STA-395-01
Duy Nguyen & Anshul Satav

Development of Convolutional Neural Network in Classifying MRI Scans of Brain Tumors

Introduction

Brain tumors are associated with significant morbidity and mortality due to their critical anatomical location and infiltrative growth patterns. Tumors of the central nervous system (CNS) are broadly classified as primary brain tumors, which originate within the brain or spinal cord, and secondary brain tumors, which arise from metastatic spread of malignancies from outside the CNS and occur five to ten times more frequently (Weller et al., 2024). According to the Central Brain Tumor Registry of the United States (CBTRUS), the average annual age-adjusted incidence rate of all malignant and non-malignant brain and other CNS tumors is 24.83 per 100,000 population. The reported five-year relative survival rates are 35.7% for malignant tumors and 91.8% for non-malignant tumors (Ostrom et al., 2023). Among malignant brain tumors, glioblastoma represents the most prevalent and aggressive subtype, accounting for 14.2% of all cases and 50.9% of malignant cases. In contrast, meningiomas, originating from the meninges surrounding the brain and spinal cord, constitute the most common non-malignant tumors (40.8%), followed by pituitary adenomas (17.2%), which originate on the pituitary gland at the base of the brain (Ostrom et al., 2023; Weller et al., 2024).

Early detection is an important determinant of therapeutic success and patient prognosis in brain tumor management. Non-invasive imaging techniques such as magnetic resonance imaging (MRI), computed tomography (CT), and positron emission tomography (PET) play a central role in tumor diagnosis and clinical decision-making, with MRI and CT being the most widely used (Kim et al., 2021). Among these, MRI is particularly advantageous due to its soft-tissue contrast and lack of ionizing radiation. MRI scans are normally acquired in three perspectives, including axial (from the top to the bottom of the brain), coronal (from the front to the back of the brain), and sagittal (sideway of the brain), that provide comprehensive anatomical coverage of the brain. The imaging process relies on the interaction between a strong external magnetic field and the nuclear magnetization of hydrogen protons in tissue water. Radiofrequency (RF) pulses perturb proton alignment, and the emitted RF signals during relaxation are reconstructed into grayscale images via Fourier transformation. Tissue contrast is governed primarily by T1 (longitudinal) and T2 (transverse) relaxation times, which reflect differences in proton realignment and phase coherence loss, respectively. This enables the differentiation between fat-rich and and water-rich components in tissues (Deoni, 2020).

In recent years, deep learning has emerged as a powerful paradigm in medical image analysis, offering automated feature extraction and high-performance classification. Numerous studies and reviews have demonstrated the transformative impact of deep learning across a range of medical imaging tasks, including disease detection, segmentation, and classification, spanning

different modalities of MRI, CT, and even ultrasound (Suganyadevi, Seethalakshimi & Balasamy, 2021). Convolutional neural networks (CNNs), in particular, have proven effective for image-based learning due to their ability to hierarchically capture spatial features directly from pixel-level data. Advances such as transfer learning have further enhanced CNN performance by leveraging pretrained architectures, especially in limited-data settings (Ilani, Shi & Banad, 2025; Khaliki & Basarlan, 2024).

Despite their successes, CNN-based approaches often require substantial computational resources and large, well-annotated datasets to achieve optimal performance, which can restrict their practical application in resource-constrained contexts. To address these challenges, this project aims to develop a computationally efficient CNN architecture with reduced depth and parameter numbers for the classification of brain tumor MRI images. The performance of this lightweight model is subsequently compared against a MobileV2-based CNN with transfer learning and feature extraction to evaluate the trade-offs between model complexity, computational efficiency, and classification accuracy.

Materials & Methods

Original Data & Code

The data used in this project along with the code are available at this repository.

Dataset & Preprocessing

The brain MRI dataset in this project was obtained from a publicly available repository hosted on Kaggle and represents a curated aggregation of three independent datasets, including Figshare, DARTAJ, and Br35H. The dataset consists of 7,023 MRI scans, presumably from different patients, from all three perspectives, providing comprehensive spatial information. Scans are provided in the portable network graphics (PNG) format and are separated into individual training and testing folders, each categorized into four classes of glioma, meningioma, pituitary adenoma, and no tumor.

