Computer and Decision Making

An International Journal

www.comdem.org
eISSN: 3008-1416

Breakthroughs in Brain Tumor Detection: Leveraging Deep Learning and Transfer Learning for MRI-Based Classification

Alireza Golkarieh[1], Sajjad Rezvani Boroujeni[2], Kiana Kiashemshaki[3], Maryam Deldadehasl[4], Hamed Aghayarzadeh[5,*], Azita Ramezani[6]

[1] Department of Computer Science and Engineering, Oakland University, Rochester, Michigan, USA.
[2] Department of Applied Statistics & Operations Research, Bowling Green State University, USA.
[3] Department of Computer Science, Bowling Green State University, Bowling Green, USA
[4] School of Electrical, computer and biomedical Engineering, Southern Illinois university, Carbondale, USA.
[5] Department of Computer Engineering, Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran.
[6] Department of Computer Science, Cleveland State Cleveland State University Cleveland, Ohio, USA.

ARTICLE INFO

Article history:
Received 17 April 2025
Received in revised form 21 June 2025
Accepted 2 July 2025
Available online 31 July 2025

Keywords:

Brain Tumor Detection; Deep Learning; Transfer Learning; MRI-Based Classification; Convolutional Neural Networks

ABSTRACT

Identifying and classifying brain tumors play a pivotal role in gaining insights into their underlying mechanisms. In contemporary medical practice, the integration of Computer-assisted Diagnosis (CAD) and machine learning, particularly deep learning, has significantly enhanced the radiologist's ability to accurately identify brain tumors. Unlike traditional machine learning methods, which often rely on manual feature engineering for classification, deep learning models can be structured to prevent the need for manual feature extraction, yielding highly accurate classification outcomes. This paper customizes advanced deep learning models including VGG19, ResNet50, InceptionV3, and EfficientNetV2 as the most powerful deep learning models aimed at the identification of both binary (normal and abnormal) and multiclass: 17 classes including Glioma, Meningioma, Neurocytoma, and other types of injuries such as Abscesses and Cysts. We utilize a publicly available dataset containing 4449 MRI images. Subsequently, we conduct a comprehensive comparative analysis of our proposed models against existing models in the literature. Our experimental findings indicate that EfficientNetV2 outperforms other state-of-the-art deep-learning models. Our experimental findings demonstrate that EfficientNetV2 outperforms other state-of-the-art deep-learning models, achieving exceptional performance metrics: 98.22% accuracy, 99.1% precision, 99.26% recall, and a 99.17% F1-score for multiclass tumor classification.

1. Introduction

---

[*] *hghayarzadeh.csu@gmail.com*

Doi: https://doi.org/10.59543/comdem.v2i.14243

Brain tumors, comprising diverse types such as Glioma, Meningioma, and Neurocytoma, represent a complex spectrum of abnormal cell growth within the brain. Gliomas, derived from glial cells, are notorious for their infiltrative nature, making complete surgical removal challenging. These tumors are further classified into various subtypes, including astrocytoma and oligodendrogliomas, each with distinct characteristics and prognoses. Meningiomas, arising from the meninges, are typically slow-growing and often benign [1-3]. However, their location within the protective layers of the brain can lead to symptoms based on their size and proximity to critical structures. Neurocytomas, though relatively rare, pose unique challenges due to their origin from neural cells. These tumors are often found in the ventricles of the brain and may necessitate intricate surgical approaches for removal. The importance of early detection of these brain tumors cannot be overstated. Gliomas, particularly high-grade ones like glioblastomas, demand swift intervention to mitigate their aggressive progression. Meningiomas, although usually benign, can exert pressure on surrounding brain tissue, causing symptoms that range from headaches to neurological deficits. Neurocytomas, while less common, require early identification to determine the optimal treatment approach. Early detection enables healthcare professionals to explore a range of treatment modalities, including surgery, radiation therapy, and chemotherapy, tailored to the specific characteristics of each tumor. Early detection enables healthcare professionals to explore a range of treatment modalities, including surgery, radiation therapy, and chemotherapy, tailored to the specific characteristics of each tumor. Recent advances in adjuvant therapies, such as thymol-enhanced 5-fluorouracil cytotoxicity for esophageal cancer [23], novel pharmaceutical management strategies for precision oncology [24], cisplatin-loaded nano-polybutylcyanoacrylate for breast cancer [44], and liposomal nanoparticle-enhanced cisplatin delivery for oral cancer [45], highlight the growing role of targeted drug delivery systems in improving therapeutic outcomes. Similarly, combinatorial approaches like sonodynamic therapy enhanced with gold nanoparticles [37], nano-liposomal paclitaxel for gastric cancer [46], and antimicrobial peptides repurposed as oncotherapeutics [47] demonstrate significant potential for tumor reduction. Innovations in synthetic methodologies—such as base-mediated cyclization to access 1,4-oxazepine derivatives [42]—and herbal phytotherapeutics [48] expand the chemical space for future neuro-oncological drug discovery, highlighting the expanding frontier of precision cancer interventions. Moreover, early intervention is pivotal in preserving neurological function and preventing complications associated with advanced tumor stages. Patients benefit not only from improved treatment outcomes but also from a better quality of life with timely and targeted therapeutic interventions. Regular screenings, advanced imaging technologies, and ongoing research efforts contribute to enhancing early detection capabilities, underscoring the critical role of proactive medical monitoring in managing these complex brain tumors. A considerable number of individuals in the United States have received a diagnosis of primary brain tumors, totaling approximately 700,000 cases. Moreover, solely within the United States in the year 2021, around 85,000 new instances of brain tumors were identified. The age of a patient is just one among several variables impacting the prognosis and survival rates in the context of brain tumors. As per the findings cited in [4], individuals aged 55–64 exhibited a one-year survival rate of 46.1%, while those in the 65–74 age bracket had a survival rate of 29.3%. The research discussed in [5] underscores the significance of early detection of tumors in augmenting the chances of survival [3].

