AUIQ Complementary Biological System
ISSN: 3007-973X
Journal homepage:
https://acbs.alayen.edu.iq

Manuscript 1046

Meningioma Detection from MRI: A Comparative Study of Deep Learning Models

A. H. Doha
Sandi Jennifer
Fatima Alabdaly

AUIQ COMPLEMENTARY BIOLOGICAL SYSTEM 3 (2026) 94–100

Scan the QR to view the full-text article on the journal website

ORIGINAL STUDY

Meningioma Detection from MRI: A Comparative Study of Deep Learning Models

A. H. Doha *, Sandi Jennifer, Fatima Alabdaly

University of southern Queensland, School of Health science and Medical, Australia, 4350

ABSTRACT

This study compares several deep learning models including DenseNet20, CNN, ResNet50, VGG19, EfficientNet-B3, MobileNetV2, NASNetMobile, Xception to automatically classify brain meningiomas from MRI scans. A public dataset was used to evaluate the proposed model. All models were trained, validated, and tested on this dataset. Because meningiomas are common and manual reads, the goal is to speed and standardize diagnosis. Our main contribution is pairing a high-accuracy classifier with deep learning models. Among all models, DenseNet201 performed best, reaching 0.991% accuracy and reliably separating tumorous from non-tumorous MRIs. Future work will expand the dataset and incorporate stronger segmentation to further boost performance and robustness.

Keywords: Meningioma, Deep learning, Magnetic resonance imaging (MRI), Explainable artificial intelligence (XAI), Grad-CAM

1. Introduction

Meningioma arises from the meninges the protective membranes around the brain and spinal cord—and is among the most common central nervous system tumors [1]. While many cases are benign, some subtypes grow aggressively [2], compressing neural tissue and causing serious neurological symptoms [3]. Given the global incidence, timely and accurate diagnosis is critical [4]; delayed detection or misread imaging can lead to more complex, late-stage surgical management [5]. Magnetic resonance imaging (MRI) is the primary tool for meningioma assessment, revealing tumor location, size, and morphology [6]. However, routine diagnosis relies on manual reads, which are time-consuming, subject to inter-reader variability, and sometimes confounded by conditions with similar imaging appearances [7–20].

Meningioma Detection from MRI is a critical task in neuro oncology, as early and accurate detection of meningiomas significantly enhance treatment planning and patient outcomes. Meningiomas are among the most common primary brain tumours, appearing from the meninges and often staging with symptoms only after substantial growth. MRI is the preferred modality for brain tumour assessment due to its superior soft tissue contrast and non-invasive nature. However, manual interpretation of MRI scans is time consuming and prone to experts variability, especially when distinguishing meningiomas from other intracranial abnormalities. Recent advancements in artificial intelligence and deep learning have enabled automated systems to analyse complex imaging patterns, improving diagnostic accuracy, and assisting radiologists in clinical decision making. Therefore, developing an efficient and accurate deep learning–based model for meningioma detection from MRI images is essential for enhancing early diagnosis, reducing workload, and advancing computer-aided medical imaging applications.

We employed different deep learning models that automatically classifies meningioma on MRI. The dataset combines patient scans from the Faculty of

Received 17 September 2025; revised 22 November 2025; accepted 22 November 2025.
Available online 16 June 2026

* Corresponding author.
E-mail address: u1175494@umail.usq.edu.au (A. H. Doha).

https://doi.org/10.70176/3007-973X.1046

Medicine with publicly available healthy controls, labelled into two classes ("tumorous" vs. "non-tumorous") and split for training, validation, and testing. We compared DenseNet20, CNN, ResNet50, VGG19, EfficientNet-B3, MobileNetV2„ NASNetMobile, Xception, under a consistent protocol [20–30]. DenseNet201 achieved the best performance. To make the model's decisions transparent, we apply Grad-CAM, which highlights image regions that drive predictions and aligns with clinically relevant areas, improving trust and interpretability.

Beyond high accuracy using MRI alone, our contribution is coupling strong classification with deep learning models to support clinical decision-making. Future work will expand the dataset and incorporate advanced segmentation to further improve robustness and generalization.

