minimizing the dimensionality of the features obtained in order to reduce the number of parameters and computational complexity of the model. The last layer employs multiple fully-connected layers that focus on converting the 2D feature maps of the preceding levels into 1D vectors. A learning or optimizer algorithm is utilized to modify network weights during training. The learning process uses loss to update the network's filters and weights. At the output layer, an activation function normalizes the output total, so all numbers add up to one [46,47].

The evolution of DCNN began in 1989, with the introduction of LeNet [48]. At the time, CNNs were limited to digit identification tasks, which could not be applied to other image analysis problems. From the 1996 to 2000, various developments in CNN architecture were created in order to make it scalable to large multi-class problems. CNN-based applications became popular following AlexNet's remarkable performance on the ImageNet dataset in 2012 [49]. Significant advancements have been made since then. Zeiler and Fergus [50] introduced a layer-by-layer representation of CNN to enhance comprehension of feature extraction stages, which shifted the paradigm toward feature extraction at low spatial resolution in DL architecture, as accomplished in VGG [51]. VGG stands for visual geometry group, which is a part of the department of science and engineering at oxford university. The Google DL group pioneered the concept of a split, transform and merge with the connecting block known as the inception block in GoogLeNet. These blocks introduced the concept called branching inside a layer, allowing for the abstraction of features at several spatial scales [52]. The idea of skip connections, proposed by residual network ResNet [53] for DCNN training, rose to prominence in 2015. Following that, most succeeding networks embraced this concept, like Inception-ResNet, Wide ResNet, and others [54]. A new network architecture called ResNeXt [55] was developed for image classification, focusing on increasing cardinality as a key factor for improving accuracy outperforming its ResNet counterpart on various datasets. MobileNet, designed for efficient mobile and embedded vision applications, brought a new level of model efficiency and portability to the field [56]. The neural architecture search (NAS) approach led to the creation of NASNet, which automates the design of CNN architectures and has produced competitive models for various tasks [57]. EfficientNet, proposed by Tan and Le in 2019, demonstrated remarkable efficiency-accuracy trade-offs by scaling model width, depth, and resolution simultaneously. It has become a popular choice for resource-constrained applications [58]. The evolution of DL architectures continued with the emergence of NFNet (normalizer-free ResNets) which built upon the success of Squeeze-and-Excitation blocks and the ReLU(rectified linear unit) activation function to achieve both computational efficiency and state of the art results in computer vision [59]. TResNet, inspired by the efficient combination of depthwise separable convolutions and spatial pyramid pooling, offers competitive performance on various computer vision tasks, including image classification [60]. The landscape of DL and CNNs has continued to evolve with the introduction of various novel architectures. Table 3 lists an overview of popular CNN architectures for image analysis.

Table 3. A summary of CNN architectures, their contribution, and limitations in image analysis.

Architecture	Year	Depth Range*	Contribution	Limitation
LeNet [48]	1998	Shallow	Pioneering CNN architecture for handwritten digit recognition	Limited capacity for complex image analysis tasks
AlexNet [49]	2012	Shallow	Popularized deep CNNs and won ImageNet competition	Prone to overfitting due to limited regularization
VGG [51]	2014	Shallow	Simplicity and uniform architecture led to strong performance	High computational requirements and memory usage
GoogLeNet (Inception) [52]	2014	Moderate	Introduced inception modules for efficient feature extraction	Complex architecture, challenging to optimize
Highway Network [61]	2015	Moderate	Use of multipath concept cross-layer connectivity mechanism	Because gates are data dependent, they may become expensive
ResNet [53]	2015	Very Deep	Introduced residual connections, enabling training of very deep networks	Some variants may suffer from overfitting
DenseNet [62]	2016	Moderate	Introduced dense connectivity patterns for feature reuse	Memory consumption increases with network growth
ResNeXt [55]	2016	Moderate	Introduced cardinality to improve representational power	Larger models can be computationally intensive
MobileNet [56]	2017	Shallow	Utilized depth-wise separable convolutions for lightweight networks	Reduced capacity for complex tasks
NASNet [57]	2017	Variable	Leveraged neural architecture search for automatic design	Computationally expensive search process
EfficientNet [58]	2019	Variable	Achieved high efficiency and accuracy via compound scaling	Some versions might require careful tuning
NFNet [59]	2020	Variable	Highly efficient due to the use of Squeeze-and-Excitation blocks and the FReLU activation function	Fine-tuning may require careful hyperparameter tuning, which can be time-consuming
TResNet [60]	2021	Variable	Efficient due to the combination of depthwise separable convolutions and spatial pyramid pooling.	Interpretability can be a challenge with TResNet, especially in larger variants, as it involves intricate operations

*Architectural depth classifications range from shallow (typically 1 to 10 layers), moderate (around 10 to 100 layers), very deep (often exceeding 100 layers), to variable, allowing significant depth variations beyond predefined ranges.

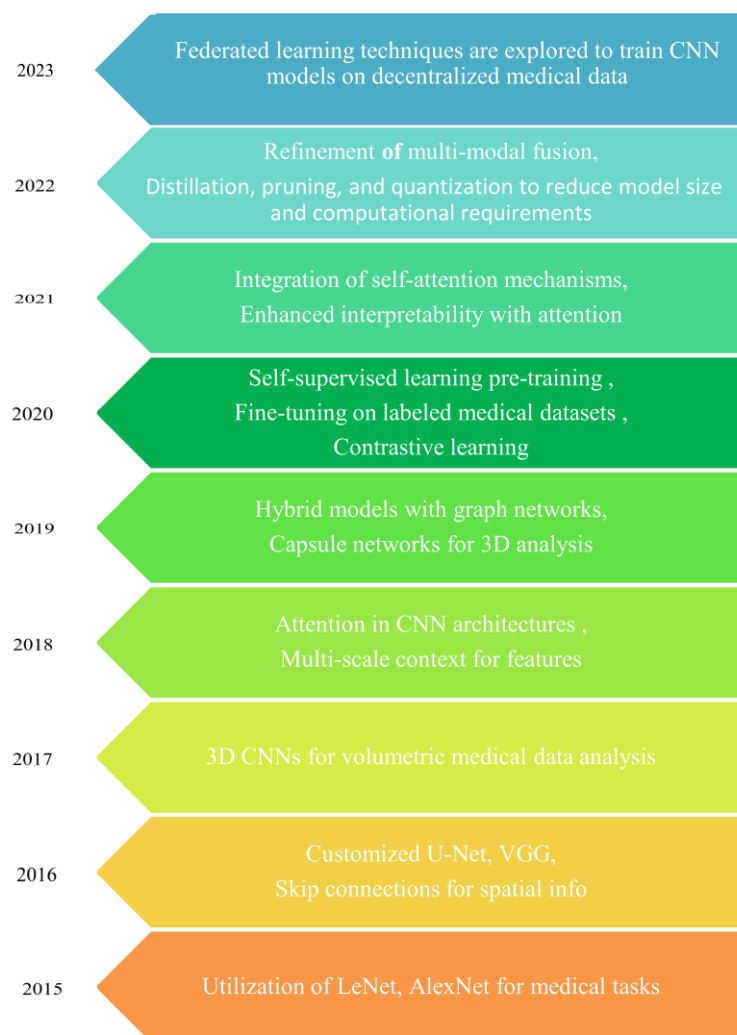


Figure 5. A timeline illustrating the development of techniques utilized by DCNN architectures and in medical imaging.

TL is currently the most widely utilized DL methodology. Training a CNN from scratch requires many labeled training samples and substantially more time and computational resources as compared to the already trained CNNs. Fine-tuning and freezing are the two main approaches [63] used in TL. Fine-tuning involves using the weights and biases of a pretrained CNN. The pretrained CNN layers are regarded as a fixed feature extractor in the freezing approach. The convolutional layer weights and biases are fixed in this case, but the fully connected layers are fine-tuned across the target dataset. Frozen layers can be any subgroup of convolutional or fully connected layers; however, the more superficial convolutional layers are usually frozen. If the training dataset is too small, an overfitting problem may occur during the training [64]. As a result, numerous research [65,66] addresses this issue by slicing 3D MRI volume into 2D slices, increasing the sample size of the original dataset and reducing the class imbalance issue. Additionally using morphological techniques such as rotation,

scaling, mirroring, translation, mirroring, and cropping [67] is another efficient technique for expanding the quantity and diversity of training data. This is known as data augmentation. Overfitting also occurs when the learning capacity of a network is so vast that it learns false characteristics rather than real patterns. This occurs when there is an abundance of information to learn. A validation dataset can be utilized throughout the training process to avoid overfitting and to achieve a steady potential of the tumor classification system on a new dataset that has not been observed in clinical practice.