Magnetic Resonance Imaging (MRI) has emerged as a cornerstone in the detection and characterization of brain tumors, revolutionizing the field of neuroimaging. This non-invasive imaging modality utilizes powerful magnets and radiofrequency pulses to generate detailed and high-resolution images of the brain's anatomy. In the context of brain tumor detection, MRI provides unparalleled advantages, offering exceptional soft tissue contrast that allows for precise visualization of tumors and their surrounding structures. Gliomas, Meningiomas, Neurocytomas, and various other brain tumors can be accurately identified, located, and characterized through different MRI sequences. The multi-parametric nature of MRI, including T1-weighted, T2-weighted, and contrast-enhanced images, enables clinicians to assess tumor composition, vascularity, and potential invasion into adjacent tissues. Additionally, advanced MRI techniques such as diffusion-weighted imaging (DWI) and magnetic resonance spectroscopy (MRS) provide valuable insights into the cellular and metabolic aspects of brain tumors. The sensitivity of MRI in detecting even subtle abnormalities makes it an indispensable tool for early diagnosis, treatment planning, and ongoing monitoring of brain tumors, contributing significantly to improved patient outcomes and the advancement of neuro-oncology. [1]

Recent advancements in artificial intelligence have driven the increasing integration of machine learning and deep learning frameworks into image-based diagnostics, particularly within medical imaging research. These technologies now play a pivotal role in tasks such as tumor recognition, tissue classification, and anatomical segmentation, as evidenced by their growing adoption in medical image analysis [6,7]. Beyond diagnostics, deep learning has also demonstrated promise in dynamic biomechanical systems, such as engineered cardiac tissues, where adaptive control of contractile resistance is critical for functional analysis [25]. Deep neural networks, through their end-to-end learning paradigms, demonstrate an ability to independently extract hierarchical features from imaging data without manual intervention. Of particular clinical relevance is their application to brain tumor identification using MRI scans, where algorithmic precision directly impacts diagnostic accuracy. This study presents a rigorous comparative evaluation of four advanced convolutional neural networks—VGG19 [8], ResNet50 [9], Inception V3 [10], and EfficientNetV2 [11]—for multiclass brain tumor classification. Each architecture contributes unique strengths: VGG19's depth enables detailed feature extraction, ResNet50's skip connections mitigate gradient degradation, Inception V3's multi-scale processing enhances pattern recognition, and EfficientNetV2 balances accuracy with computational efficiency. Rather than relying on predefined features, these models autonomously learn discriminative patterns from pixel-level data, adapting their hierarchical representations to tumor morphology. To assess performance, we evaluate these models on a publicly available brain tumor dataset comprising 17 distinct pathological classes. Our methodology emphasizes not only classification accuracy but also architectural efficiency and generalizability across tumor subtypes. By benchmarking against contemporary standards, we identify critical trade-offs between model complexity, computational demands, and diagnostic precision—factors essential for clinical deployment.

Section 2 critically reviews seminal works in brain tumor analysis, focusing on evolving deep learning architectures and their diagnostic applications. Building on this foundation, Section 3 outlines our methodological framework, including a comprehensive breakdown of the dataset and the four deep learning models selected for comparative evaluation. Section 4 then analyzes experimental outcomes, benchmarking model performance through statistical metrics and robustness testing.

2. Related Works

Recent advances in neuroimaging diagnostics have intensified interest in machine learning frameworks for early brain tumor detection and classification. Abd-Ellah et al. [12] performed a systematic review of MRI-based diagnostic systems, contrasting traditional machine learning pipelines with deep neural networks to highlight their comparative advantages in accuracy and scalability. Expanding on this work, the authors later proposed a semi-automated segmentation framework optimized for brain tumor MRI analysis [13], leveraging T1-weighted (T1W) MRI sequences to enhance 3D volumetric precision. Parallel research by Somasundaram et al. [14] addressed preprocessing challenges in MR image analysis, benchmarking performance metrics across 3D CNNs, SVMs, and multiclass SVM variants for tumor boundary delineation. Deep learning architectures, particularly CNNs, now dominate medical image processing due to their hierarchical feature extraction capabilities—a critical advantage in detecting subtle tumor morphology. For instance, Latif et al. [15] developed a hybrid architecture combining deep learning-derived radiomic features with ensemble classifiers like Naïve Bayes, Multilayer Perceptrons, and SVMs. Notably, when paired with SVMs, their model achieved 96% classification accuracy, demonstrating the synergy between learned representations and classical machine learning. The diversity of tumor types analyzed in this study, including gliomas, meningiomas, and neurocytomas (Figure 1), underscores the need for adaptable deep learning architectures capable of distinguishing subtle morphological variations across pathological classes.

Recent methodological innovations in brain tumor diagnostics highlight the growing role of multimodal imaging integration. Yahyaoui et al. [16] pioneered a fusion framework for 2D and 3D MRI analysis, pairing DenseNet for classification with specialized 3D CNNs for cross-modal segmentation. Their approach demonstrated robust generalization, attaining 92% classification accuracy with DenseNet and 85% segmentation precision using task-specific 3D architectures. Parallel advancements in reinforcement learning, such as secrecy rate optimization in smart grid communications through improved IChOA-based frameworks [26], illustrate the broader applicability of adaptive machine learning systems in critical infrastructure domains, suggesting potential synergies with medical imaging optimization. Parallel advancements in foundation models, such as SAM-based architectures for precise iris segmentation [27], demonstrate the broader potential of adaptable deep learning systems in medical image analysis, particularly for fine-grained anatomical delineation. Similarly, machine learning continues to expand into diverse domains, including cybersecurity for intrusion detection in mobile social networks [35] and business economics for analyzing multifactorial impacts on financial systems [36]. Beyond medical imaging, NLP-enabled automated assessment frameworks show promise in reducing linguistic biases in scientific evaluation [39], highlighting AI's potential to enhance objectivity in multimodal research contexts. Parallel advances by Kang et al. [17] focused on optimizing classifier selection, combining deep CNN-derived features with machine learning ensembles. Their empirical evaluation across heterogeneous datasets revealed radial basis function SVMs to outperform both conventional classifiers and end-to-end deep learning models, suggesting hybrid architectures may bridge domain adaptability gaps. Beyond tumor diagnostics, machine learning has demonstrated efficacy in diverse neurological applications, including depression detection through EEG signal analysis using optimized time-frequency feature extraction [28], and investigations into chronic tobacco smoking’s impact on electrical brain activity [29], highlighting its versatility across neuroimaging and neurobehavioral studies.