2. Experimental dataset

The Brain Tumour MRI dataset, sourced from Rob flow Universe, is an extensive collection intended for developing and evaluating computer vision models in brain tumour detection and classification. It contains 3,903 MRI scans, organized into four major categories:

- Glioma: Tumours that develop from the brain's glial cells.
- Meningioma: Tumours that originate in the meninges, the protective membranes surrounding the brain and spinal cord.
- Pituitary Tumour: Growths occurring in the pituitary gland, which can disrupt hormonal regulation.
- No Tumour: MRI images showing normal brain tissue without any signs of tumour formation.

The images and data utilized in this study were acquired from Medical Image DataSet: Brain Tumor Detection. All data and images were fully anonymized in compliance with ethical guidelines before use. Figure displays a sample of the collected images.

3. Methodology

To overcome dataset limitations and improve the model's generalization capability, various data augmentation methods were applied. During preprocessing, all images were scaled to meet the input specifications of the selected CNN architectures, and the data pool was separated into training (70%), validation (20%), and testing (10%) sections. Several strategies, including arbitrary rotations up to 30 degrees, horizontal flipping, brightness adjustments, and the addition of Gaussian noise, were applied to optimize the model's performance. Additionally,

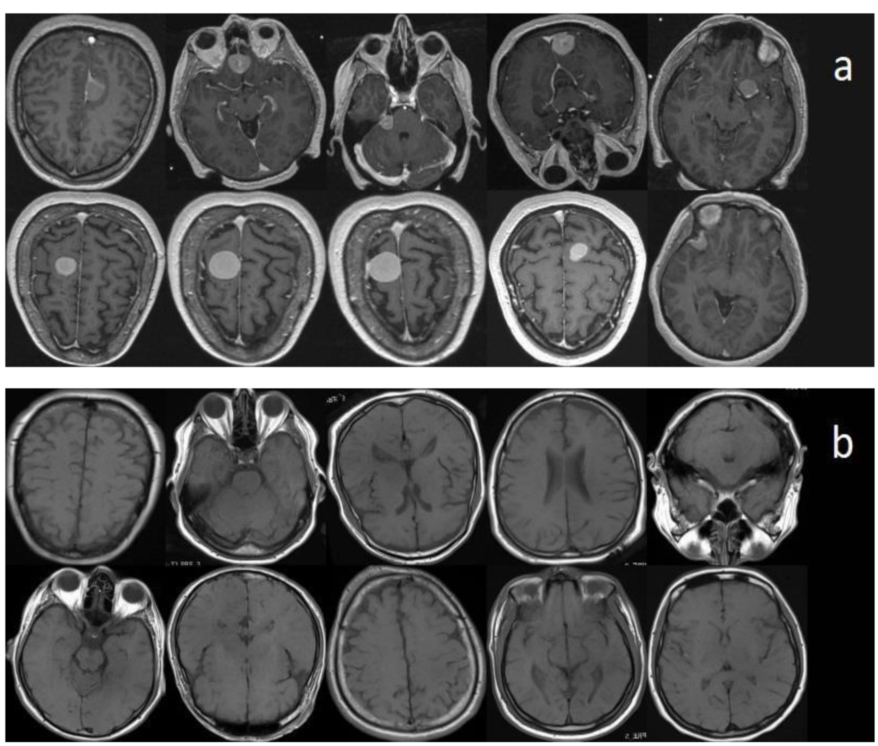

Fig. 1. This figure presents T1-weighted axial MRI scans illustrating the presence and absence of meningioma tumours: (a) Meningioma-positive cases, (b) Meningioma-negative cases.

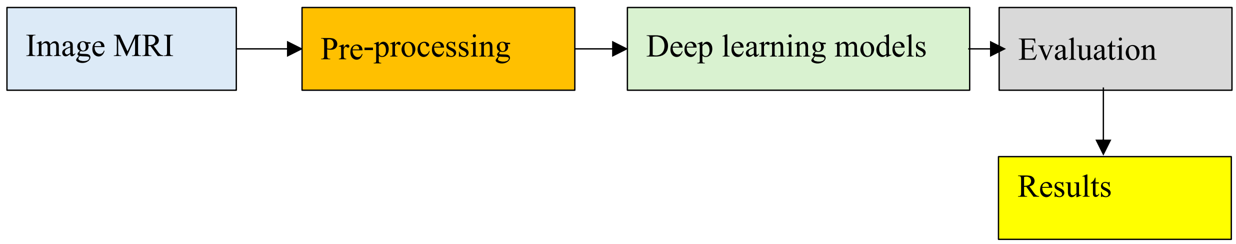

Fig. 2. The proposed model for Meningioma detection from MRI.