Similar to TL, ensemble algorithm-based architectures [68] have gained prominence in the realm of DL due to their ability to enhance model performance and robustness. Ensembles combine the predictions of several individual models, often using techniques like bagging, boosting or stacking [69,70]. Bagging trains models on distinct data subsets, reducing overfitting risk. Boosting iteratively emphasizes weak learners, constructing a robust ensemble. Stacking combines diverse models' predictions via a meta-learner for intricate decision-making. These ensembles elevate DL model accuracy and generalization, especially in complex or data-scarce scenarios. Nonetheless, they demand additional computational resources and meticulous tuning. As with TL, selecting the best ensemble strategy hinges on the specific task, available resources, and managing overfitting using validation data. Figure 5 illustrates a chronological timeline depicting the utilization of different techniques by DCNN architectures and in medical imaging.

4.1. Performance metrics

Performance metrics are specific guidelines that give us scientific proof of the authenticity of a particular model. The metrics most used by multiple authors for classification in this study are outlined in Table 4, along with their respective functionalities.

Table 4. Performance metrics [71,72] used to evaluate studies in this review.

Metrics	Formula*
Accuracy (Ac)	$\frac{T_p + T_n}{T_p + F_n + F_p + T_n}$
Specificity (Sp)	$\frac{T_n}{T_n + F_p}$
Sensitivity (Sn)/Recall	$\frac{T_p}{T_p + F_n}$
Precision (Pr)	$\frac{T_p}{T_p + F_p}$
F1 Score (F1)	$\frac{2 * T_p}{2 * T_p + F_p + F_n}$
AUC	$\int_0^1 Sn(Sp)^{-1} d(Sp)$

* T_p = True Positive, T_n = True Negative, F_p = False Positive, F_n = False Negative, AUC = Area under the Curve

5. DL in glioma grading

The application of DCNNs to the classification of gliomas is an area of current investigation in the field of imaging science. To create a predictive model that can effectively categorize an image, a CNN may learn radiologic properties and their relative relevance with enough high-quality data [73]. The flowchart in Figure 6 provides an overview of a brain tumor diagnostic system, employing a generic DCNN. The process initiates with the collection involves collecting MRI scans of the brain. These scans are typically obtained from various sources, including hospitals, research institutions, and public datasets. Following data acquisition, the dataset is split into training and testing sets to facilitate model development and evaluation. Preprocessing steps are then employed to enhance the quality and utility of the MRI images. This includes normalization to ensure consistent intensity values across images and augmentation techniques to expand the dataset and improve model robustness. Additionally, preprocessing may involve cropping to focus on relevant brain regions and bias correction to mitigate inconsistencies in image acquisition. The subsequent stage involves model training, where different DCNN architectures are considered based on the specific task and data characteristics. Hyperparameter optimization is conducted to fine-tune the model parameters, such as learning rate, batch size, and number of epochs, aiming to maximize its ability to accurately classify brain tumors while minimizing errors and biases. This iterative process typically employs techniques like grid search or random search. Once the model is trained, it undergoes evaluation on the testing set using various performance metrics, including accuracy (Ac), specificity (Sp), sensitivity (Sn)/recall, precision (Pr), F1 score (F1), and area under the curve (AUC). These metrics provide insights into overall performance and its capability to correctly classify brain tumors across different classes. Throughout the process of configuring the model hyperparameters, the validation set gives an objective evaluation of a classification model on the training dataset.

Recent advancements in DL have significantly advanced the field of medical imaging, particularly in the areas of segmentation and classification. Researchers have been dedicated to improving the accuracy and efficiency of DL models for medical image segmentation. For example, Rehman et al. [74] introduced BU-Net, a modified U-Net architecture for brain tumor segmentation, which leverages residual extended skip and wide context to extract diverse features and enhance the valid receptive field. The researchers also employed a custom loss function to extract contextual information, resulting in improved segmentation performance. Addressing the challenge of information loss in deeper layers, Rehman et al. [75] proposed BrainSeg-Net, an encoder-decoder model that strategically shares pertinent details from shallow layers with deeper ones, enhancing tumor identification. Additionally, Rehman et al. [76] introduced a novel encoder-decoder architecture, RAAGR2-Net, which utilizes residual spatial pyramid pooling and attention gate modules to capture rich feature representations and retain local information, particularly in fine segmentation. Another study by Lin et al. [77] explored the integration of EfficientNetV2 as an encoder in combination with U-Net for brain tumor segmentation, significantly enhancing the model's performance. Furthermore, DL models have been successfully applied to tasks such as supraspinatus extraction from MRI, demonstrating high segmentation accuracy by Wang et al. [78]. Additionally, Yin et al. [79] proposed a double-branch flat bottom U-Net for efficient medical image segmentation, which achieved outstanding performance in the challenging task of pancreatic segmentation. These recent studies highlight the potential of DL models for medical image segmentation and underscore the importance of developing efficient and accurate models for clinical applications. Much of the ongoing research is confined to brain

segmentation, with only a limited amount of work done for tumor grading. Therefore, there is considerable potential to explore grade estimation for brain tumor DL approaches. In this section, we have discussed some of the existing DL-based glioma grading methods.

Recent studies have found that utilizing DCNNs to predict tumor grade and long-term survival is highly successful. Banerjee et al. [80] investigated the feasibility of using DL-based techniques to grade gliomas from MRIs. They used VGGNet and ResNet architectures to assess the appropriateness of transfer learning, achieving an accuracy of 84 and 90%, respectively. The study by Muneer et al. [81] contrasts the glioma classification performance of two DL systems, WNDCHRM (weighted neighbor distance using compound hierarchy of algorithms representing morphology) and VGG-19 DCNN. It was observed that VGG-19 CNN achieved higher accuracy than WNDCHRM. Ge et al. [82] proposed a glioma classification multistream CNN and fusion network. T1Wc, T2W, and FLAIR images were extracted from the BraTS 2017 dataset and put in their own CNN. The collected information was then combined with the extracted features. They were able to achieve a precision of 90.87% by using three distinct data points. Individually, the T1Wc images were the best at distinguishing between HGGs and LGGs. In another study, Yang et al. [83] investigated AlexNet and GoogLeNet's ability to distinguish between LGGs and HGGs. They compared the accuracy of these two CNNs when trained from scratch versus pretrained CNNs with fine-tuning using T1Wc images from glioma patients. According to the results, pretrained CNNs outperform untrained CNNs, with GoogleNet outperforming AlexNet. Gutta et al. [84] built a DCNN model and compared it to ML models trained solely on traditional radiomic data. With an accuracy of 87%, the proposed DCNN model significantly outperforms ML models.

Lu et al. [85] classified gliomas using the ResNet model. Pyramid dilated convolution is added to ResNet to increase classification performance. The proposed method achieves 80.1% accuracy; however, this method can only interpret 2D MRI. Also, manual labeling of the training set was required. Mzoughi et al. [86] proposed a fully automatic 3D CNN architecture with a T1Wc sequence to distinguish between LGGs and HGGs. The accuracy of this 3D-CNN model was 96.49%. Zhuge et al. [87] used conventional MRI to compare 3DConvNet and 2D Mask R-CNN (region-based CNNs) for glioma classification. The results showed that the 3DConvNet outperformed the 2D Mask R-CNN, with a test accuracy of 97.1% versus 96.3%. Khawaldeh et al. [88] utilizes a modified version of AlexNet. The 12-layer ConvNet model proposed in this research study comprises convolutional, subsampling, dense, and fully connected layers. Overall accuracy achieved by this model is 91.16 % on FLAIR MRIs. Chenjie et al. [89] proposed an MRI-based multimodality glioma classification system. To make use of unlabeled data, the authors used deep semi-supervised learning. Generative adversarial networks generated synthetic MRIs to mitigate overfitting in the intermediate dataset. Using CNN, the suggested system achieved a test accuracy of 86.53% on the TCGA (the cancer genome atlas program) dataset and 90.70% on the BraTS dataset. Liang et al. [90] proposed the more advanced DenseNet to predict IDH mutations. Their approach was also used to grade gliomas, with a 91.4% accuracy. As a result, its potential application can be extended to additional multimodal radiogenomics challenges. Some recent applications of DCNN-based methods for automated glioma grading research are summarized in Table 5.

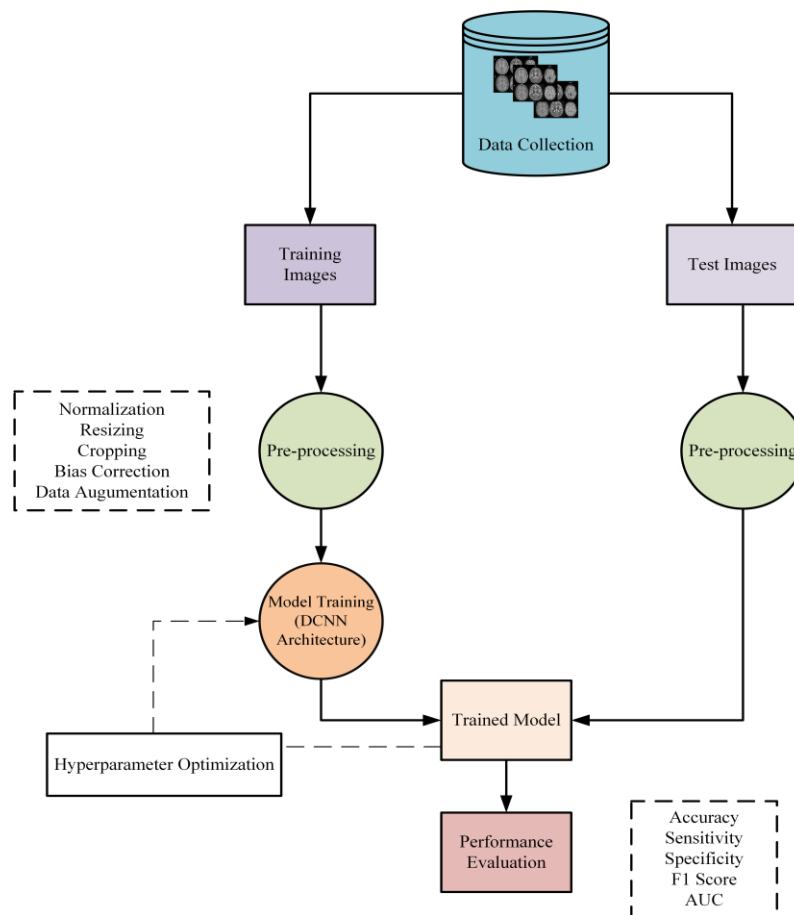


Figure 6. A flowchart of a generic DCNN-based brain tumors diagnosing system.