All scans are subjected to a preprocessing pipeline adapted from Kaggle to standardize the input size and reduce background variability. Each scan was converted to grayscale from RGB and smoothed using a Gaussian filter to suppress high-frequency noise. Binary thresholding followed by morphological erosion and dilation was applied to isolate the brain region from background areas. The largest connected contour was identified and used to determine the extreme boundaries points, allowing for tight cropping around the brain region while excluding the extraneous image content. Cropped scans were subsequently resized to a uniform resolution of 256 x 256 pixels to ensure consistent input dimensions for CNN training.

To mitigate redundancy, duplicate scans in each folder were identified using a hash-based approach in which an MD5 hash was computed from the raw pixel values of each scan. 215 duplicate scans in the training folder (3.8%) and 37 duplicate scans in the testing folder (2.8%)

were removed. In addition, to prevent data leakage, 182 duplicate scans (14.3%) in both folders were further removed in the testing folders. After duplicate removal, the two folders were combined and stratified into a new training and testing set at the ratio of 75%-25% with preserved class proportions across both sets. The final training set contained 4,941 scans, including 1,274 glioma scan, 1,215 meningioma scans, 1,147 scans without tumors, and 1,305 pituitary adenoma scans. The final testing set contained 1,648 scans, including 425 glioma scans, 405 meningioma scans, 383 scans without tumors, and 435 pituitary adenoma scans.

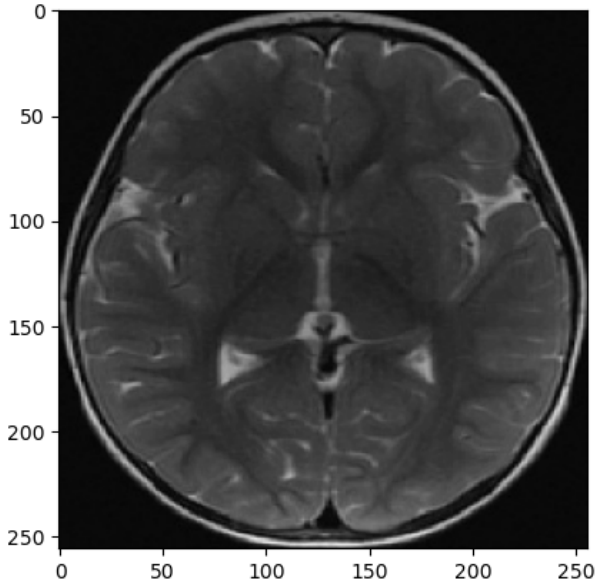

Figure 1. An MRI scan from the training set after preprocessing

Data Augmentation & Tensor Preparation

To enhance model generalization and reduce overfitting, data augmentation was applied to the training scans using the transformation pipeline in PyTorch. Augmentations were designed to introduce realistic variability while preserving relevant features of the MRI scans. In particular, scans were subjected to Gaussian blurring with randomly sampled blur intensities to simulate acquisition-related noise, followed by random horizontal flipping to account to bilateral anatomical symmetry. Additional geometric variability was introduced through random rotation of 30 degrees and affine transformations with shear, capturing possible variations in patient positioning. Random resized cropping was also added by sampling 90%-100% of the original scan area and resizing back to 256 x 256 pixels, promoting spatial robustness while maintaining consistent input dimensions.

In terms of tensor preparation, the scan arrays and labels were converted into PyTorch tensors and then reshaped into the correct format required as the inputs for CNN before being stored in contiguous memory to ensure compatibility with convolution operations. To ensure reproducibility, a random seed was fixed prior to tensor construction. The processed tensors were

then wrapped into a TensorDataset object and loaded into a DataLoader with a batch size of 32 for both training and testing tensors.