pixel values were scaled within the [0, 1] range to ensure stable learning. Preprocessing layers were added to maintain compatibility with the model's input specifications, and each step was executed systematically to optimize accuracy and efficiency. Finally, the performance of several CNN-based deep learning models, including ResNet50, VGG19, EfficientNet, MobileNetV2, DenseNet201, NASNetMobile, Xception, and InceptionV3, was evaluated [30–40].

- ResNet50 is a deep learning model that utilizes residual connections to mitigate the diminishing gradient issue observed in deep neural networks [10].
- VGG19 features a 19-layer architecture and is characterized by its use of small filter sizes [11]. EfficientNet enhances parameter efficiency through a model scaling strategy that systematically balances factors such as network size, width, and depth [12].
- MobileNetV2 is specifically optimized for mobile and embedded applications, reducing computational overhead using depthwise separable convolutions and linear bottleneck layers [13].
- DenseNet201 improves gradient flow and enhances parameter efficiency by establishing dense connections between layers [14].
- InceptionV3 is a computationally efficient model that leverages inception modules to enable multiscale feature extraction [15].
- NASNetMobile, designed through neural architecture search (NAS), is optimized for mobile platforms, utilizing efficient factorized convolutions to achieve high accuracy with lower computational complexity [25].
- Xception extends the Inception architecture by replacing conventional convolutions with depth wise separable convolutions, improving efficiency and performance while reducing parameter count [24]. These models were chosen due to their high accuracy in medical image classification applications and their varied architectural methodologies.

The models were refined to optimize medical image classification by adjusting their final layers. To ensure accurate differentiation between 'tumorous' and 'non-tumorous' cases in binary classification, the Sigmoid activation function was integrated. Various layer configurations and compilation techniques were implemented to enhance efficiency and maximize classification performance [16]. Initially, a Global Average Pooling layer was incorporated to improve feature extraction and reduce redundant parameters. The output unit employed the Sigmoid activation function to estimate predicted class scores (tumorous/non-tumorous). To establish a stable and efficient training process, the Stochastic Gradient Descent optimization algorithm was utilized

[17]. The Binary Cross-Entropy loss function was applied to minimize classification errors [18]. To assess the model's classification efficiency, key performance indicators including accuracy, recall, precision, Receiver Operating Characteristic (ROC), F1-Score, and Area Under the Curve (AUC) were calculated, and the model's overall capability was examined in depth [1725]. The training process was meticulously designed to ensure effective learning for deep learning models. Most models were trained for a total of 10 epochs, with cross-validation applied in certain cases to enhance overall performance [19]. During training, fine-tuning techniques were utilized by unfreezing specific layers of pre-trained models. This approach leveraged transfer learning to improve the models' adaptability to the training dataset. To Mitigate Overfitting, early stopping and data augmentation techniques were consistently implemented throughout the learning process [20]. The weight optimization process was conducted using the SGD algorithm, while the Binary Cross-Entropy loss function was implemented as the error measure to effectively guide the learning process.

Upon the completion of the learning stage, the accuracy of the primary model, which incorporates a pre-trained base architecture, was assessed using the test dataset. A thorough assessment was conducted to rigorously analyse the model's classification effectiveness. To achieve this, a wide selection of performance

Table 1. Performance evaluation of deep learning models using Accuracy, Loss, Precision, Recall, F1-Score, Specificity and AUC.