Table 5. An overview of DCNN-based high-accuracy glioma grading studies featuring dataset specifics, class labels, sample sizes, MRI modalities, architectural methods, validation techniques, performance metrics, and study limitations.

Study	Dataset	Class Label	Sample Size	MRI Modalities	Architecture	Validation	Performance	Limitations
[91]	BraTS 2017	LGG HGG	LGG: 75 HGG: 210	T1Wc	3D Multiscale CNN	unspecified	Ac = 89.47	Limited literature available to validate the results.
[92]	Radiopaedia	Grade I Grade II Grade III Grade IV	Grade I: 1080 Grade II: 960 Grade III: 750 Grade IV: 840	T1Wc	Pretrained VGG-19	unspecified	Ac = 90.67 Grade I = 95.54 Grade II = 92.66 Grade III = 87.77 Grade IV = 86.71	Without data augmentation accuracy is very low compared to the literature.

continued on next page

Study	Dataset	Class Label	Sample Size	MRI Modalities	Architecture	Validation	Performance	Limitations
[83]	ClinicalTrials.org	LGG HGG	LGG: 52 HGG: 61	T1Wc	Pretrained AlexNet	5-fold	Ac = 92.7 Ac = 94.7	No automated tumor extraction before classification.
[93]	Private	LGG HGG	LGG: 50 HGG: 54	T2W FLAIR	Modified DCNN	5-fold	Ac = 97.1 Sn = 98.0 Sp = 96.3 Pr = 96.1 F1 = 97.0	Difficult to classify glioma with heterogenous lesions containing cystic morphological formation.
[63]	BraTS 2019	LGG HGG	LGG: 76 HGG: 259	T1W T2W T1Wc	Pretrained AlexNet	unspecified	AUC = 82	Maynot be generalized to other medical imaging dataset.
[89]	BraTS 2017	LGG HGG	LGG: 75 HGG: 210	T1W T2W T1Wc FLAIR	Graph based semi-supervised learning	unspecified	Ac = 90.70 Sn = 84.35 Sp = 93.01	Imbalance of training data between two classes affected average test performance.
[86]	BraTS 2018	LGG HGG	LGG: 75 HGG: 209	T1Wc	Multiscale 3D CNN	unspecified	Ac = 96.49	Not enough state of the art literature for CNN available to validate the result.
[87]	TCIA BraTS 2018	LGG HGG	LGG: 108 LGG: 75 HGG: 210	T1W T2W T1Wc FLAIR	3D ConvNet	5-fold	Ac = 97.1 Sn = 94.7 Sp = 96.8	GPU limitation
[94]	BraTS 2019	LGG HGG	LGG: 75 HGG: 210	T1W T2W T1Wc FLAIR T1-GD	7 stacked pretrained CNN	10-fold	Ac = 98.06 Sn = 98.64 Sp = 98.67 Pr = 98.67 F1 = 98.62	Dataset heterogeneity.
[95]	TCIA Private	LGG GBM LGG GBM	LGG:121 GBM:16 4 LGG:49 GBM:91	T1Wc T2W FLAIR T1Wc T2W FLAIR	3D UNet CNN	unspecified	Ac = 90.0 Sn = 93.48 Sp = 87.04 Ac = 90.0 Sn = 90.16 Sp = 89.80	False positive of the model in case IDH mutant astrocytoma. Small dataset.

continued on next page

Study	Dataset	Class Label	Sample Size	MRI Modalities	Architecture	Validation	Performance	Limitations
[96]	TCIA	Grade II Grade III Grade IV	Grade II: 30 Grade III: 43 Grade IV: 57	T1Wc	Pretrained DCNN	10-fold	Ac = 97.9 AUC = 99.9	Use of data augmentation with transfer learning is questionable.
[81]	Private	Grade I Grade II Grade III Grade IV	Grade I: 130 Grade II: 269 Grade III: 103 Grade IV: 155	T2W	Pretrained Vgg-19	5-fold	Ac = 98.25	Smaller dataset was used. No data augmentation was done to improve data size.
[97]	TCIA[REM BRANDT]	LGG HGG	LGG: 484 HGG: 631	T2W	MajVot	5-fold	Ac = 98.43	Classification was performed without segmentation.
[98]	TCIA	LGG- GRADE I LGG- GRADE II Unknown GRADE	LGG- GRADE I: 50 LGG- GRADE II: 58 Unknown GRADE: 2	T1W T1Wc FLAIR	Pretrained VGG-19	5-fold	Ac = 95.0 Sn = 93.0 Sp = 98.0	No independent dataset available for testing.
[88]	TCIA[REM BRANDT]	LGG HGG Healthy Subjects	LGG: 41 HGG: 67 Healthy Subjects:22	FLAIR	12layer AlexNet	unspecified	Ac = 91.1 Pr = 91.79 Re = 92.25 F1 = 92.05	Ground truth was only provided for 126 subjects.
[99]	Radiopaedia	Grade I Grade II Grade III Grade IV	Grade I: 36 Grade II: 32 Grade III: 25 Grade IV: 28	Unspecified	DCNN	5-fold	Ac = 93.71 Grade I = 96.32 Grade II = 95.31 Grade III = 96.81 Grade IV = 99.61	Smaller dataset

continued on next page

Study	Dataset	Class Label	Sample Size	MRI Modalities	Architecture	Validation	Performance	Limitations
[100]	BraTS 2019	LGG HGG	LGG: 259 HGG: 76	T1CE, T2, FLAIR	EfficientNetB 0	unspecified	Ac = 98.8	Due to hardware limitations each image is processed to 50 epochs only.
[101]	TCIA	LGG GBM	LGG:159 GBM: 163	T2W FLAIR	Fusion of ResNet/18/50/ 101/152 using DST Dempster- shafer theory	5-fold	Ac = 95.87 Pr = 89.12 Re = 95.12 F1 = 91.91	Smaller dataset.
[102]	REMBRA NDT	Grade I Grade II Grade III Grade IV	Grade I: 2 Grade II: 110, Grade III: 93 Grade IV: 140	T1W T2W FLAIR	DCNN	unspecified	Ac = 98.91	No data augmentation was done to improve data size.
[103]	TCIA (REMBRA NDT)	LGG HGG	LGG: 44 HGG: 68	T1W T2W FLAIR	MajVot (AlexNet, VGG16, ResNet18, GoogleNet, ResNet50)	5-fold	Ac = 97.21	Very small dataset from a single institution.
[104]	BratTS 2020	LGG HGG	LGG :76 HGG: 293	T1W T2W T1Wc FLAIR	DenseNet121	monte carlo	Ac = 97	Dataset heterogeneity

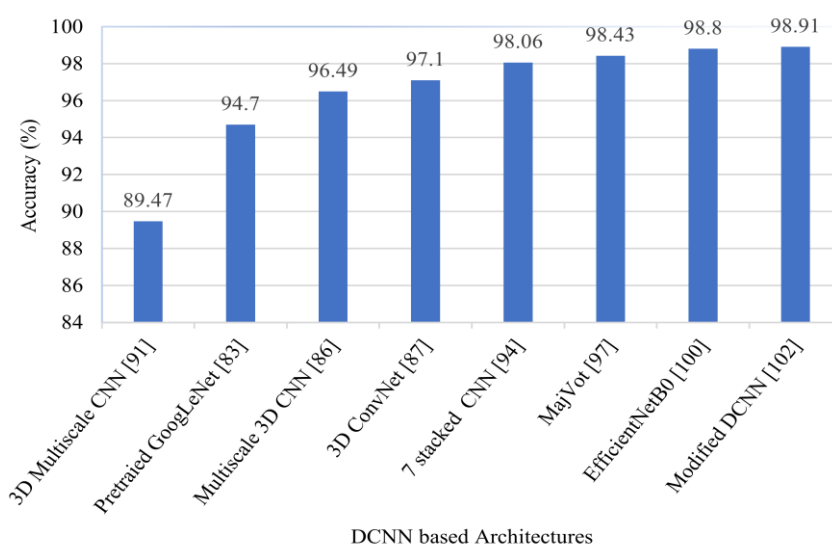


Figure 7. Comparative analysis of glioma grading accuracy across different DCNNs.