Custom CNN Architecture (GBMnet)

A custom CNN, termed GBMnet, was developed as the primary lightweight model in this project and serves as the baseline architecture against which a MobileNetV2-based transfer learning model was later compared. The network design was inspired by the architectural principles described by Hashemzehi et al. (2020) and prioritizes computational efficiency while maintaining sufficient classification capacity for brain tumors.

GBMnet consists of 4 convolutional blocks, each employing a 3 x 3 convolution with stride 1 and padding 1 to learns spatially localized features, followed by batch normalization to stabilize feature distributions and reduce sensitivity to variability in intensity, rectified linear unit (ReLU) activation, and 2 x 2 max-pooling for progressive spatial down-sampling. The number of feature channels increases from 8 to 64 across successive layers, allowing for hierarchical feature extraction from low-level intensity patterns to higher-level structural patterns. For an input resolution of 256 x 256 pixels, the network reduces spatial dimensionality to 16 x 16 pixels at the final convolutional stage.

The extracted feature maps are flattened and passed through a fully connected layer of 512 neurons with ReLU activation before mapping the output to 4 classes. Dropout rate of 0.3 was used to mitigate overfitting. Lower numbers of neurons (256 neurons, 128 neurons, and 64 neurons) were also attempted but they lowered the training accuracy, potentially due to the insufficiency of parameters.

Transfer Learning MobileNetV2 CNN Architecture (GBM_MobileNetV2net)

MobileNetV2 is a deep yet computationally efficient CNN originally developed for large-scale natural image recognition on resource-constrained devices. At the standard input resolution of 224 x 224 pixels, the architecture includes an initial convolutional layer followed by 18 inverted residual bottleneck blocks and contains approximately 3.4 million parameters, making it substantially deeper than the custom GBMnet model. Its defining design relies on depthwise separable convolutions combined with inverted residuals and linear bottlenecks, wherein feature dimensionality is first expanded via pointwise convolutions to enable non-linear transformation using ReLU6 activations with subsequent depthwise spatial filtering and projection back into a compact feature space (Sandler et al., 2018). This architecture reduces parameter count and computational cost while preserving representational capacity, with batch normalization applied throughout to stabilize the training. In this project, a transfer learning strategy was employed by initializing MobileNetV2 with ImageNet-pretrained weights and freezing all convolutional feature extraction layers to retain the general-purpose hierarchical visual features learned from millions of natural images, inspired by the approach described by Khan et al. (2025). This transfer learning model is termed GBM_MobileNetV2net. The original classifier was replaced with a task-specific full connected layer of 64 neurons with ReLU activation and dropout rate of 0.5, followed by the

mapping of the output to 4 classes. The input MRI scans were resized to 224 x 224 pixels to conform to the expected input dimensions of MobileNetV2.

Training Setup

Both the custom GBMnet and the transfer learning GBM_MobileNetV2net were trained under an identical optimization framework for fair performance comparison. To ensure reproducibility, a fixed random seed was set prior to training. Models were trained for 200 epochs using the Adam optimizer with a learning rate of 0.00001and batch size of 32. Training was performed on Apple Metal Performance Shaders (MPS) on an Macbook Air with M2 chip. Cross-entropy cost was used as the objective function for the evaluation of multi-class classification. During each training epoch, models were set to training mode and optimized using mini-batch gradient descent, with data augmentation applied directly to the input scans. Forward propagation was followed by backpropagation and parameter updates at each iteration, while training cost and accuracy were tracked across epochs to monitor convergence. Final-epoch predictions were stored for downstream evaluation and learning dynamics were visualized by plotting the cost and accuracy over time.

Results

GBMnet Performance

The GBMnet model demonstrated stable convergence and strong classification performance across both the training and testing sets. Training accuracy increased steadily from 83.42% at epoch 20 to 95.55% after 200 epochs, with a corresponding gradual decrease in loss (Figure 2.). The model also achieved a final balanced accuracy of 95.41%, indicating the minimal impact of imbalances between classes. The training confusion matrix (Figure 3.) shows that most scans were correctly classified, with particularly strong performance for glioma (1,255/1,274) and pituitary adenoma (1,280/1,305). Misclassifications were relatively limited and primarily occurred between meningioma or pituitary adenoma.