Base Model	Acc	Loss	Precision	Recall	F1-Score	Specificity	AUC
VGG19	0.944	0.154	0.963	0.933	0.951	0.991	0.950
ResNet50	0.922	0.256	0.960	0.930	0.930	0.980	0.920
EfficientNet-B3	0.972	0.113	0.950	0.970	0.950	0.970	0.970
DenseNet201	**0.991**	**0.0679**	**0.950**	**0.970**	**0.970**	**0.980**	**0.990**
MobileNet-V2	0.981	0.1911	0.950	0.960	0.970	0.980	0.980
Inception-V3	0.982	0.1739	0.970	0.970	0.970	0.980	0.970
NASNetMobile	0.991	0.0719	0.970	0.970	0.970	0.970	0.970
Xception	0.990	0.1090	0.960	0.980	0.980	0.980	0.980

measures was utilized. These metrics are derived from fundamental classification outcomes and are defined as follows:

4. Experimental results

This study was conducted entirely on a local machine, leveraging high-performance computing resources to train and evaluate deep learning models. Eight different deep learning architectures were employed for the classification of brain MRI images including VGG19, EfficientNetB3, ResNet50, DenseNet201, MobileNetV2, NASNetMobile, Xception, and InceptionV3. The MRI images were preprocessed to enhance feature extraction and improve model performance. Unlike multiclass classification tasks, this study focused on binary classification. Accordingly, the final layer of each model was designed with a sigmoid activation function instead of SoftMax.

To evaluate model performance, the data sample was divided into training, validation, and test sets, guaranteeing that the models were tested on unseen data. Critical assessment measures like accuracy, loss, precision, recall, F1-score, AUC, and cross-entropy loss were computed for each model to facilitate comparisons. Results from eight different architectures are summarized in Table 1, which presents a summary of classification performance on the test dataset, serving as a comparative reference for measuring the effectiveness of each model.

Table 1 offers a systematic comparison eight deep learning models according to fundamental assessment metrics, such as accuracy, loss, precision, recall, F1-score, specificity, and AUC. Amongst the assessed models, DenseNet201 emerged as the top-performing architecture, it was closely followed by NASNetMobile These results highlight the strong generalization capabilities of these models. Additionally, Xception and Inception-V3 demonstrated notable performance, both attaining an F1-score close to 0.97.

EfficientNet-B3 achieved a respectable accuracy of 0.972%, though its precision and recall values were

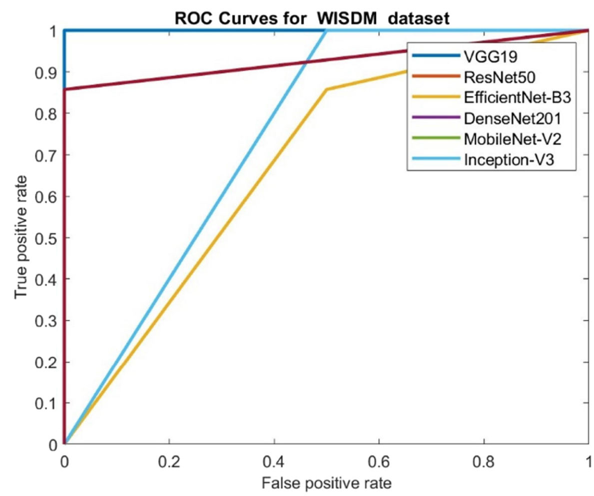

Fig. 3. performance evaluation using ROC curves for all models.

marginally lower compared to the top-performing models.

On the other hand, VGG19 exhibited relatively lower performance compared to other models. Specifically, VGG19 achieved an accuracy of 0.944%, while ResNet50 recorded the lowest accuracy among all models at 0.922%. This suggests that these architectures face challenges in maintaining balanced classification performance, particularly regarding precision and recall. Fig. 3 reports the results in term of ROC. The results in Fig. 3 support our findings in Table 1.

This work developed and compared deep learning models for meningioma classification from MRI scans [26, 27]. The main findings were discussed in this section. The complexity time of all models was reported in Fig. 4. From the results obtained, it can be noticed that all models showed accepted complexity time.

Across metrics accuracy, precision, recall, DenseNet201 outperformed all models. It delivered the highest results among all models. Relative to prior studies that often examine a single CNN or hand-engineered features, our multi models' comparison achieved ≥0.99 % top-line accuracy and incorporated explainable several deep learning models. The dataset combined real patient scans from

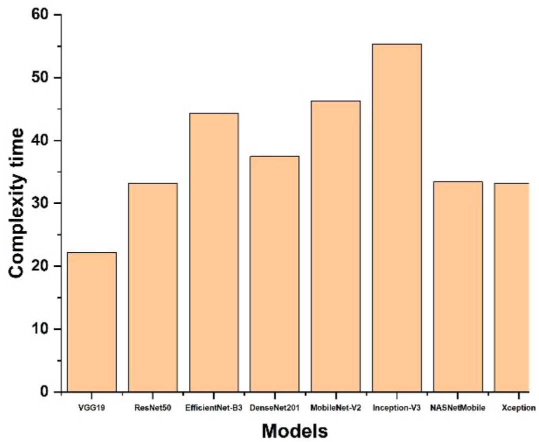

Fig. 4. Complexity time comparison of all models used in this study.