Figure 7 presents a comparative evaluation of the efficacy of various DCNN architectures in the task of glioma grading. The graph highlights the high level of accuracy attained by several DCNNs using TL, underlining their expertise in grading gliomas. Moreover, EfficientNetB0 [100] also proves to be a robust performer with an accuracy rate of 98.8%, signifying its capability in managing this intricate task. The ensemble algorithm, which employs majority voting (MajVot), demonstrates remarkable accuracy [97,103]. This superior performance can be ascribed to the ensemble approach capacity to utilize the strengths of multiple models, thereby augmenting the overall predictive ability. The custom-built DCNN [102] demonstrates the highest accuracy of 98.91% in the task of glioma grading.

These findings indicate that DL models, particularly those that incorporate ensemble techniques and custom architectures, exhibit substantial potential in improving the accuracy and reliability of glioma grading. This enhancement is pivotal as it directly influences clinical decision-making and patient care. By delivering more precise grading, these models can support clinicians in formulating more effective treatment strategies, ultimately leading to better patient outcomes. However, it is important to note that while these DCNNs shows promising results, further validation and testing on broader datasets are necessary to confirm its effectiveness and generalizability in real-world clinical settings.

5.1. Computing environment

The utilization of CNN classifiers in complex image classification tasks such as grading brain tumors offers distinctive advantages. CNNs autonomously extract relevant features, eliminating the need for separate feature extraction and classification steps. Despite their compact architecture, CNNs excel in intricate classification, although they entail higher computing complexity compared to traditional methods like SVM or logistic regression. In the realm of high-level programming environments, Python and MATLAB are prominent choices for DL implementation due to user-

friendliness. Two primary approaches for evaluating ML models are development-based and production-based. Python, particularly when used with platforms like Google Colab, has an advantage over MATLAB due to its faster training times, made possible by accessible GPUs (graphics processing units) and cloud-based storage. However, the longer training times in MATLAB can be offset by a powerful workstation. The performance of glioma grading algorithms is significantly influenced by computing power. Adequate memory is essential for loading and preprocessing large medical imaging datasets, which include intermediate activations and gradients. GPUs, which perform essential matrix operations in parallel, speed up the training process. Workstations equipped with substantial memory and multiple GPUs can hasten the training, tuning, and evaluation processes. Factors such as model complexity, dataset size, hardware, batch size, and optimization techniques all influence training time. Despite their parallel processing capabilities, GPUs often necessitate manual synchronization in frameworks like OpenMP (open multi-processing) and CUDA (compute unified device architecture). Nevertheless, the use of GPUs for parallel algorithms holds potential for efficient big data processing, especially in high-performance computing applications such as cancer research and AI [105]. In terms of GPU vs CPU (central processing unit) performance, it has been observed that GPUs are generally faster than CPUs. However, for smaller networks with only two hidden layers, CPUs can be faster than GPUs if there are less than 1000 neurons in each hidden layer. This highlights the importance of considering the specific requirements and characteristics of the model when choosing between GPU and CPU for prediction.

6. Implementation challenges

Brain tumors remain a popular research topic in medical image processing. Advanced glioma classification techniques in HGGs and LGGs are constantly evolving. For such problems, DL has emerged as a critical research tool for improving the performance of standard ML approaches. DL facilitates multiple levels of representation and abstraction, thereby providing more comprehensive information about MRIs and their attributes [98]. This research focused on the DCNN-based glioma classification architectures. Table 5 summarizes the findings of several studies that show that DCNN-based architectures can handle a wide range of glioma classification tasks effectively and efficiently. It demonstrates that TL using DCNN models such as ResNet, VGG, and GoogleNet outperforms other models developed from scratch. However, some challenges must be resolved before DL can be used in oncology as, shown in Figure 8.

The lack of an objective dataset was one of the most common issues identified in this study. DCNNs are based on supervised learning techniques, requiring a large volume of labeled data to learn properly. The problem with small datasets is that the DL algorithm may produce absurdly inflated algorithm accuracy due to the millions of parameters that must be overfit to a single specialized training population [83]. This issue is critical given the scarcity of curated datasets, particularly in radiography research. In addition to collecting large, heterogeneous datasets, various methods have been developed to solve this issue, such as the addition of feature dropout, L2 regularization and batch normalization [106,107].

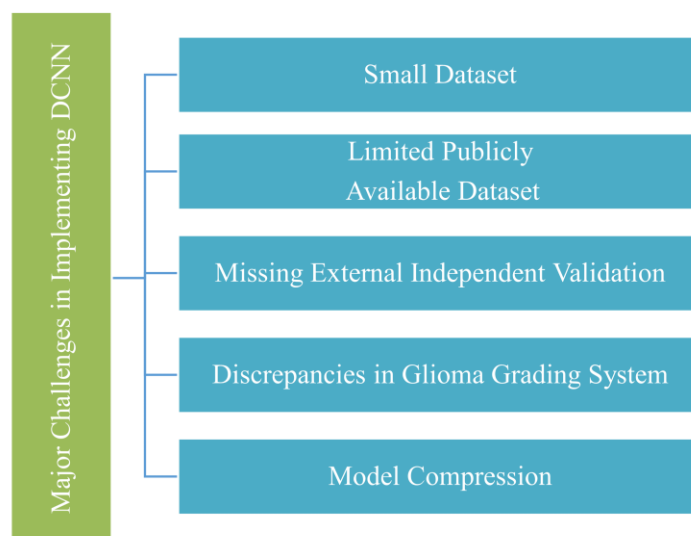


Figure 8. Major challenges in adopting DCNN-based classification algorithms for glioma grading in clinical settings.

This review also revealed a startling gap in precision among certain researchers when defining the dataset, tumor type, and the accuracy, sensitivity, and specificity performance measures of the algorithm. In addition, research based on meticulously managed datasets, such as BraTS or TCIA (the cancer imaging archive), demonstrated algorithms trained without external validation that might not produce reproducible findings in clinical practice despite their consistently high accuracy. Most publications did not do validation, which is the most significant problem with ML and DL that should be considered. In some cases, only cross-validation was done. Validation is essential in accordance with the standards for constructing and reporting ML/DL prediction models in biomedical research [108].

Although this analysis highlighted several contributions that independently concentrate on the three primary phases of tumor identification, it did not identify any diagnostic approaches that encompass all the phases. The absence of a comprehensive diagnostic system in a single package presents two issues: the lack of a fully automated procedure and the lack of integration between the three processes. The development of a complete and automated system should facilitate the process of diagnosing brain tumors for physicians and radiologists, as well as translating research-based diagnostic algorithms into clinical practice. Additionally, uniform criteria of glioma grade should be utilized when developing DL models. It is interesting to note that there were discrepancies between the LGGs and HGGs definitions, with some research identifying Grade III gliomas as HGGs and others as LGGs. Lack of a standard classification strategy may hinder the performance on independent datasets given that the images used for segmentation, feature extraction, and model training/testing are labelled as LGGs or HGGs based on nonuniform criteria. As glioma grade influences clinical therapy, it is vital that the outputs of LGG and HGG algorithms reflect a universal definition congruent with current WHO standards. Another key observation regarding the DL architecture is that in the current context, GPU-based systems with a lot of memory are essential since DL models need a lot of data [37], which is linked to millions and trillions of parameters [71]. Also, to enable practical deployment of

well-trained DL models, addressing their extensive memory and computational demands is essential. Particularly in data-intensive domains like healthcare and environmental science, these requirements limit their usage in resource-constrained settings, hindering healthcare applications due to escalating data complexity. Solutions like FPGA (field programmable gate arrays) and GPU hardware accelerators have emerged, while techniques like parameter pruning, knowledge distillation, compact filters, and low-rank factorization offer model compression strategies to mitigate computational challenges [109]. Table 6 summarizes the challenges encountered in DL based research related to glioma grading and their influence on algorithm performance.

Table 6. Challenges in DL-based research on glioma grading and impact on algorithm performance.

Challenge	Description
Lack of Objective Dataset	DCNNs require large labelled datasets for supervised learning, which is problematic when dealing with limited data in radiography research.
Precision Gap in Definitions	Variation in dataset definitions, tumor types, and evaluation metrics leads to inconsistent algorithm performance assessment.
Absence of Comprehensive Diagnostic Approaches	The deep models lack in their ability of interoperability and automation. Very few architectures are available till date that are fully automated and can adapt to model changes
Uniform Criteria for Glioma Grade	Inconsistent classification of LGGs and HGGs across datasets due to non-standardized definitions. Model might fail on independent dataset due to grade inconsistency.
Imbalanced Datasets	Accuracy favoured majority class, leading to higher misclassification rates for minority classes.
Model Compression	DL model complexity demands intensive memory and computation, challenging deployment on limited computational-power machines, especially in healthcare.
Lack of External Validation	Many studies lack external validation, leading to limited generalizability of findings in real-world clinical practice. Algorithm might achieve high accuracy on in-house dataset, but fails in clinical setting due to unverified generalizability.
Standard Pre-processing Technique	Pre-processing is required to make data clean from every type of noise and more acceptable for the required task at hand. Lack of standardized pre-processing affects data quality; software choices degrade image quality.
Gradient Explode and Gradient Vanishing	Deep architectures, aimed at achieving high accuracy, encounter challenges such as gradient vanishing, where the error that needs to be propagated diminishes, and gradient explosion, caused by suboptimal optimizer selection.