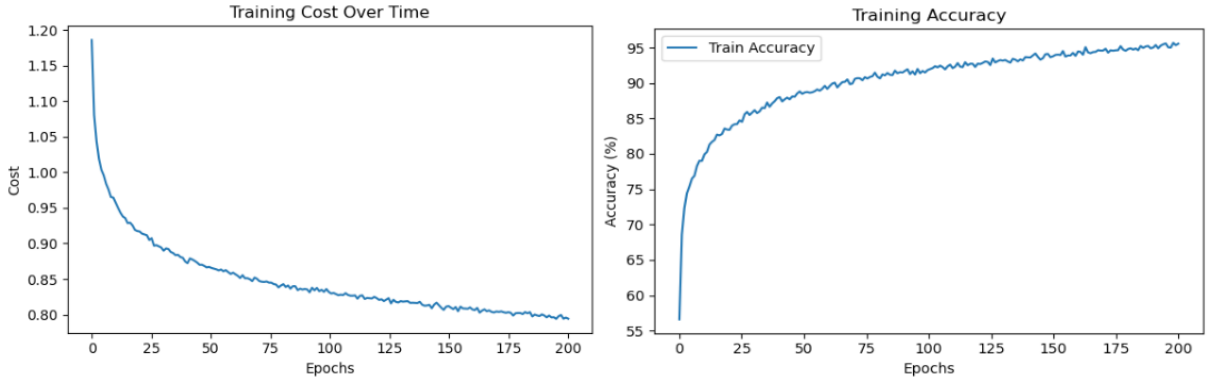

Figure 2. The training cost and accuracy of GBMnet over time

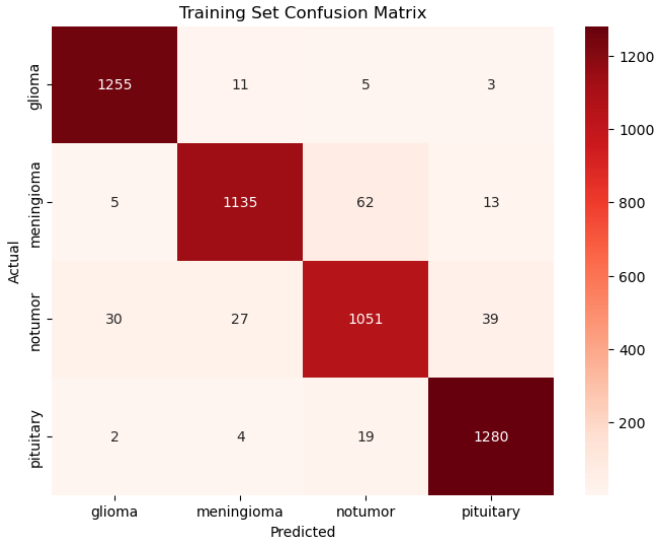

Figure 3. Confusion matrix for GBMnet on the training set

Class	Precision	Recall	F1 Score	Scan Number
Glioma	0.97	0.99	0.98	1,274
Meningioma	0.96	0.93	0.95	1,215
No tumor	0.92	0.92	0.92	1,147
Pituitary	0.96	0.98	0.97	1,305

Table 1. Classification report for GBMnet on the training set

On the testing set, GBMnet achieved an accuracy of 90.66% and a balanced accuracy of 90.15%, revealing good generalization with modest performance drop relative to training. The testing confusion matrix (Figure 4.) shows robust performance for glioma (417/425) and pituitary adenoma (428/435), consistent with the training set. In contrast, the no-tumor class hints at increased confusion, with a significant number of samples misclassified as glioma or meningioma, suggesting while GBMnet effectively captures the discriminative features for tumor subtypes, distinguishing normal scans from pathological scans remains the main source of errors. The precision, recall, and F1 score of each class further confirmed these patterns across both sets (Table 1. & 2.).