In glioblastoma grading, Bhatele and Bhadauria [18] implemented an automated extreme gradient boosting (XGBoost) system for CNS tumor stratification, achieving 90% accuracy in high-grade glioma detection and 95% in low-grade cases—a critical improvement for clinical decision-making. Meanwhile, Mehrotra et al. [19] systematically evaluated transfer learning strategies for tumor malignancy detection, benchmarking Adam, RMSprop, and SGD optimizers across pre-trained CNNs. Their findings underscored the importance of architecture-task alignment, with a fine-tuned AlexNet configuration surpassing newer models in medical imaging contexts.

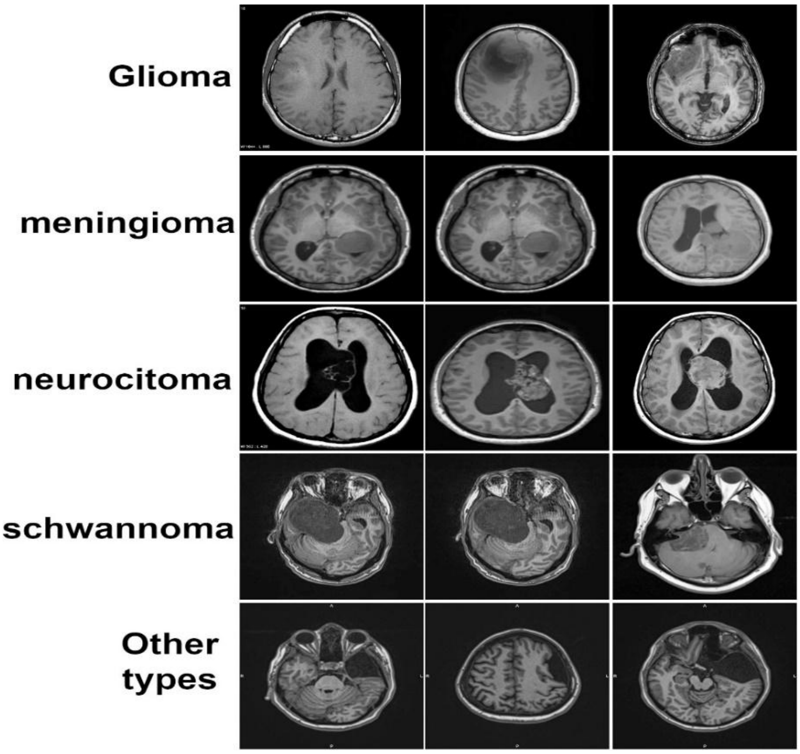


Fig. 1. Different types of tumors in this study

Evolutionary approaches also show promise: Anaraki et al. [20] developed a neuroevolutionary framework using genetic algorithms to autonomously design CNN architectures for glioblastoma detection. This self-optimizing system achieved 90.9% accuracy in glioma identification and 94.2% in pituitary tumor classification, demonstrating the viability of meta-learning for medical image analysis.

3. Methodology

3.1. *Brain Tumor Dataset*

The provided database consists of 4,449 authentic magnetic resonance imaging (MRI) images of the skull, captured in axial planes, and weighted in T1, T1 with contrast, and T2. The images are meticulously organized into folders based on their distinct types or categories, comprising a total of 17 classes [21], employing strategies akin to tailored partitioning techniques for healthcare big data management [30] to ensure efficient retrieval and processing. These examinations have been interpreted by qualified radiologists and are made available for study and research purposes. These classes include:

1. Glioma (Subtypes: Astrocytoma, Ganglioglioma, Glioblastoma, Oligodendroglioma, Ependymoma)
2. Meningioma (Subtypes: Low Grade, Atypical, Anaplastic, Transitional)
3. Neurocytoma (Subtypes: Central - Intraventricular, Extra ventricular)
4. NORMAL
5. Other Types of Injuries (Including Abscesses, Cysts, Miscellaneous Encephalopathies)
6. Schwannoma (Subtypes: Acoustic, Vestibular - Trigeminal)

It's important to note that these images are genuine samples obtained from real examinations and have been categorized without revealing any patient-specific medical records, ensuring the preservation of patient identity. These examinations have been interpreted by qualified radiologists and are made available for study and research purposes.

3.2. Deep Learning Models

To address the challenges of precise brain tumor diagnostics, this study leverages four convolutional neural networks—VGG19, ResNet50, Inception V3, and EfficientNetV2—selected for their complementary strengths in hierarchical feature extraction. Rather than deploying these architectures as static feature extractors, we implemented a two-stage adaptive fine-tuning strategy. First, we replaced each model's final classification layer with a task-specific binary classifier optimized for tumor/non-tumor discrimination. Second, we trained the composite architecture end-to-end, allowing both pretrained base layers and the novel classifier to co-adapt during gradient optimization. This approach ensures the model retains generalizable visual representations while refining domain-specific decision boundaries critical for medical imaging.

The fine-tuning process proved particularly vital for balancing transfer learning efficiency with diagnostic precision. By selectively adjusting layer-wise learning rates, we preserved low-level edge detectors in early convolutional blocks while intensively training higher layers to recognize tumor morphology patterns. Such targeted parameter optimization enables robust performance across heterogeneous MRI datasets without requiring exhaustive retraining.

3.2.1. Pre-trained models

VGG-19 adopts a layered hierarchy of 19 weight-bearing stages—16 convolutional and 3 fully connected layers—prioritizing simplicity through uniform 3×3 kernels (stride=1, padding=1) to balance parameter efficiency with spatial feature extraction. Its architecture reinforces VGGNet's foundational principle: depth enhances hierarchical visual interpretation, exemplified by 2×2 max-pooling layers (stride=2) that progressively abstract tumor morphology.

ResNet50 advances this paradigm with 50-layer residual blocks, integrating skip connections to bypass vanishing gradients—a critical innovation for training stability. By preserving gradient flow across 60.8 million parameters, ResNet50 captures subtle textural anomalies in noisy medical imaging data, cementing its role in precision-driven diagnostics.

Inception V3 diverges through multi-scale processing, deploying parallel 48-layer pathways with factorized convolutions (asymmetric and symmetric) to isolate tumor features across spatial resolutions. Google's pre-trained implementation—refined via Adam optimization and batch-normalized inputs—leverages 20 million parameters to balance specificity and generalization, with Softmax-driven classification ensuring robust probabilistic outputs.