7. Limitations

The study presented here offers valuable insights into the performance of DL algorithms in terms of glioma grading. However, it is important to acknowledge certain limitations that might affect the generalizability and robustness of the findings. This study's limitations include the possibility of missing recent and unpublished works due to the timing and criteria of the search, potentially affecting the comprehensiveness of the findings. Additionally, the focus on accuracy as the primary performance metric resulted in the exclusion of studies lacking accuracy results, limiting the overall assessment.

Furthermore, the presence of heterogeneity, inconsistent definitions, evolving standards, high variability across studies, and the absence of confidence intervals in the reviewed literature hindered the aggregation of results, introducing uncertainties in the study's conclusions. These challenges are compounded by the sensitivity of DCNNs to subtle variations in medical images, stemming from factors such as patient anatomy, acquisition conditions, and disease presentation. While the human eye can adapt to such nuances effortlessly, DCNNs may struggle, resulting in misdiagnoses or missed diagnoses. This highlights the necessity for evaluation metrics that can capture the model's ability to manage these complexities. The articles reviewed in our study primarily concentrate on conventional evaluation metrics such as accuracy, sensitivity, and specificity. These metrics provide a comprehensive evaluation of model performance across the entire dataset. However, they may not adequately highlight local discrepancies, potentially leading to misleading interpretations of model performance. A recent survey [110] reiterates these concerns, emphasizing the crucial role of model uncertainty and interpretability in building confidence in medical diagnoses. To overcome the limitations of traditional evaluation metrics, future research should also concentrate on local discrepancy analysis using localization metrics like intersection over union or dice similarity coefficient. These metrics measure the spatial overlap between predicted and actual regions. The incorporation of region-based evaluation techniques, such as precision-recall curves or localization error analysis, can offer a more nuanced understanding of model performance. Techniques like gradient-weighted class activation mapping (Grad-CAM) or attention mechanisms can provide visual explanations of the model's decision-making process, assisting in comprehending the model's behavior and identifying potential regions prone to errors. Additionally, task-specific metrics, customized for specific clinical tasks, can offer more pertinent insights than generic metrics, steering development towards clinically relevant applications.

While our current work leverages traditional CNNs, we acknowledge the potential of fuzzy logic to address uncertainty challenges raised in glioma diagnosis, an area that has been scarcely explored. Medical images inherently contain ambiguity due to imaging artifacts, partial volume effects, and inter-observer variability that can lead to misclassifications. Traditional performance metrics like accuracy may not adequately capture these nuances. Fuzzy logic, capable of handling ambiguity and incorporating expert knowledge, offers a promising alternative. A recent study [111] presents an intriguing application of fuzzy logic to address uncertainties in evaluating external loads on steel structures. This method, based on divergence computations, achieves better classification, and reduces ambiguity compared to traditional approaches. Similarly, rather than regarding ambiguous features as binary certainties, a DCNN empowered by fuzzy logic could interpret them as possibilities with varying degrees of truth. This could lead to more nuanced and robust classifications, particularly in cases with subtle variations or overlapping tumor regions. We believe exploring this integration holds immense potential for advancing glioma diagnosis.

8. Conclusions

The ability for medic specialists to categorize brain tumor scans quickly and accurately has never been more crucial. Recently CNNs have accomplished astonishing achievements in categorizing brain tumors such as gliomas. This study examined the most recent DCNN-based glioma classification architecture, datasets, and the efficacy of each suggested model for brain MRIs over the period from 2015 to 2023. Table 5 shows a compilation of pertinent data, applied approaches, DL networks,

and their performance. The research findings highlight the potential of DCNN architectures, particularly hybrid and ensemble DCNNs, which have achieved accuracy levels as high as 98.9%. These results underscore the considerable potential of advanced DL models in augmenting the accuracy and reliability of glioma grading. However, despite the undeniable successes of DCNNs, challenges remain in incorporating them into clinical practice. The study also found that preprocessing and segmentation were not always used in the surveyed articles before categorization. No single system can do all the functions automatically and with high precision.

While there is ongoing work to enhance the utility of DL in tumor identification and classification, the need for standardized databases for these purposes remains evident. The varied use of databases and benchmarks by many researchers underscores the need for standardization. Additionally, the black-box nature of DCNNs has constrained their application beyond research contexts. DL holds great promise for the future of brain tumor research. By focusing on the right strategies, these studies could transition from research labs to clinical settings. These methods could also be applied to the classification of other brain disorders, including alzheimer's, parkinson's, stroke, and autism. We hope that our review will guide researchers toward potential future directions for efficient grading techniques.

9. Future directions

Addressing the implementation challenges of DCNNs in radiography research requires a strategic approach. To mitigate the impact of limited datasets, collaborative efforts should focus on creating objective, diverse datasets, potentially incorporating data augmentation techniques. Establishing standardized definitions and evaluation metrics across tumor types would enhance algorithm assessment consistency. The development of a unified diagnostic framework spanning tumor identification phases holds promise for increased automation and integration. Overcoming the hurdle of inconsistent glioma grading could involve adopting universally accepted grading criteria. Additionally, to ensure broader applicability, it is essential to explore hardware-efficient solutions, such as model compression techniques, thereby ensuring accessibility to necessary resources. External validation is also very crucial for real-world utility; thus, incorporating rigorous external validation protocols in research design would enhance clinical relevance. As research progresses, accounting for these future directions will refine the robustness and practicality of DCNN implementation in radiography, ultimately benefiting patient care and diagnosis quality.

We propose the following novel directions for shaping forthcoming models:

- For robust clinical applicability, future investigations must embrace expansive multicenter datasets, gauging model efficacy across diverse populations independently.
- Elevating CNN performance hinges on meticulous hyperparameter selection, underscoring their pivotal role. In future designs, adept optimization techniques must be employed to navigate this critical aspect.
- Standardizing imaging methodologies [112] is still a crucial problem to solve since even the best CNNs may prove ineffective when tested on real-life data. This involves ensuring consistency across institutions and modalities, accounting for real-world variability, and enhancing robustness to achieve effective performance on diverse clinical data.
- Incorporating explainability in AI models is essential for improving the trust and understanding of AI software. Future models should aim to provide clear, understandable reasoning for their predictions and decisions. This will not only enhance user trust but also facilitate troubleshooting

and refinement of the models.

- In a dynamic landscape, the WHO has refined glioma classification, transitioning to molecular insights from conventional histopathology in 2016, further accentuated in 2021 by emphasis of cIMPACT-NOW (the consortium to inform molecular and practical approaches to CNS tumor taxonomy) on molecular markers. This evolving scenario introduces flux in defining LGGs and HGGs, impeding inter-comparison of ML/DL models anchored on differing grading criteria. To enhance accuracy and coherence, future research should converge on glioma grading standards alongside molecular subtypes, assuring enduring and accurate prognostications.
- Developing precise data augmentation methods to expand and diversify training datasets for improved model performance.
- Investigate integrating divergence-based fuzzy logic into existing DCNN architectures for glioma grading to improve classification robustness and address inherent image uncertainty.
- Beyond these technical aspects, it is also important to address the clinical issues regarding the adoption of DCNNs for tumor grading. It is important to consider factors such as the interpretability of the model's predictions, the integration of the model into existing workflows, and the training and support provided to healthcare professionals using the technology. Additionally, ethical considerations, such as patient consent and data privacy, must also be addressed. These factors are all critical for the successful adoption of AI technologies in clinical settings.

These novel directions, coupled with the previously outlined strategies, underscore the evolving landscape of radiography research, and hold significant potential for advancing both diagnostic accuracy and patient care.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Conflict of interest

The authors declare there is no conflict of interest.