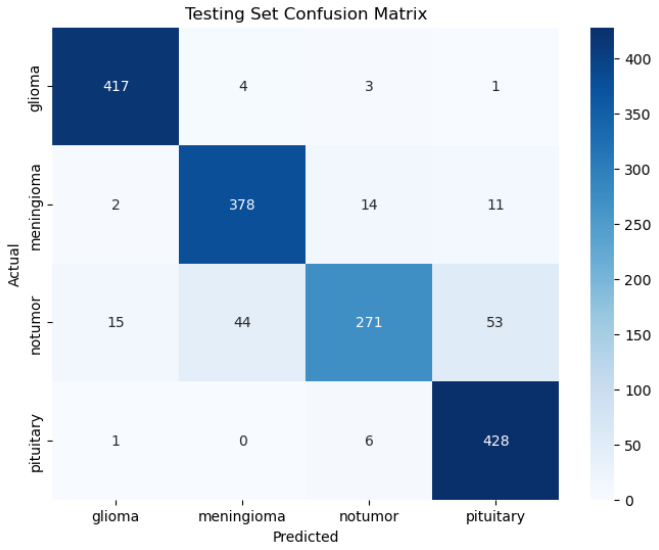

Figure 4. Confusion matrix for GBMnet on the testing set

Class	Precision	Recall	F1 Score	Scan Number
Glioma	0.96	0.98	0.97	425
Meningioma	0.89	0.93	0.91	405
No tumor	0.92	0.71	0.80	383
Pituitary	0.87	0.98	0.92	435

Table 2. Classification report for GBMnet on the testing set

GBM_MobileNetV2net Performance

The transfer learning model, GBM_MobileNetV2net, showed a slower convergence and consistently lower performance than the custom GBMnet model. During training, accuracy increased gradually from 75.69% at epoch 20 to a final accuracy of 84.21% and balanced accuracy of 83.89% at epoch 200, with the cost plateauing around 0.92 (Figure 5.), indicating that the frozen MobileNetV2 feature extraction restrained the ability of the model to fully adapt to domain-specific MRI features. The training confusion matrix (Figure 6.) shows that most glioma (1,196/1,274) and pituitary adenoma (1,154/1,305) scans were correct, whereas meningioma (947/1,215) and no-tumor (864/1,147) exhibited higher misclassification rates.