EfficientNetV2 redefines efficiency, combining neural architecture search (NAS) with compound scaling to optimize its 25.7 million parameters. By expanding convolutional block diversity while constraining computational overhead, it achieves state-of-the-art accuracy with reduced training latency—an asset for resource-constrained clinical environments. Recent advances in robust subgraph learning [40] and information-theoretic graph summarization [41] offer promising avenues for optimizing feature hierarchies in non-IID medical data, potentially enhancing model generalizability. The architectural distinctions between VGG19, ResNet50, and Inception V3—visually contrasted in Figure 2—directly inform their hierarchical feature extraction strategies, with VGG19's uniform 3×3 convolutions enabling spatial coherence, ResNet50's skip connections preserving gradient fidelity, and Inception V3's multi-scale pathways capturing tumor heterogeneity. Each architecture addresses distinct diagnostic challenges:

- VGG-19: Prized for structural uniformity and interpretability.
- ResNet50: Excels in gradient preservation for deep feature learning.
- Inception V3: Optimizes multi-scale feature fusion.
- EfficientNetV2: Balances accuracy with deployability.

Model selection hinges on task-specific trade-offs: computational budgets, dataset complexity, and interpretability needs. While VGG-19 and Inception V3 suit exploratory studies requiring transparency, ResNet50 and EfficientNetV2 dominate scenarios demanding robustness to noise or rapid inference.

3.2.2. Customized network

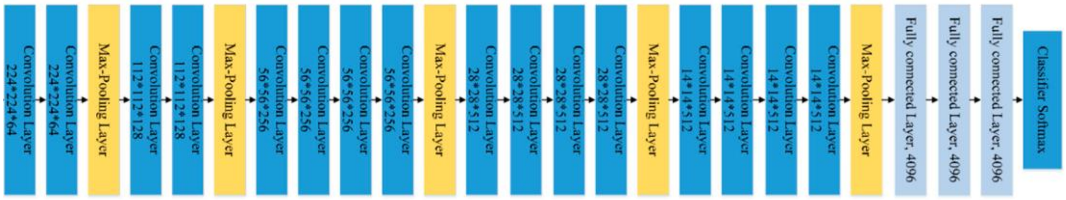
To address the complexity of multi-class tumor classification, we adopted a transfer learning framework leveraging pretrained CNNs as adaptive feature extractors. The fine-tuning protocol involved two critical modifications: first, replacing each model's final classification layer with a custom classification head engineered for 17-class differentiation, and second, training the hybrid architecture end-to-end to refine both pretrained backbone layers and the novel classifier. This dual-phase optimization ensures the networks retain generalized visual representations while adapting to domain-specific tumor morphology patterns.

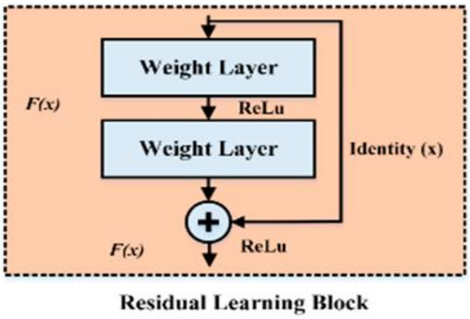
The redesigned classification layer—architected as a task-specific fully connected network—serves as a discriminative bottleneck, transforming high-dimensional CNN features into probabilistic predictions across 17 tumor subtypes. During training, we optimized the composite model (pretrained layers + classifier) via backpropagation, allowing gradient updates to fine-tune early convolutional filters for tumor edge detection while intensively training deeper layers for pathological feature abstraction. This hierarchical parameter adjustment proved essential for balancing model plasticity (adapting to new data) and stability (preserving pretrained knowledge), particularly given the dataset's inter-class heterogeneity.

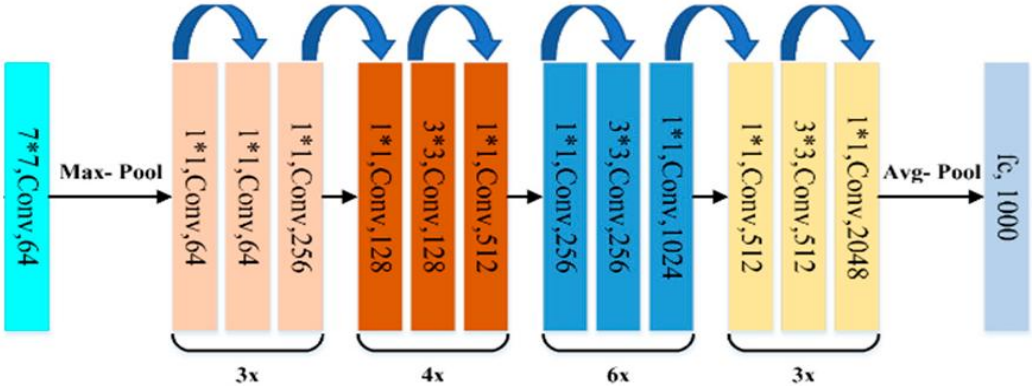
While originally designed for binary classification, our framework demonstrates scalability to multiclass settings through layer-wise learning rate modulation—a strategy that mitigates catastrophic forgetting while enhancing diagnostic precision.

3.3. Parameter Setting

We implemented the proposed deep learning framework using Keras, with all experiments conducted on a workstation equipped with an Intel i7-6900 CPU (3.7GHz, 8 cores), 32GB RAM, and an NVIDIA GeForce RTX 3090 GPU. The training protocol employed a margin setting of 0.2, randomized sampling, and 100 epochs to balance computational efficiency with convergence stability. For validation, we utilized a publicly available brain tumor MRI dataset comprising 17 distinct pathological classes.