References

1. M. L. Goodenberger, R. B. Jenkins, Genetics of adult glioma, *Cancer Genet.*, **205** (2012), 613–621. <https://doi.org/10.1016/j.cancergen.2012.10.009>
2. D. N. Louis, A. Perry, G. Reifenberger, A. Deimling, D. Figarella-Branger, W. K. Cavenee, et al., The 2016 World Health Organization classification of tumors of the central nervous system: a summary, *Acta Neuropathol.*, **131** (2016), 803–820. <https://doi.org/10.1007/s00401-016-1545-1>
3. D. N. Louis, A. Perry, P. Wesseling, D. J. Brat, I. A. Cree, D. Figarella-Branger, et al., The 2021 WHO classification of tumors of the central nervous system: a summary, *Neuro-oncology*, **23** (2021), 1231–1251. <https://doi.org/10.1093/neuonc/noab106>
4. A. Munshi, Central nervous system tumors: Spotlight on India, *South Asian J. Cancer*, **5** (2016) 146–147. <https://doi.org/10.4103/2278-330x.187588>

5. F. Zaccagna, J. T. Grist, N. Quartuccio, F. Riemer, F. Fraioli, C. Caracò, Imaging and treatment of brain tumors through molecular targeting: Recent clinical advances, *Eur. J. Radiol.*, **142** (2021), 109842. <https://doi.org/10.1016/j.ejrad.2021.109842>
6. D. Aquino, A. Gioppo, G. Finocchiaro, M. G. Bruzzone, V. Cuccarini, MRI in glioma immunotherapy: evidence, pitfalls, and perspectives, *J. Immunol. Res.*, **2017** (2017), 5813951. <https://doi.org/10.1155/2017/5813951>
7. A. Maier, C. Syben, T. Lasser, C. Riess, A gentle introduction to deep learning in medical image processing, *Z. fur Med. Phys.*, **29** (2019), 86–101. <https://doi.org/10.1016/j.zemedi.2018.12.003>
8. K. Yasaka, H. Akai, A. Kunimatsu, S. Kiryu, O. Abe, Deep learning with convolutional neural network in radiology, *Jpn. J. Radiol.*, **36** (2018), 257–272. <https://doi.org/10.1007/s11604-018-0726-3>
9. M. I. Razzak, S. Naz, A. Zaib, Deep learning for medical image processing: Overview, challenges and the future, *Classif. BioApps: Autom. Decis. Making*, **26** (2018), 323–350. https://doi.org/10.1007/978-3-319-65981-7_12
10. N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, et al., Convolutional neural networks for medical image analysis: Full training or fine tuning?, *IEEE Trans. Med. Imaging*, **35** (2016), 1299–1312. <https://doi.org/10.1109/tmi.2016.2535302>
11. W. Jin, M. Fatehi, K. Abhishek, M. Mallya, B. Toyota, G. Hamarneh, Artificial intelligence in glioma imaging: Challenges and advances, *J. Neural Eng.*, **17** (2020), 021002. <https://doi.org/10.1088/1741-2552/ab8131>
12. J. H. Park, N. Jung, S. J. Kang, H. S. Kim, E. Kim, H. J. Lee, et al., Survival and prognosis of patients with pilocytic astrocytoma: a single-center study, *Brain Tumor Res. Treat.*, **7** (2019), 92–97. <https://doi.org/10.14791/btrt.2019.7>
13. S. G. Berntsson, R. T. Merrell, E. S. Amirian, G. N. Armstrong, D. Lachance, A. Smits, et al., Glioma-related seizures in relation to histopathological subtypes: A report from the glioma international case-control study, *J. Neurol.*, **265** (2018), 1432–1442. <https://doi.org/10.1007/s00415-018-8857-0>
14. W. Taal, J. E. Bromberg, M. J. V. den Bent, Chemotherapy in glioma, *CNS Oncol.*, **4** (2015), 179–192. <https://doi.org/10.2217/cns.15.2>
15. M. Glas, C. Happold, J. Rieger, D. Wiewrodt, O. Bähr, J. P. Steinbach, et al., Long-term survival of patients with glioblastoma treated with radiotherapy and lomustine plus temozolomide, *J. Clin. Oncol.*, **27** (2009), 1257–1261. <https://doi.org/10.1200/jco.2008.19.2195>
16. N. Wijethilake, D. Meedeniya, C. Chitraranjan, I. Perera, M. Islam, H. Ren, Glioma survival analysis empowered with data engineering—a survey, *IEEE Access*, **9** (2021), 43168–43191. <https://doi.org/10.1109/access.2021.3065965>
17. Q. T. Ostrom, L. Bauchet, F. G. Davis, I. Deltour, J. L. Fisher, C. E. Langer, et al., The epidemiology of glioma in adults: A “state of the science” review, *Neuro-Oncology*, **16** (2014), 896–913. <https://doi.org/10.1093/neuonc/nou087>
18. D. Salles, G. Laviola, A. C. de M. Malinverni, J. N. Stávale, Pilocytic astrocytoma: A review of general, clinical, and molecular characteristics, *J. Child Neurol.*, **35** (2020), 852–858. <https://doi.org/10.1177/0883073820937225>
19. S. Florian. S. Şuşman, Diffuse astrocytoma and oligodendroglioma: An integrated diagnosis and management, in *Glioma-Contemporary Diagnostic and Therapeutic*, IntechOpen, 2019. <https://doi.org/10.5772/intechopen.76205>

20. C. Balañá, M. Alonso, A. Hernandez, P. Perez-Segura, E. Pineda, A. Ramos, et al., SEOM clinical guidelines for anaplastic gliomas, *Clin. Transl. Oncol.*, **20** (2017), 16–21. <https://doi.org/10.1007/s12094-017-1762-7>
21. O. G. Taylor, J. S. Brzozowski, K. A. Skelding, Glioblastoma multiforme: An overview of emerging therapeutic targets, *Front. Oncol.*, **9** (2019) 963. <https://doi.org/10.3389/fonc.2019.00963>
22. M. K. Abd-Ellah, A. I. Awad, A. A. Khalaf, H. F. Hamed, A review on brain tumor diagnosis from MRI images: Practical implications, key achievements, and lessons learned, *Magn. Reson. Imaging*, **61** (2019), 300–318. <https://doi.org/10.1016/j.mri.2019.05.028>
23. A. Lasocki, A. Tsui, M. A. Tacey, K. J. Drummond, K. M. Field, F. Gaillard, et al., MRI grading versus histology: Predicting survival of World Health Organization Grade II-IV astrocytomas, *AJNR Am. J. Neuroradiol.*, **36** (2015), 77–83. <https://doi.org/10.3174/ajnr.a4077>
24. G. Mohan, M. Subashini, MRI based medical image analysis: Survey on brain tumor grade classification, *Biomed. Signal Process. Control.*, **39** (2018), 139–161. <https://doi.org/10.1016/j.bspc.2017.07.007>
25. A. Vamvakas, S. C. Williams, K. Theodorou, E. Kapsalaki, K. Fountas, C. Kappas, et al., Imaging biomarker analysis of advanced multiparametric MRI for glioma grading, *Phys. Medica*, **60** (2019), 188–198. <https://doi.org/10.1016/j.ejmp.2019.03.014>
26. M. Rizwan, A. Shabbir, A. R. Javed, M. Shabbir, T. Baker, D. A-J. Obe, Brain tumor and glioma grade classification using Gaussian convolutional neural network, *IEEE Access*, **10** (2022), 29731–29740. <https://doi.org/10.1109/access.2022.3153108>
27. J. E. Villanueva-Meyer, M. C. Mabray, S. Cha, Current clinical brain tumor imaging, *Neurosurgery*, **81** (2017), 397–415. <https://doi.org/10.1093/neuros/nyx103>
28. CIP, The Cancer Genome Atlas (TCGA), 2012. Available from: <http://cancergenome.nih.gov>.
29. *The Cancer Genome Atlas Glioblastoma Multiforme Collection*, <https://www.cancerimagingarchive.net/collection/tcga-gbm/>
30. L. Scarpace, T. Mikkelsen, S. Cha, S. Rao, S. Tekchandani, D. Gutman, et al., The Cancer Genome Atlas Glioblastoma Multiforme Collection (TCGA-GBM) (Version 5) [Data set], *Cancer Imaging Arch.*, 2016. <https://doi.org/10.7937/K9/TCIA.2016.RNYFU9>
31. *The Cancer Genome Atlas Glioblastoma Multiforme Collection*, <https://wiki.cancerimagingarchive.net/display/Public/TCGA-LGG>
32. N. Pedano, A. E. Flanders, L. Scarpace, T. Mikkelsen, J. M. Eschbacher, B. Hermes, et al., The Cancer Genome Atlas Low Grade Glioma Collection (TCGA-LGG) (Version 3) [Dataset], *Cancer Imaging Arch.*, 2016. <https://doi.org/10.7937/K9/TCIA.2016.L4LTD3TK>
33. K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel, et al., The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository, *J. Digit. Imaging*, **26** (2013), 1045–1057. <https://doi.org/10.1007/s10278-013-9622-7>.
34. *REMBRANDT- The Cancer Imaging Archive (TCIA) Public Access*, <https://wiki.cancerimagingarchive.net/display/Public/REMBRANDT>.
35. L. Scarpace, A. E. Flanders, R. Jain, T. Mikkelsen, D. W. Andrews, Data from REMBRANDT [Data set], *Cancer Imaging Arch.*, 2019. <https://doi.org/10.7937/K9/TCIA.2015.588OZUZB>
36. *MICCAI BRATS- The Multimodal Brain Tumor Segmentation*, <http://braintumorsegmentation.org/>