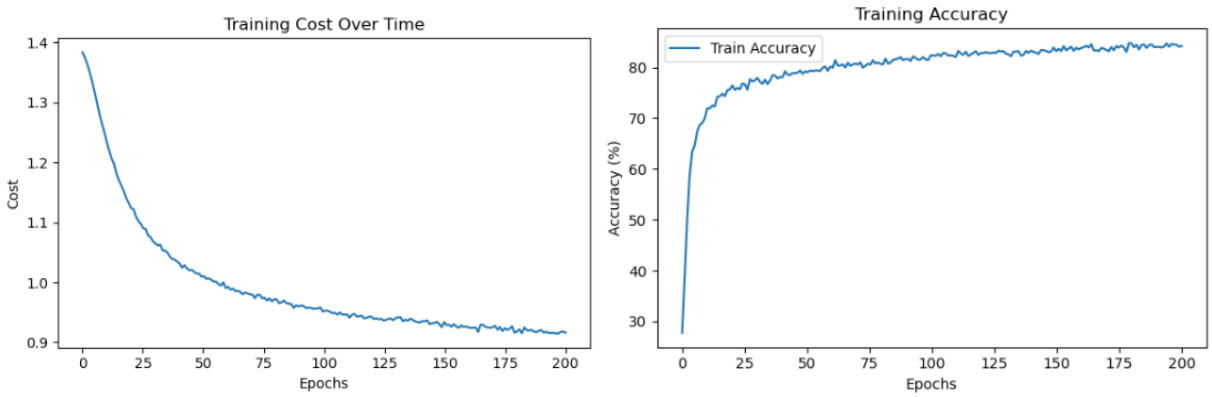

Figure 5. The training cost and accuracy of GBM_MobileNetV2net over time

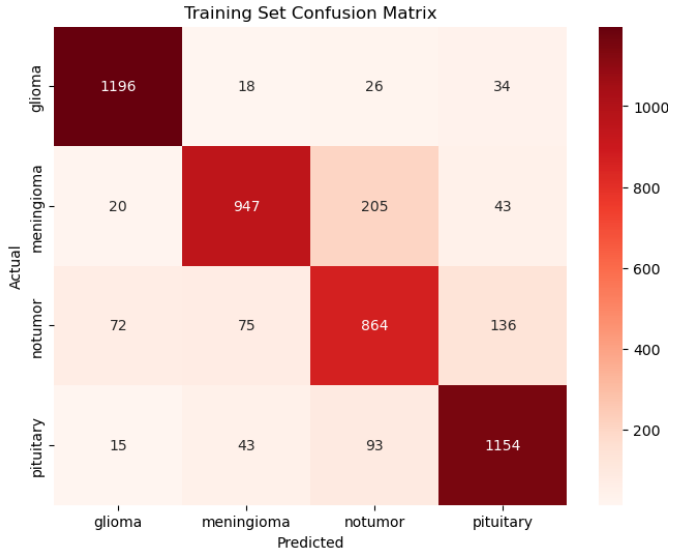

Figure 6. Confusion matrix for GBM_MobileNetV2net on the training set

Class	Precision	Recall	F1 Score	Scan Number
Glioma	0.92	0.94	0.93	1,274
Meningioma	0.87	0.78	0.82	1,215
No tumor	0.73	0.75	0.74	1,147
Pituitary	0.84	0.88	0.86	1,305

Table 3. Classification report for GBM_MobileNetV2net on the training set

On the testing set, the model achieved an overall accuracy of 77.31% and a balanced accuracy of 76.13%, which is a noticeable drop in generalization compared to training. The testing confusion matrix (Figure 7.) demonstrates that glioma (394/425) and pituitary adenoma (403/435) scans were still reliably classified, while meningioma (378/405) and especially no-tumor (99/383) scans were significantly misclassified, often predicted as tumor-bearing. These results were also reflected in the precision, recall, and F1 score as alternative classification metrics (Table 3. & 4.).

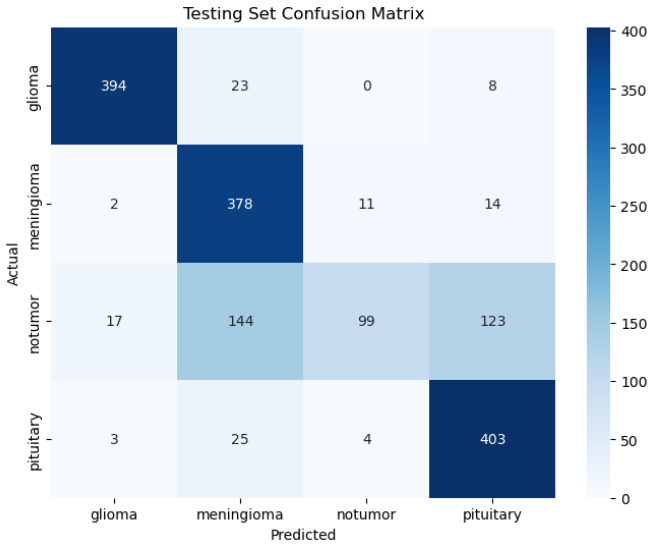

Figure 7. Confusion matrix for GBM_MobileNetV2net on the testing set

Class	Precision	Recall	F1 Score	Scan Number
Glioma	0.95	0.93	0.94	425
Meningioma	0.66	0.93	0.78	405
No tumor	0.87	0.26	0.40	383
Pituitary	0.74	0.93	0.82	435