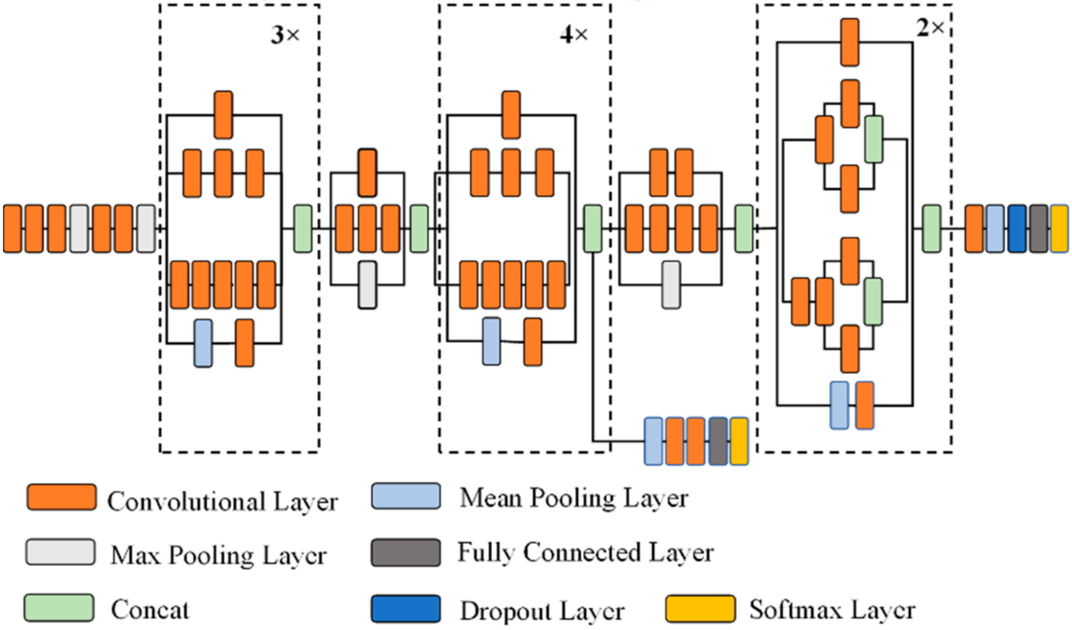


Fig. 2. Top: VGG19, Middle: ResNet 50, Down: Inception V3 [22]

3.4. Performance Metrics

In evaluating the classification performance, we rely on a set of core metrics directly involving the model's outcomes, encompassing true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). True Positives (TP) indicate instances correctly classified as the positive class, representing the accurate identification of brain tumor disease cases. TP reflects the model's successful recognition of actual instances of the condition. True Negatives (TN) signify instances correctly classified as the negative class, showcasing the model's accurate identification of cases without brain tumor disease.

False Positives (FP) quantify instances erroneously classified as positive when they are, in reality, negative for the condition, revealing the model's tendency to produce incorrect positive predictions. False Negatives (FN) measure instances erroneously classified as negative when they should be positive, indicating the model's shortfall in capturing actual positive cases. Analyzing the interplay of TP, TN, FP, and FN provides a detailed and precise understanding of the model's performance in distinguishing between brain tumor disease and its absence. These metrics allow a nuanced assessment of the model's effectiveness, considering both correct classifications and errors or omissions in the classification process.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F_{1_{score}} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{4}$$

4. Results

Table 1 presents a summary of the deep learning models applied in this study, providing insights into their architectural intricacies and scale. The table outlines crucial details, such as the number of convolutional layers and the total parameter count measured in millions. VGG19, recognized for its simplicity and uniform structure, encompasses 19 convolutional layers with a substantial 143 million parameters. In contrast, ResNet50 stands out for its impressive depth with 50 convolutional layers, maintaining a relatively compact parameter count of 23 million. InceptionV3, acknowledged for its versatility, features 48 convolutional layers and 21 million parameters, highlighting its capability to capture features at diverse scales. The EfficientNetV2 family, represented by B0, illustrates its computational efficiency with an impressive 237 convolutional layers and a remarkably modest parameter count of 25 million. This table serves as a quick reference to the scale and complexity of these deep learning models, laying the groundwork for their application in the subsequent analysis.

Tables 2 and 3 offer an in-depth overview of our analysis, where we systematically assess the performance of diverse models in the domain of brain tumor disease detection, utilizing a varied set of metrics. Table 2 presents the outcomes derived from the utilization of pre-trained models on the MRI Dataset, while Table 3 provides a detailed insight into the effectiveness of our specifically devised model. In Table 3, the unique achievements of each model become evident, enabling a thorough examination of their respective performances. In the evaluation of brain tumor disease detection, VGG19 achieves a substantial accuracy of 90.22%, demonstrating proficiency in accurately categorizing patches. Its commendable precision score of 95.5% showcases its ability to minimize false positive classifications. With a recall rate of 89.1%, it effectively identifies genuine brain tumor disease patches, resulting in a balanced F1 score of 92.18. ResNet50 stands out with an impressive accuracy of 94.27% and a remarkable precision of 96.85%, indicating effectiveness in mitigating false positives. A recall rate of 92.68% demonstrates the model's capability to capture authentic brain tumor disease patches, leading to an F1 score of 94.73. InceptionV3 delivers commendable results with an accuracy of 93.15% and a precision of 97.39%, slightly lower in recall at 93.36. The resulting F1 score, at 95.45, effectively balances precision and recall. Notably, EfficientNetV2 excels in this evaluation, achieving an exceptional accuracy of 98.22% and an exceptionally high precision of 99.1%, signifying proficiency in reducing false positives. With a recall rate of 99.26%, it captures the most authentic brain tumor disease patches, resulting in an outstanding F1 score of 99.17. This comprehensive analysis emphasizes the extraordinary performance of EfficientNetV2 in the realm of brain tumor disease detection.

In conclusion, in the evaluation of brain tumor disease detection, EfficientNetV2 emerges as the top-performing model across all metrics. It achieves the highest accuracy at 98.22%, the highest precision at 99.1%, the highest recall at 99.26%, and the highest F1 score at 99.17%. This signifies EfficientNetV2's exceptional proficiency in reducing false positives, capturing genuine brain tumor disease patches, and maintaining an outstanding balance between precision and recall. Following closely, ResNet50 also demonstrates strong performance.

4.1. Statistically Analysis

To inspect potential statistically noteworthy distinctions in the classification performance metrics across the models, we can employ statistical approaches like Analysis of Variance (ANOVA), crafted for simultaneous comparison of multiple groups. ANOVA empowers us to investigate whether substantial differences exist within the means of the diverse models. The p-values derived from ANOVA will serve as indicators to establish if meaningful variations are present among the models for each metric. Should the p-values descend below a predetermined significant threshold (e.g., 0.05), it enables us to infer that indeed there are notable differences among the models.

Table 1

Deep Learning Models: Number of layers and parameters.