37. S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. S. Kirby, et al., Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features, *Sci. Data*, **4** (2017). <https://doi.org/10.1038/sdata.2017.117>
38. B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, et al., The multimodal Brain Tumor Image Segmentation Benchmark (BRATS), *IEEE Trans. Med. Imaging*, **34** (2015), 1993–2024. <https://doi.org/10.1109/tmi.2014.2377694>
39. *ClinicalTrials.gov.*, <https://clinicaltrials.gov/>, Accessed: Apr. 26, 2022.
40. *Radiopaedia.org, the Wiki-based Collaborative Radiology Resource*, <https://radiopaedia.org/>. Accessed: Jun. 26, 2022.
41. Q. D. Buchlak, N. Esmaili, J. C. Leveque, C. Bennett, F. Farrokhi, M. Piccardi, Machine learning applications to neuroimaging for glioma detection and classification: An artificial intelligence augmented systematic review, *J. Clin. Neurosci.*, **89** (2021), 177–198. <https://doi.org/10.1016/j.jocn.2021.04.043>
42. J. Amin, M. Sharif, A. Haldorai, M. Yasmin, R. S. Nayak, Brain tumor detection and classification using machine learning: A comprehensive survey, *Complex Intell. Syst.*, **8** (2022), 3161–3183. <https://doi.org/10.1007/s40747-021-00563-y>
43. Y. LeCun, Y. Bengio, G. Hinton, Deep learning, *Nature*, **521** (2015), 436–444. <https://doi.org/10.1038/nature14539>
44. E. Lotan, R. Jain, N. Razavian, G. M. Fatterpekar, Y. W. Lui, State of the art: machine learning applications in glioma imaging, *Am. J. Roentgenol.*, **212** (2019), 26–37. <https://doi.org/10.2214/ajr.18.20218>
45. D. Shen, G. Wu, H. I. Suk, Deep learning in medical image analysis, *Annu. Rev. Biomed. Eng.*, **19** (2017), 221–248. <https://doi.org/10.1146/annurev-bioeng-071516-044442>
46. R. Yamashita, M. Nishio, R. K. G. Do, K. Togashi, Convolutional neural networks: An overview and application in radiology, *Insights Imaging*, **9** (2018), 611–629. <https://doi.org/10.1007/s13244-018-0639-9>
47. L. Alzubaidi, J. Zhang, A. J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma, et al., Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions, *J. Big Data*, **8** (2021), 53. <https://doi.org/10.1186/s40537-021-00444-8>
48. Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, et al., Backpropagation applied to handwritten zip code recognition, *Neural Comput.*, **1** (1989), 541–551. <https://doi.org/10.1162/neco.1989.1.4.541>
49. G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, R. R. Salakhutdinov, Improving neural networks by preventing co-adaptation of feature detectors, preprint, arXiv:1207.0580.
50. M. D. Zeiler, R. Fergus, Visualizing and understanding convolutional networks, preprint, arXiv:1311.2901.
51. K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, preprint, arXiv:1409.1556.
52. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, et al., Going deeper with convolutions, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (2015), 1–9. <https://doi.org/10.1109/cvpr.2015.7298594>
53. C. Szegedy, S. Ioffe, V. Vanhoucke, A. Alemi, Inception-v4, Inception-ResNet and the impact of residual, preprint, arXiv:1602.07261.

54. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, preprint, arXiv:1512.03385.
55. S. Xie, R. Girshick, P. Dollár, Z. Tu, K. He, Aggregated residual transformations for deep neural networks, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (2017), 5987–5995. <https://doi.org/10.1109/cvpr.2017.634>
56. A. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, et al., Mobilenets: Efficient convolutional neural networks for mobile vision applications, preprint, arXiv:1704.04861.
57. B. Zoph, V. Vasudevan, J. Shlens, Q. V. Le, Learning transferable architectures for scalable image recognition, in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018. <https://doi.org/10.1109/cvpr.2018.00907>
58. M. Tan, Q. V. Le, Efficientnet: Rethinking model scaling for convolutional neural networks, in *Proceedings of the 36th International Conference on Machine Learning*, **97** (2019), 6105–6114. <http://proceedings.mlr.press/v97/tan19a.html>
59. A. Brock, S. De, S. L. Smith, K. Simonyan, High-performance large-scale image recognition without normalization, preprint, arXiv:2102.06171
60. T. Ridnik, H. Lawen, A. Noy, E. Ben, B. G. Sharir, I. Friedman, Tresnet: High performance gpu-dedicated architecture, in *IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2021. <https://doi.org/10.1109/wacv48630.2021.00144>
61. R. K. Srivastava, K. Gref, J. Schmidhuber, Highway networks, preprint, arXiv:1505.00387.
62. G. Huang, Z. Liu, L. Maaten, K. Q. Weinberger, Densely connected convolutional networks, preprint, arXiv:1608.06993.
63. R. Hao, K. Namdar, L. Liu, F. Khalvati, A transfer learning–based active learning framework for brain tumor classification, *Front. Artif. Intell.*, **4** (2021), 635766. <https://doi.org/10.3389/frai.2021.635766>
64. K. Muhammad, S. Khan, J. D. Ser, V. H. C. de Albuquerque, Deep learning for multigrade brain tumor classification in smart healthcare systems: A prospective survey, *IEEE Trans. Neural Networks Learn. Syst.*, **32** (2021), 507–522. <https://doi.org/10.1109/tnnls.2020.2995800>
65. R. Miotto, F. Wang, S. Wang, X. Jiang, J. T. Tudley, Deep learning for healthcare: review, opportunities and challenges, *Brief. Bioinf.*, **19** (2018), 1236–1246. <https://doi.org/10.1093/bib/bbx044>
66. A. H. Morad, H. M. Al-Dabbas, Classification of brain tumor area for MRI images, in *Journal of Physics: Conference Series*, **1660** (2020). <https://doi.org/10.1088/1742-6596/1660/1/012059>
67. Z. Huang, H. Xu, S. Su, T. Wang, Y. Luo, X. Zhao, et al., A computer-aided diagnosis system for brain magnetic resonance imaging images using a novel differential feature neural network, *Comput. Biol. Med.*, **121** (2020), 103818. <https://doi.org/10.1016/j.combiomed.2020.103818>
68. X. Dong, Z. Yu, W. Cao, Y. Shi, Q. Ma, A survey on ensemble learning, *Front. Comput. Sci.*, **14** (2019), 241–258. <https://doi.org/10.1007/s11704-019-8208-z>
69. M. Ganaie, M. Hu, A. K. Malik, M. Tanveer, P. N. Suganthan, Ensemble deep learning: A review, *Eng. Appl. Artif. Intell.*, **115** (2022), 105151. <https://doi.org/10.1016/j.engappai.2022.105151>
70. Y. Yang, H. Lv, N. Chen, A Survey on ensemble learning under the era of deep learning, *Artif. Intell. Rev.*, **56** (2022), 5545–5589. <https://doi.org/10.1007/s10462-022-10283-5>
71. A. Tharwat, Classification assessment methods, *Appl. Comput. Inf.*, **17** (2020), 168–192. <https://doi.org/10.1016/j.aci.2018.08.003>

72. M. Nazi, S. Shakil, K. Khurshid, Role of deep learning in brain tumor detection and classification (2015 to 2020): A review, *Comput. Med. Imaging Graph*, **91** (2021). <https://doi.org/10.1016/j.compmedimag.2021.101940>
73. P. Bulla, L. Anantha, S. Peram, Deep neural networks with transfer learning model for brain tumors classification, *Trait. Du Signal*, **37** (2020), 593–601. <https://doi.org/10.18280/ts.370407>
74. M. U. Rehman, S. Cho, J. H. Kim, K. T. Chong, Bu-net: Brain tumor segmentation using modified u-net architecture, *Electronics*, **9** (2020), 2203. <https://doi.org/10.3390/electronics9122203>
75. M. U. Rehman, S. Cho, J. H. Kim, K. T. Chong, Brainseg-net: Brain tumor mr image segmentation via enhanced encoder–decoder network, *Diagnostics*, **11** (2021), 169. <https://doi.org/10.3390/diagnostics11020169>
76. M. U. Rehman, J. Ryu, I. F. Nizami, K. T. Chong, RAAGR2-Net: A brain tumor segmentation network using parallel processing of multiple spatial frames, *Comput. Biol. Med.*, **152** (2023), 106426. <https://doi.org/10.1016/j.compbiomed.2022.106426>
77. S. Y. Lin, C. L. Lin, Brain tumor segmentation using U-Net in conjunction with EfficientNet, *PeerJ Comput. Sci.*, **10** (2024). <https://doi.org/10.7717/peerj-cs.1754>
78. P. Wang, Y. Liu, Z. Zhou, Supraspinatus extraction from MRI based on attention-dense spatial pyramid UNet network, *J. Orthop. Surg. Res.*, **19** (2024). <https://doi.org/10.1186/s13018-023-04509-7>
79. H. Yin, Y. Wang, J. Wen, G. Wang, B. Lin, W. Yang, et al., DFBU-Net: Double-branch flat bottom U-Net for efficient medical image segmentation, *Biomed. Signal Process. Control.*, **90** (2024), 105818. <https://doi.org/10.1016/j.bspc.2023.105818>
80. S. Banerjee, S. Mitra, F. Masulli, S. Rovetta, Deep radiomics for brain tumor detection and classification from multi-sequence MRI, preprint, arXiv:1903.09240
81. K. V. Muneer, V. R. Rajendran, P. K. Joseph, Glioma tumor grade identification using artificial intelligent techniques, *J. Med. Syst.*, **43** (2019). <https://doi.org/10.1007/s10916-019-1228-2>
82. C. Ge, I. Y. H. Gu, A. S. Jakola, J. Yang, Deep learning and multi-sensor fusion for glioma classification using multistream 2D convolutional networks, in *40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2018. <https://doi.org/10.1109/embc.2018.8513556>
83. Y. Yang, L. F. Yan, X. Zhang, Y. Han, H. Y. Nan, Y. C. Hu, et al., Glioma grading on conventional MR images: a deep learning study with transfer learning, *Front. Neurosci.*, **12** (2018), 804. <https://doi.org/10.3389/fnins.2018.00804>
84. S. Gutta, J. Acharya, M. Shiroishi, D. Hwang, K. Nayak, Improved glioma grading using deep convolutional neural networks, *Am. J. Neuroradiol.*, **42** (2020), 233–239. <https://doi.org/10.3174/ajnr.a6882>
85. Z. Lu, Y. Bai, Y. Chen, C. Su, S. Lu, T. Zhan, et al., The classification of gliomas based on a Pyramid dilated convolution resnet model, *Pattern Recognit. Lett.*, **133** (2020), 173–179. <https://doi.org/10.1016/j.patrec.2020.03.007>
86. H. Mzoughi, I. Njeh, A. Wali, M. B. Slima, A. B. Hamida, C. Mhiri, et al., Deep multi-scale 3D Convolutional Neural Network (CNN) for MRI Gliomas brain tumor classification, *J. Digit. Imaging*, **33** (2020), 903–915. <https://doi.org/10.1007/s10278-020-00347-9>
87. Y. Zhuge, H. Ning, P. Mathen, J. Y. Cheng, A. V. Krauze, K. Camphausen, et al., Automated glioma grading on conventional MRI images using deep convolutional neural networks, *Med. Phys.*, **47** (2020), 3044–3053. <https://doi.org/10.1002/mp.14168>