Table 4. Classification report for GBM_MobileNetV2net on the testing set

Discussion

The comparative evaluation of the custom GBMnet model and the transfer learning GBM_MobileNetV2 model highlights both the strengths and weaknesses of these approaches for brain tumor classification. GBMnet, a shallow, end-to-end CNN tailored for brain MRI data, achieved superior performance on both the training and testing sets, demonstrating stable convergence, high accuracy, and minimal class imbalance effects. Its strong generalization, particularly for glioma and pituitary adenoma scans, indicates that designing networks specific to domain features can capture discriminative patterns without the need for significant computational resources. In contrast, GBM_MobileNetV2net, despite leveraging the hierarchical features learned

from millions of natural images via transfer learning, showed a lower performance. This may partly result from the decision to freeze all convolutional layers due to restrictive capability of a personal laptop, preventing the network from adapting pretrained parameters to the new MRI domain. It is also important to recognize that both models were trained on a relatively small dataset of static MRI scans (normal MRI would produce multiple slices of the brain so the scans would be slightly different in real-word situation), which may further limit generalization. From a practicality standpoint, the lightweight design and high accuracy of the GBMnet make it a better option for deployment in clinical context as it would require less computational resources while achieving reliable predictions. Future directions could involve selectively unfreezing layers in the MobileNetV2 to allow for domain-specific fine-tuning, expanding the dataset with multiple slices of MRI from the same patient to provide additional spatial information, and exploring feature engineering to improve the discrimination between normal and pathological scans.

References

1. Deoni, S. (2020). Biophysical and physiological principles of T1 and T2. *Advances in Magnetic Resonance Technology and Applications*, 3-17. https://doi.org/10.1016/b978-0-12-817057-1.00003-2
2. Hashemzehi, R., Mahdavi, S. J., Kheirabadi, M., & Kamel, S. R. (2020). Detection of brain tumors from MRI images base on deep learning using hybrid model CNN and Nade. *Biocybernetics and Biomedical Engineering*, *40*(3), 1225–1232. https://doi.org/10.1016/j.bbe.2020.06.001
3. Ilani, M. A., Shi, D., & Banad, Y. M. (2025). T1-weighted MRI-based brain tumor classification using hybrid deep learning models. *Scientific Reports*, *15*(1). https://doi.org/10.1038/s41598-025-92020-w
4. Khaliki, M. Z., & Başarslan, M. S. (2024). Brain tumor detection from images and comparison with transfer learning methods and 3-layer CNN. *Scientific Reports*, *14*(1). https://doi.org/10.1038/s41598-024-52823-9
5. Khan, M. A., Hussain, M. Z., Mehmood, S., Khan, M. F., Ahmad, M., Mazhar, T., Shahzad, T., & Saeed, M. M. (2025). Transfer learning for accurate brain tumor classification in MRI: A step forward in medical diagnostics. *Discover Oncology*, *16*(1). https://doi.org/10.1007/s12672-025-02671-4
6. Kim, B., Kim, H., Kim, S., & Hwang, Y. (2021). A brief review of non-invasive brain imaging technologies and the near-infrared optical bioimaging. *Applied Microscopy*, *51*(1). https://doi.org/10.1186/s42649-021-00058-7
7. Ostrom, Q. T., Price, M., Neff, C., Cioffi, G., Waite, K. A., Kruchko, C., & Barnholtz-Sloan, J. S. (2023). CBTRUS statistical report: Primary Brain and other central nervous system tumors diagnosed in the United States in 2016-2020. *Neuro-Oncology*, *25*. https://doi.org/10.1093/neuonc/noad149
8. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). MobileNetV2: Inverted residuals and linear bottlenecks. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 4510-4520. https://doi.org/10.1109/cvpr.2018.00474
9. Suganyadevi, S., Seethalakshmi, V., & Balasamy, K. (2021). A review on deep learning in medical image analysis. *International Journal of Multimedia Information Retrieval*, *11*(1), 19-38. https://doi.org/10.1007/s13735-021-00218-1
10. Weller, M., Wen, P. Y., Chang, S. M., Dirven, L., Lim, M., Monje, M., & Reifenberger, G. (2024). Glioma. *Nature Reviews Disease Primers*, *10*(1). https://doi.org/10.1038/s41572-024-00516-y