Model	Number of Convolutional Layers	Number of Parameters (Millions)
VGG19 model	19	143
ResNet50 model	50	23
InceptionV3 model	48	21
EfficeinNetV2 model	237 (B0)	25

Table 2

Comparison of the different models: the pre-trained models (without training)

Model	Accuracy	Precision	Recall	$F1_{score}$
VGG19 model	0.881±0.0057	0.912±0.0022	0.848±0.019	0.878±0.0027
ResNet50 model	0.942±0.010	0.948±0.0033	0.921±0.0055	0.934±0.0036

InceptionV3 model	0.920±0.0021	0.947±0.0071	0.945±0.062	0.945±0.0085
EfficientNetV2 model	0.954±0.0029	0.989±0.0036	0.991±0.0027	0.989±0.0071

Table 3
Comparison of the different models: fine tuning the pre-trained models.

Model	Accuracy	Precision	Recall	F1$_{score}$
VGG19 model	0.902±0.0030	0.955±0.0010	0.891±0.0074	0.921±0.0065
ResNet50 model	0.942±0.0067	0.968±0.0032	0.926±0.0044	0.947±0.017
InceptionV3 model	0.931±0.0031	0.973±0.0096	0.933±0.0057	0.954±0.00074
EfficientNetV2 model	0.982±0.0038	0.991±0.0039	0.992±0.0033	0.991±0.0040

The analysis of variance (ANOVA) results in notable p-values, indicating substantial differences among the models concerning performance metrics. Specifically, the p-value for Accuracy is 2.9e-03, signaling a significant divergence in accuracy among the models. Similarly, the Precision metric yields a p-value of 4.7e-03, reaffirming a significant difference in precision across the models. Additionally, the Recall metric exhibits a small p-value of 7.4e-04, emphasizing a notable variance in recall among the models. Lastly, the F1 Score registers a p-value of 4.7e-03, confirming a significant difference in the F1 Score across the models. In summary, the statistical analysis highlights significant performance disparities among the models across all four metrics, underscoring the impactful role of model selection in influencing classification performance.

5. Conclusion

Our investigation systematically evaluated four fine-tuned deep learning architectures for detecting brain tumors in MRI scans, each enhanced with tailored fully connected layers. Among these models, EfficientNetV2 emerged as the superior performer, achieving exceptional diagnostic metrics: 98.2% accuracy, 99.1% precision, 99.2% recall, and a harmonized F1 score of 99.1%. These results highlight its capability to reliably identify tumor pathology while striking an optimal balance between sensitivity and specificity—a critical advantage for clinical applications.

The robust performance of EfficientNetV2 underscores its potential as a scalable tool for medical imaging, particularly in settings demanding high diagnostic precision. Future work could explore its integration into multimodal diagnostic pipelines to further enhance clinical decision-making. Extending these frameworks to generate 3D tumor reconstructions from 2D MRI slices, akin to point cloud estimation methods applied in satellite imagery analysis [31], could improve volumetric assessments for surgical planning. Additionally, incorporating environmental risk factors, such as air pollution exposure linked to CNS disease outcomes [32], and lifestyle interventions like nutritional supplementation for mitigating physiological stressors [33], could enable more holistic prognostic models that account for clinical, ecological, and behavioral determinants of tumor progression. Parallel investigations into material durability for instance, the impact of alcohol-containing mouthwashes on dental composites [43] may also inform supportive care protocols for patients undergoing neurotoxic treatments. Privacy-preserving frameworks like federated learning [38] could further enable secure multi-institutional collaboration while addressing data sensitivity constraints in medical imaging. Furthermore, combining such diagnostic models with robotic-assisted systems, as demonstrated in rehabilitation contexts through electromyography-driven detection of wrist and hand movements [34], could expand their utility in surgical planning and patient-specific treatment delivery.

Author Contributions

Conceptualization, A.G., S.R.B., K.K., M.D., H.A., A.R.; methodology, A.G., S.R.B., K.K., M.D., H.A., A.R.; software, A.G., S.R.B., K.K., M.D., H.A., A.R.; validation, A.G., S.R.B., K.K., M.D., H.A., A.R.; formal analysis, A.G., S.R.B., K.K., M.D., H.A., A.R.; investigation, A.G., S.R.B., K.K., M.D., H.A., A.R.; writing—original draft preparation, A.G., S.R.B., K.K., M.D., H.A., A.R.; visualization, A.G., S.R.B., K.K., M.D., H.A., A.R.; supervision, A.G., S.R.B., K.K., M.D., H.A., A.R.

Funding

This research received no external funding

Data Availability Statement

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgement

This research was not funded by any grant

References

[1] Gore, D. V., & Deshpande, V. (2020, June). Comparative study of various techniques using deep Learning for brain tumor detection. In *2020 International conference for emerging technology (INCET)* (pp. 1-4). IEEE. https://doi.org/10.1109/INCET49848.2020.9154030

[2] Sapra, P., Singh, R., & Khurana, S. (2013). Brain tumor detection using neural network. *International Journal of Science and Modern Engineering (IJISME) ISSN*, 2319-6386.

[3] Abdusalomov, A. B., Mukhiddinov, M., & Whangbo, T. K. (2023). Brain tumor detection based on deep learning approaches and magnetic resonance imaging. *Cancers*, 15(16), 4172. https://doi.org/10.3390/cancers15164172

[4] Yavuz, B. B., Kanyilmaz, G., & Aktan, M. (2021). Factors affecting survival in glioblastoma patients below and above 65 years of age: A retrospective observational study. *Indian Journal of Cancer*, *58*(2), 210-216. https://doi.org/10.4103/ijc.ijc_36_19

[5] Adel Fahmideh, M., & Scheurer, M. E. (2021). Pediatric brain tumors: descriptive epidemiology, risk factors, and future directions. Cancer Epidemiology, *Biomarkers & Prevention*, 30(5), 813-821. https://doi.org/10.1158/1055-9965.EPI-20-1443

[6] Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., & Terzopoulos, D. (2021). Image segmentation using deep learning: A survey. *IEEE transactions on pattern analysis and machine intelligence*, 44(7), 3523-3542. https://doi.org/10.1109/tpami.2021.3059968

[7] Zadeh, F. S., Molani, S., Orouskhani, M., Rezaei, M., Shafiei, M., & Abbasi, H. (2023). Generative Adversarial Networks for Brain Images Synthesis: A Review. arXiv preprint arXiv:2305.15421. https://doi.org/10.48550/arXiv.2305.15421

[8] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556. https://doi.org/10.48550/arXiv.1409.1556

[9] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *In Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