88. S. Khawaldeh, U. Pervaiz, A. Rafiq, R. Alkhawaldeh, Noninvasive grading of glioma tumor using magnetic resonance imaging with convolutional neural networks, *Appl. Sci.*, **8** (2017), 27. <https://doi.org/10.3390/app8010027>
89. C. Ge, I. Y. H. Gu, A. S. Jakola, J. Yang, Deep semi-supervised learning for brain tumor classification, *BMC Med. Imaging*, **20** (2020). <https://doi.org/10.1186/s12880-020-00485-0>
90. S. Liang, R. Zhang, D. Liang, T. Song, T. Ai, C. Xia, et al., Multimodal 3D DenseNet for IDH genotype prediction in gliomas, *Genes*, **9** (2018), 382. <https://doi.org/10.3390/genes9080382>
91. C. Ge, Q. Qu, I. Y. H. Gu, A. Jakola, 3D multi-scale convolutional networks for glioma grading using MR images, in *25th IEEE International Conference on Image Processing (ICIP)*, 2018. <https://doi.org/10.1109/icip.2018.8451682>
92. M. Sajjad, S. Khan, K. Muhammad, W. Wu, A. Ullah, S. W. Baik, Multi-grade brain tumor classification using deep CNN with extensive data augmentation, *J. Comput. Sci.*, **30** (2018), 174–182. <https://doi.org/10.1016/j.jocs.2018.12.003>
93. H. Özcan, B. G. Emiroglu, H. Sabuncuoğlu, S. Özdoğan, A. Soyer, T. Saygı, A comparative study for glioma classification using deep convolutional neural networks, *Math. Biosci. Eng.*, **18** (2021), 1550–1572. <https://doi.org/10.3934/mbe.2021080>
94. H. E. Hamdaoui, A. Benfares, S. Boujraf, N. E. H. Chaoui, B. Alami, M. Maaroufi, et al., High precision brain tumor classification model based on deep transfer learning and stacking concepts, *Indones. J. Electr.*, **24** (2021), 167–177. <https://doi.org/10.11591/ijeecs.v24.i1.pp167-177>
95. M. Decuyper, S. Bonte, K. Deblaere, R. V. Holen, Automated MRI based pipeline for segmentation and prediction of grade, IDH mutation and 1p19q co-deletion in glioma, *Comput. Med. Imaging Graph*, **88** (2021), 101831. <https://doi.org/10.1016/j.compmedimag.2020.101831>
96. C. M. Lo, Y. C. Chen, R. C. Weng, K. L. C. Hsieh, Intelligent glioma grading based on deep transfer learning of MRI radiomic features, *Appl. Sci.*, **9** (2019), 4926. <https://doi.org/10.3390/app9224926>
97. G. S. Tandel, A. Tiwari, O. Kakde, Performance optimisation of deep learning models using majority voting algorithm for brain tumour classification, *Comput. Biol. Med.*, **135** (2021). <https://doi.org/10.1016/j.compbiomed.2021.104564>
98. M. A. Naser, M. J. Deen, Brain tumor segmentation and grading of lower-grade glioma using deep learning in MRI images, *Comput. Biol. Med.*, **121** (2020). <https://doi.org/10.1016/j.compbiomed.2020.103758>
99. W. Ayadi, W. Elhamzi, I. Charfi, M. Atri, Deep CNN for brain tumor classification, *Neural Process. Lett.*, **53** (2021), 671–700. <https://doi.org/10.1007/s11063-020-10398-2>
100. Y. Xie, F. Zaccagna, L. Rundo, C. Testa, R. Agati, R. Lodi, et al., Convolutional neural network techniques for brain tumor classification (from 2015 to 2022): Review, challenges, and future perspectives, *Diagnostics*, **12** (2022), 1850. <https://doi.org/10.3390/diagnostics12081850>
101. P. C. Tripathi, S. Bag, A computer-aided grading of glioma tumor using deep residual networks fusion, *Comput. Methods Programs Biomed.*, **215** (2022). <https://doi.org/10.1016/j.cmpb.2021.106597>
102. S. Gull, S. Akbar, S. M. Naqi, A deep learning approach for multi-stage classification of brain tumor through magnetic resonance images, *Int. J. Imaging Syst. Technol.*, **33** (2023), 1745–1766. <https://doi.org/10.1002/ima.22897>

103. G. S. Tandel, A. Tiwari, O. G. Kakde, N. Gupta, L. Saba, J. S. Suri, Role of ensemble deep learning for brain tumor classification in multiple magnetic resonance imaging sequence data, *Diagnostics*, **13** (2023), 481. <https://doi.org/10.3390/diagnostics13030481>
104. S. V. Rubio, M. T. García-Ordás, O. García-Olalla Olivera, H. Alaiz-Moretón, M. I. González-Alonso, J. A. Benítez-Andrades, Survival and grade of the glioma prediction using transfer learning, *PeerJ Comput. Sci.*, **9** (2023), 1723. <https://doi.org/10.7717/peerj-cs.1723>
105. L. Alzubaidi, J. Zhang, A. J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma, et al., Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions, *J. Big Data*, **8** (2021). <https://doi.org/10.1186/s40537-021-00444-8>
106. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov,] Dropout: a simple way to prevent neural networks from overfitting, *J. Mach. Learn. Res.*, **15** (2014), 1929–1958. <http://jmlr.org/papers/v15/srivastava14a.html>
107. R. C. Moore, J. DeNero, L1 and L2 regularization for multiclass hinge loss models, in *Proceedings of the Symposium on Machine Learning in Speech and Natural Language Processing*, 2011.
108. S. Ioffe, C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift, in *Proceedings of the 32nd International Conference on Machine Learning*, (2015), 448–456. <https://doi.org/10.48550/arXiv.1502.03167>.
109. Z. Khazaei, M. Langarizadeh, M. E. S. Ahmadabadi, Developing an artificial intelligence model for tumor grading and classification, based on mri sequences of human brain gliomas, *Int. J. Cancer Manage.*, **15** (2022). <https://doi.org/10.5812/ijcm.120638>
110. D. R. Sarvamangala, R. V. Kulkarni, Convolutional neural networks in medical image understanding: a survey, *Evol. Intell.*, **15** (2022), 1–22. <https://doi.org/10.1007/s12065-020-00540-3>
111. M. Versaci, G. Angiulli, F. LaForesta, F. Laganà, Palumbo, A. Annunziata, Intuitionistic fuzzy divergence for evaluating the mechanical stress state of steel plates subject to bi-axial loads, *Integr. Comput. Aided Eng.*, (2024), 1–17. <https://doi.org/10.3233/ica-230730>
112. W. Luo, D. Phung, T. Tran, S. Gupta, S. Rana, C. Karmakar, et al., Guidelines for developing and reporting machine learning predictive models in biomedical research: a multidisciplinary view, *J. Med. Internet Res.*, **18** (2016). <https://doi.org/10.2196/jmir.5870>.



AIMS Press

©2024 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>)