Medical Image Dataset. Future directions. Integrating tumour segmentation prior to classification, leveraging multi-modal imaging (e.g., MRI + CT), and adding attention mechanisms may further improve accuracy and stability. Larger, multi-centre datasets and standardized evaluation will be key to reliable deployment.

5. Conclusion

Deep learning models can classify meningioma on MRI with high accuracy. In our comparison, DenseNet201 gained the best accuracy (0.997%) compared to several models. Beyond accuracy, pairing classification with seep learning models offers interpretable outputs that can support radiological workflows. To strengthen clinical reliability, models should be trained on broader, more diverse cohorts and, where possible, multi-modal data (MRI, CT, PET). Hybrid systems that combine deep learning with classical image processing may add robustness. Future studies should extend comparative evaluations to other brain tumour types. Explore transfer learning for data-efficient training and incorporate expert-in-the-loop explainability to align model rationales with clinical judgment. These steps will help translate AI-assisted diagnosis into routine practice.

Conflict of interest

Ethical approval

Data availability

Funding statement

Author contribution

References

1. Ogasawara C, Philbrick BD, Adamson DC. Meningioma: A review of epidemiology, pathology, diagnosis, treatment, and future directions. *Biomedicines*, 2021;9(3):319. https://doi.org/10.3390/biomedicines9030319
2. Harter PN, Braun Y, Plate KH. Classification of meningiomas—advances controversies. *Chinese Clinical Oncology*, 2017;6(Suppl 1):S2–S2. https://doi.org/10.21037/cco.2017.05.02
3. Poshattiwar RS, Acharya S, Shukla S, Kumar S. Neurological manifestations of connective tissue disorders. *Cureus*, 2023;15(10). https://doi.org/10.7759/cureus.47108
4. Goldbrunner R, Stavrinou P, Jenkinson MD, Sahm F, Mawrin C, Weber DC, Weller M. EANO guideline on the diagnosis and management of meningiomas. *Neuro-oncology*, 2021;23(11):1821–1834. https://doi.org/10.1093/neuonc/noab150
5. Kahraman-Koytak P, Bruce BB, Peragallo JH, Newman NJ, Biousse V. Diagnostic errors in initial misdiagnosis of optic nerve sheath meningiomas. *JAMA Neurology*, 2019;76(3):326–332. https://doi.org/10.1001/jamaneurol.2018.3989
6. Saloner D, Uzelac A, Hetts S, Martin A, Dillon W. Modern meningioma imaging techniques. *Journal of neuro-oncology*, 2010;99:333–340. https://doi.org/10.1007/s11060-010-0367-6
7. Mohan G, Subashini MM. MRI based medical image analysis: Survey on brain tumor grade classification. *Biomedical Signal Processing and Control*, 2018;39:139–161. https://doi.org/10.1016/j.bspc.2017.07.007
8. Tsuneki M. Deep learning models in medical image analysis. *Journal of Oral Biosciences*, 2022;64(3):312–320. https://doi.org/10.1016/j.job.2022.03.003
9. Yu H, Yang LT, Zhang Q, Armstrong D, Deen MJ. Convolutional neural networks for medical image analysis: State-of-the-art, comparisons, improvement and perspectives. *Neurocomputing*, 2021;444:92–110. https://doi.org/10.1016/j.neucom.2020.04.157
10. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, 2016;770778. https://doi.org/10.1109/CVPR.2016.90
11. Simonyan K. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014, https://doi.org/10.48550/arXiv.1409.1556
12. Tan M, Le Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In International conference on machine learning, 2019; 6105–6114. PMLR. https://doi.org/10.48550/arXiv.1905.11946
13. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC. Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE conference on computer vision and pattern recognition, 2018;4510–4520. https://doi.org/10.48550/arXiv.1801.04381
14. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition, 2017;4700–4708.

15. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. In Proceedings of the IEEE conference on computer vision and pattern recognition, 2016;2818–2826. https://doi.org/10.1109/CVPR.2016.308
16. Menghani G. Efficient deep learning: A survey on making deep learning models smaller, faster, and better. *ACM Computing Surveys*, 2023;55(12):1–37. https://doi.org/10.1145/3578938
17. Yang J, Yang G. Modified convolutional neural network based on dropout and the stochastic gradient descent optimizer. *Algorithms*, 2018;11(3):28. https://doi.org/10.3390/a11030028
18. Ruby U, Yendapalli V. Binary cross entropy with deep learning technique for image classification. *Int. J. Adv. Trends Comput. Sci. Eng*, 2020;9(10). https://doi.org/10.30534/ijatcse/2020/175942020
19. Lopez-del Rio A, Nonell-Canals A, Vidal D, and Perera-Lluna A. Evaluation of cross-validation strategies in sequence-based binding prediction using deep learning. *Journal of chemical information and modeling*, 2019;59(4):1645–1657. https://doi.org/10.1021/acs.jcim.8b00663
20. Maharana K, Mondal S, and Nemade B. A review: Data pre-processing and data augmentation techniques. *Global Transitions Proceedings*, 2022;3(1):91–99. https://doi.org/10.1016/j.gltp.2022.04.020
21. Chaddad A, Peng J, Xu J, Bouridane A. Survey of explainable AI techniques in healthcare. Sensors, 2023;23(2):634. https://doi.org/10.3390/s23020634
22. Talaat FM, Gamel SA, El-Balka RM, Shehata M, ZainEldin H. GradCAM enabled breast cancer classification with a 3D inception-ResNet V2: Empowering radiologists with explainable insights. *Cancers*, 2024;16(21):3668. https://doi.org/10.3390/cancers16213668
23. Suara S, Jha A, Sinha P, Sekh AA. Is grad-cam explainable in medical images?. In International Conference on Computer Vision and Image Processing, 2023;124–135, Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-03158181-6_11
24. Shaheed K, Mao A, Qureshi I, Kumar M, Hussain S, Ullah I, Zhang X. DS-CNN: A pre-trained Xception model based on depth-wise separable convolutional neural network for finger vein recognition. Expert Systems with Applications, 2022;191:116288. https://doi.org/10.1016/j.eswa.2021.116288
25. Kateb Y, Khebli A, Meglouli H. Classification of surface defects in steel sheets using developed NasNet-mobile CNN and few samples. *Revue d'Intelligence Artificielle*, 2024;38(2). https://doi.org/10.18280/ria.380232
26. Saeedi S, Rezayi S, Keshavarz H, Niakan Kalhori R. MRI-based brain tumor detection using convolutional deep learning methods and chosen machine learning techniques. BMC Medical Informatics and Decision Making, 2023;23(1):16.
27. Abdusalomov AB, Mukhiddinov M, Whangbo TK. Brain tumor detection based on deep learning approaches and magnetic resonance imaging. *Cancers*, 2023;15(16):4172.
28. Huang ML, Liao YC. Stacking ensemble and ECA-EfficientNetV2 convolutional neural networks on classification of multiple chest diseases including COVID-19. *Academic Radiology*, 2023;30(9):1915–1935.
29. Chow LS, Tang GS, Solihin MI, Gowdh NM, Ramli N, Rahmat K. Quantitative and qualitative analysis of 18 deep convolutional neural network (CNN) models with transfer learning to diagnose COVID-19 on chest X-ray (CXR) images. *SN Computer Science*, 2023;4(2):141.
30. Ennab M, Mcheick H. Advancing AI interpretability in medical imaging: A comparative analysis of pixel-level interpretability and grad-CAM models. *Machine Learning and Knowledge Extraction*, 2025;7(1):12.
31. Zhu H, Wang W, Ulidowski I, Zhou Q, Wang S, Chen H, Zhang Y. MEEDNets: Medical image classification via ensemble bio-inspired evolutionary DenseNets. *Knowledge-Based Systems*, 2023;280:111035.
32. Okuboyejo DA, Olugbara OO. Classification of skin lesions using weighted majority voting ensemble deep learning. *Algorithms*, 2022;15(12):443.