[10] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. *In Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2818-2826). https://doi.org/10.1109/CVPR.2016.308

[11] Tan, M., & Le, Q. (2021, July). Efficientnetv2: Smaller models and faster training. In *International conference on machine learning* (pp. 10096-10106). PMLR. https://doi.org/10.48550/arXiv.2104.00298

[12] Abd-Ellah, M. K., Awad, A. I., Khalaf, A. A., & Hamed, H. F. (2019). A review on brain tumor diagnosis from MRI images: Practical implications, key achievements, and lessons learned. *Magnetic resonance imaging*, *61*, 300-318. https://doi.org/10.1016/j.mri.2019.05.028

[13] Yushkevich, P. A., Gao, Y., & Gerig, G. (2016, August). ITK-SNAP: An interactive tool for semi-automatic segmentation of multi-modality biomedical images. In *2016 38th annual international conference of the IEEE engineering in medicine and biology society (EMBC)* (pp. 3342-3345). IEEE. https://doi.org/10.1109/EMBC.2016.7591443

[14] Somasundaram, S., & Gobinath, R. (2019, February). Current trends on deep learning models for brain tumor segmentation and detection–a review. In *2019 International conference on machine learning, big data, cloud and parallel computing (COMITCon)* (pp. 217-221). IEEE. https://doi.org/10.1109/COMITCon.2019.8862209

[15] Latif, G., Ben Brahim, G., Iskandar, D. A., Bashar, A., & Alghazo, J. (2022). Glioma Tumors' classification using deep-neural-network-based features with SVM classifier. *Diagnostics*, *12*(4), 1018. https://doi.org/10.3390/diagnostics12041018

[16] Yahyaoui, H., Ghazouani, F., & Farah, I. R. (2021, July). Deep learning guided by an ontology for medical images classification using a multimodal fusion. In *2021 International congress of advanced technology and engineering (ICOTEN)* (pp. 1-6). IEEE. https://doi.org/10.1109/ICOTEN52080.2021.9493469

[17] Kang, J., Ullah, Z., & Gwak, J. (2021). MRI-based brain tumor classification using ensemble of deep features and machine learning classifiers. *Sensors*, *21*(6), 2222. https://doi.org/10.3390/s21062222

[18] Bhatele, K. R., & Bhadauria, S. S. (2022). Machine learning application in glioma classification: review and comparison analysis. *Archives of Computational Methods in Engineering*, *29*(1), 247-274. https://doi.org/10.1007/s11831-021-09572-z

[19] Mehrotra, R., Ansari, M. A., Agrawal, R., & Anand, R. S. (2020). A transfer learning approach for AI-based classification of brain tumors. *Machine Learning with Applications*, *2*, 100003. https://doi.org/10.1016/j.mlwa.2020.100003

[20] Anaraki, A. K., Ayati, M., & Kazemi, F. (2019). Magnetic resonance imaging-based brain tumor grades classification and grading via convolutional neural networks and genetic algorithms. *biocybernetics and biomedical engineering*, *39*(1), 63-74. https://doi.org/10.1016/j.bbe.2018.10.004

[21] https://www.kaggle.com/datasets/fernando2rad/brain-tumor-mri-images-17-classes

[22] Ali, L., Alnajjar, F., Jassmi, H. A., Gocho, M., Khan, W., & Serhani, M. A. (2021). Performance evaluation of deep CNN-based crack detection and localization techniques for concrete structures. *Sensors*, *21*(5), 1688. https://doi.org/10.3390/s21051688

[23] Pouyamanesh, G., Ameli, N., Metanat, Y., Khorrami, A., Abbasinezhad-Moud, F., Qoorchi Moheb Seraj, F., Ferns, G.A. and Bahrami, A. (2024). Thymol Enhances 5-Fluorouracil Cytotoxicity by Reducing Migration and Increasing Apoptosis and Cell Cycle Arrest in Esophageal Cancer Cells: An In-vitro Study. Indian Journal of Clinical Biochemistry, pp.1-12. https://doi.org/10.1007/s12291-024-01219-7

[24] Alishahi, F., Soudmand, N., Goki, T. G., & Rashidoleslami, T. S. (2025). Optimal Pharmaceutical Management Strategies in Cancer Treatment: Novel Approaches. Asian Pacific Journal of Cancer Nursing, 20250308-20250308. https://doi.org/10.31557/apjcn.1740.20250308

[25] Schuftan D, Kooh YKG, Guo J, Sun Y, Aryan L, Stottlemire B, Berkland C, Genin GM, Huebsch N. Dynamic control of contractile resistance to iPSC-derived micro-heart muscle arrays. J Biomed Mater Res A. 2024 Apr;112(4):534-548. https://doi.org/10.1002/jbm.a.37642

[26] Shoeibi, M., Nevisi, M. M. S., Khatami, S. S., Martín, D., Soltani, S., & Aghakhani, S. (2024). Improved IChOA-Based Reinforcement Learning for Secrecy Rate Optimization in Smart Grid Communications. Computers, Materials and Continua, 81(2), 2819-2843. https://doi.org/10.32604/cmc.2024.056823

[27] Farmanifard, P., & Ross, A. (2024, June). Iris-SAM: Iris Segmentation Using a Foundation Model. In International Conference on Pattern Recognition and Artificial Intelligence (pp. 394-409). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-97-8702-9_27

[28] Rafiei, M., Pour, M. R., & Akbari, H. (2024). Comparing linear, nonlinear and time frequency-based features in the EEMD domain for depression detection application using EEG signals. Signal Processing and Renewable Energy, 8(4), 1-11. https://doi.org/10.57647/j.spre.2024.0804.19

[29] Taebi, M., Taghavizanjani, F., Parsaei, M., Ershadmanesh, M., Beikmarzehei, A., Gorjestani, O., Rezaei, Z., Hasanzadeh, A. and Moghaddam, H.S. (2025). Chronic Effects of Tobacco Smoking on Electrical Brain Activity: A Systematic Review on Electroencephalography Studies. Behavioural Brain Research, 115479. https://doi.org/10.1016/j.bbr.2025.115479

[30] Soltanmohammadi, E., Dilek, A., & Hikmet, N. (2024). Tailored Partitioning for Healthcare Big Data: A Novel Technique for Efficient Data Management and Hash Retrieval in RDBMS Relational Architectures. Journal of Data Analysis and Information Processing, 13(1), 46-65. https://doi.org/10.4236/jdaip.2025.131003

[31] Rezaei, Y., & Lee, S. (2022). sat2pc: Estimating point cloud of building roofs from 2d satellite images. arXiv preprint arXiv:2205.12464. https://doi.org/10.48550/arXiv.2205.12464

[32] Karami, M., Hamzehei, S., Rastegari, F., & Akbarzadeh, O. (2023, July). Exploring the Relationship Between Air Pollution and CNS Disease Mortality in Italy: A Forecasting Study with ARIMA and XGBoost. In 2023 Congress in Computer Science, Computer Engineering, & Applied Computing (CSCE) (pp. 46-52). IEEE. https://doi.org/10.1109/CSCE60160.2023.00013

[33] Dehghani, E., Beba, M., Danandeh, K., Memari, A., Ershadmanesh, M.J., Rasoulian, P., Danandeh, A. and Djafarian, K. (2025). The effect of tart cherry juice (TCJ) supplementation on exercise-induced muscle damage (EIMD) in an athletic population. Annals of Medicine and Surgery, 87(2), 880-890. https://doi.org/10.1097/ms9.0000000000002914

[34] Savoji, K., Soleimani, M., & Moshayedi, A. J. (2024). A Comprehensive Review of Electromyography in Rehabilitation: Detecting Interrupted Wrist and Hand Movements with a Robotic Arm Approach. EAI Endorsed Transactions on AI and Robotics, 3. https://doi.org/10.4108/airo.7377

[35] Rivandi, E., & Jamili Oskouie, R. (2024). A Novel Approach for Developing Intrusion Detection Systems in Mobile Social Networks. Available at SSRN 5174811. https://dx.doi.org/10.2139/ssrn.5174811

[36] Bagherabad, M. B., Rivandi, E., & Mehr, M. J. (2025). Machine Learning for Analyzing Effects of Various Factors on Business Economic. Authorea Preprints. https://doi.org/10.36227/techrxiv.174429010.09842200/v1

[37] Neshastehriz, A., Hormozi-Moghadda, Z., Amini, S. M., Taheri, S. M., & Kichi, Z. A. (2024). Combined sonodynamic therapy and X-ray radiation with methylene blue and gold nanoparticles coated with apigenin: Impact on MCF7 cell viability. International Journal of Radiation Research, 22(2), 509-513. http://dx.doi.org/10.61186/ijrr.22.2.515

[38] Mahdimahalleh, S. E. (2024). Revolutionizing Wireless Networks with Federated Learning: A Comprehensive Review. European Journal of Electrical Engineering and Computer Science, 8(6), 33-45. https://doi.org/10.24018/ejece.2024.8.6.671

[39] Kim, C., Passonneau, R. J., Lee, E., Sheikhi Karizaki, M., Gnesdilow, D., & Puntambekar, S. (2025). NLP-enabled automated assessment of scientific explanations: Towards eliminating linguistic discrimination. British Journal of Educational Technology, 00, 1–33. https://doi.org/10.1111/bjet.13596

[40] Neshatfar, S., & Sekeh, S. Y. (2024). Robust Subgraph Learning by Monitoring Early Training Representations. arXiv preprint arXiv:2403.09901. https://doi.org/10.48550/arXiv.2403.09901

[41] Neshatfar, S., Magner, A., & Sekeh, S. Y. (2023). Promise and Limitations of Supervised Optimal Transport-Based Graph Summarization via Information Theoretic Measures. IEEE Access, 11, 87533-87542. https://doi.org/10.1109/ACCESS.2023.3302830

[42] Yazzaf, R., Asadi, M., & Mahdavi, M. Base Mediated 7-exo-dig Intramolecular Cyclization of Betti-propargyl Precursors: An Efficient Approach to 1, 4-oxazepine Derivatives. Current organic synthesis. Epub ahead of print. PMID: 39950461. https://doi.org/10.2174/0115701794353226241209175136.

[43] Ayatollahi, S., Davoudi, A., & Momtazi, H. (2025). In vitro comparative effects of alcohol-containing and alcohol-free mouthwashes on surface roughness of bulk-fill composite resins. BMC Research Notes, 18(1), 146. https://doi.org/10.1186/s13104-025-07213-3

[44] Gorgzadeh, A., Hheidari, A., Ghanbarikondori, P., Arastonejad, M., Goki, T.G., Aria, M., Allahyartorkaman, A., Moazzam, F., Velisdeh, Z.J., Saberian, E. and Habiba, A. (2023). Investigating the properties and cytotoxicity of cisplatin-loaded nano-polybutylcyanoacrylate on breast cancer cells. Asian Pacific Journal of Cancer Biology, 8(4), 345-350. https://doi.org/10.31557/apjcb.2023.8.4.345-350

[45] Ghanbarikondori, P., Aliakbari, R. B. S., Saberian, E., Jenča, A., Petrášová, A., Jenčová, J., & Khayavi, A. A. (2024). Enhancing cisplatin delivery via liposomal nanoparticles for oral cancer treatment. Indian Journal of Clinical Biochemistry, 1-7. https://doi.org/10.1007/s12291-024-01239-3

[46] Heidari, Z.A.C., Ghanbarikondori, P., Mamaghani, E.M., Hheidari, A., Saberian, E., Mozaffari, E., Alizadeh, M. and Allahyartorkaman, M. (2023). Characteristics and cytotoxic effects of nano-liposomal paclitaxel on gastric cancer cells. Asian Pacific journal of cancer prevention: APJCP, 24(9), p.3291. https://doi.org/10.31557/APJCP.2023.24.9.3291

[47] Zare-Zardini, H., Saberian, E., Jenča, A., Ghanipour-Meybodi, R., Petrášová, A., & Jenčová, J. (2024). From defense to offense: antimicrobial peptides as promising therapeutics for cancer. Frontiers in Oncology, 14, 1463088. https://doi.org/10.3389/fonc.2024.1463088

[48] Jenča, A., Mills, D.K., Ghasemi, H., Saberian, E., Jenča, A., Karimi Forood, A.M., Petrášová, A., Jenčová, J., Jabbari Velisdeh, Z., Zare-Zardini, H. and Ebrahimifar, M. (2024). Herbal therapies for cancer treatment: A review of phytotherapeutic efficacy. Biologics: Targets and Therapy, pp.229-255. https://doi.org/10.2147/BTT.S484068