33. Ahmed AE. A novel architecture of depthwise separable CNN and multi-level pooling for detection and classification of myopic maculopathy. International Journal of Advanced Computer Science & Applications, 2024;15(7).
34. Briouya H, Briouya A, Choukri A. Surveying lightweight neural network architectures for enhanced mobile performance. In *International Conference on Smart Applications and Data Analysis*, 2024;187–199. Cham: Springer Nature Switzerland.
35. Hosseini A, Eshraghi MA, Taami T, Sadeghsalehi H, Hoseinzadeh Z, Ghaderzadeh M, Rafiee M. A mobile application based on efficient lightweight CNN model for classification of B-ALL cancer from non-cancerous cells: A design and implementation study. Informatics in Medicine Unlocked, 2023;39:101244.
36. Du W, He Y, Li Y, Wu Z. Brain tumor diagnosis using EfficientNet. In *Second IYSF Academic Symposium on Artificial Intelligence and Computer Engineering*, 2021;12079:18–25. SPIE.
37. Solanki S, Singh UP, Chouhan SS, Jain S. Brain tumor detection and classification using intelligence techniques: an overview. *IEEE Access*, 2023;11:12870–12886.
38. Liu Y, Pu H, Sun DW. Efficient extraction of deep image features using convolutional neural network (CNN) for applications in detecting and analysing complex food matrices. *Trends in Food Science & Technology*, 2021;113:193–204.
39. Vassantachart A, Cao Y, Gribble M, Guzman S, Ye JC, Hurth K, Yang W. Automatic differentiation of grade I and II meningiomas on magnetic resonance image using an asymmetric convolutional neural network. *Scientific Reports*, 2022;12(1):3806.
40. Maki S, Furuya T, Horikoshi T, Yokota H, Mori Y, Ota J, Ohtori S. A deep convolutional neural network with performance comparable to radiologists for differentiating between spinal schwannoma and meningioma. *Spine*, 2020;45(10):694–700.
41. Ambrogio S, Narayanan P, Tsai H, Shelby RM, Boybat I, Di Nolfo C, Burr GW. Equivalent-accuracy accelerated neural-network training using analogue memory. *Nature*, 2018;558(7708):60–67.
42. Ennab M, Mcheick H. Advancing AI interpretability in medical imaging: A comparative analysis of pixel-level interpretability and grad-CAM models. *Machine Learning and Knowledge Extraction*, 2025;7(1):12.
43. Wang T, Roberts A, Hesslow D, Le Scao T, Chung HW, Beltagy I, Raffel C. What language model architecture and pretraining objective works best for zero-shot generalization?. In *International Conference on Machine Learning*, 2022;2296422984, PMLR.
44. Sadri AR, Janowczyk A, Zhou R, Verma R, Beig N, Antunes J, Viswanath SE. MRQy—An open-source tool for quality control of MR imaging data. *Medical physics*, 2020;47(12):6029–6038.
45. Mathivanan SK, Sonaimuthu S, Murugesan S, Rajadurai H, Shivahare BD, Shah MA. Employing deep learning and transfer learning for accurate brain tumor detection. *Scientific Reports*, 2024;14(1):7232.

46. Alam MA, Sohel A, Hasan KM, and Ahmad I. Advancing brain tumor detection using machine learning and artificial intelligence: A systematic literature review of predictive models and diagnostic accuracy. *Strategic Data Management and Innovation*, 2024;1(01):37–55.
47. Ennab M, Mcheick H. Advancing AI interpretability in medical imaging: A comparative analysis of pixel-level interpretability and grad-CAM models. *Machine Learning and Knowledge Extraction*, 2025;7(1):12.
48. Bento M, Fantini I, Park J, Rittner L, Frayne R. Deep learning in large and multi-site structural brain MR imaging datasets. *Frontiers in Neuroinformatics*, 2022;15:805669.
49. Saleh AW, Gupta G, Khan SB, Alkhaldi NA, and Verma A. An Alzheimer's disease classification model using transfer learning densenet with embedded healthcare decision support system. *Decision Analytics Journal*, 2023;9:10