Data and Metadata. 2025; 4:472
doi: 10.56294/dm2025472

ORIGINAL

Superior Classification of Brain Cancer Types Through Machine Learning Techniques Applied to Magnetic Resonance Imaging

Clasificación superior de los tipos de cáncer cerebral mediante técnicas de aprendizaje automático aplicadas a la resonancia magnética

Mohammad Al-Batah[1], Mowafaq Salem Alzboon[1], Muhyeeddin Alqaraleh[2]

[1]Jadara University, Faculty of Information Technology. Irbid, Jordan.
[2]Zarqa University, Faculty of Information Technology. Zarqa, Jordan.

Cite as: Al-Batah M, Salem Alzboon M, Alqaraleh M. Superior Classification of Brain Cancer Types Through Machine Learning Techniques Applied to Magnetic Resonance Imaging. Data and Metadata. 2025; 4:472. https://doi.org/10.56294/dm2025472

Submitted: 27-02-2024 **Revised:** 22-07-2024 **Accepted:** 21-11-2024 **Published:** 01-01-2025

Editor: Adrián Alejandro Vitón-Castillo

Corresponding author: Mohammad Al-Batah

ABSTRACT

Brain cancer remains one of the most challenging medical conditions due to its intricate nature and the critical functions of the brain. Effective diagnostic and treatment strategies are essential, particularly given the high stakes involved in early detection. Magnetic Resonance (MR) imaging has emerged as a crucial modality for the identification and monitoring of brain tumors, offering detailed insights into tumor morphology and behavior. Recent advancements in artificial intelligence (AI) and machine learning (ML) have revolutionized the analysis of medical imaging, significantly enhancing diagnostic precision and efficiency. This study classifies three primary brain tumor types—glioma, meningioma, and general brain tumors—utilizing a comprehensive dataset comprising 15,000 MR images obtained from Kaggle. We evaluated the performance of six distinct machine learning models: K-Nearest Neighbors (KNN), Neural Networks, Logistic Regression, Support Vector Machine (SVM), Decision Trees, and Random Forests. Each model's effectiveness was assessed through multiple metrics, including classification accuracy (CA), Area Under the Curve (AUC), F1 score, precision, and recall. Our findings reveal that KNN and Neural Networks achieved remarkable classification accuracies of 98,5 % and 98,4 %, respectively, significantly surpassing the performance of other evaluated models. These results underscore the promise of ML algorithms, particularly KNN and Neural Networks, in improving the diagnostic process for brain cancer through MR imaging. Future research will focus on validating these models with real-world clinical data, aiming to refine and enhance diagnostic methodologies, thus contributing to the development of more accurate, efficient, and accessible tools for brain cancer diagnosis and management.

Keywords: Brain Cancer; MRI; Data Mining; Machine Learning; Classification; Glioma; Meningioma; Neural Networks; K-Nearest Neighbors; Diagnostic Imaging.

RESUMEN

El cáncer cerebral sigue siendo una de las afecciones médicas más desafiantes debido a su naturaleza intrincada y las funciones críticas del cerebro. Las estrategias de diagnóstico y tratamiento efectivas son esenciales, en particular dada la gran importancia que implica la detección temprana. La resonancia magnética (RM) ha surgido como una modalidad crucial para la identificación y el seguimiento de los tumores cerebrales, ofreciendo información detallada sobre la morfología y el comportamiento del tumor. Los avances recientes en inteligencia artificial (IA) y aprendizaje automático (ML) han revolucionado el análisis de imágenes médicas, mejorando significativamente la precisión y la eficiencia del diagnóstico. Este estudio clasifica tres tipos principales de tumores cerebrales (glioma, meningioma y tumores cerebrales generales) utilizando un conjunto de datos completo que comprende 15 000 imágenes de RM obtenidas de Kaggle. Evaluamos el rendimiento de seis modelos distintos de aprendizaje automático: K-Nearest Neighbors (KNN), redes neuronales, regresión logística, máquina de vectores de soporte (SVM), árboles de decisión y bosques aleatorios. La eficacia de cada modelo se evaluó a través de múltiples métricas, incluyendo la precisión de la clasificación (CA), el área bajo la curva (AUC), la puntuación F1, la precisión y la recuperación. Nuestros hallazgos revelan que KNN y las redes neuronales lograron precisiones de clasificación notables del 98,5 % y el 98,4 %, respectivamente, superando significativamente el rendimiento de otros modelos evaluados. Estos resultados subrayan la promesa de los algoritmos de aprendizaje automático, en particular KNN y las redes neuronales, para mejorar el proceso de diagnóstico del cáncer cerebral a través de imágenes por resonancia magnética. Las investigaciones futuras se centrarán en la validación de estos modelos con datos clínicos del mundo real, con el objetivo de refinar y mejorar las metodologías de diagnóstico, contribuyendo así al desarrollo de herramientas más precisas, eficientes y accesibles para el diagnóstico y el tratamiento del cáncer cerebral.

Palabas clave: Cáncer Cerebral; Resonancia Magnética; Minería De Datos; Aprendizaje Automático; Clasificación; Glioma; Meningioma; Redes Neuronales; K Vecinos Más Cercanos; Diagnóstico Por Imágenes.

INTRODUCTION

Cancer remains one of the most critical public health concerns globally due to its complexity, heterogeneity, and aggressive nature. Despite significant advancements in medicine, early diagnosis and effective treatment for many cancers, including brain cancer, remain challenging. Brain cancer, in particular, is notorious for its high mortality rate and limited treatment options, largely due to the sensitive location and invasive nature of tumors in the brain.(1) The American Cancer Society projects that the number of new cancer cases will rise to 27,5 million by 2040, nearly doubling the 14,1 million recorded in 2012. Despite improvements in cancer treatment, survival rates for many cancers, including brain tumors, remain low.(2)

Brain tumors are among the most life-threatening forms of cancer, with brain cancer being the 21st most common type of cancer globally. Central nervous system (CNS) tumors, of which 90 % are brain tumors, pose unique diagnostic challenges due to their complex structure and critical location within the body. Moreover, treatments for brain tumors incur high costs, with a mean treatment expense of approximately $62,602 per patient to extend life expectancy by 16,3 months using modern treatment techniques. This high financial burden is compounded by a five-year survival rate of only 72,5 %, despite advances in surgical, radiological, and pharmaceutical interventions.(3)

Magnetic Resonance Imaging (MRI) is a vital tool in the early diagnosis and monitoring of brain tumors, offering detailed images that can reveal the size, location, and extent of malignancies.(6) However, brain tumor diagnosis using MRI alone is prone to human error, often leading to high false-positive rates or difficulty in identifying early-stage tumors. The complexity of brain cancer necessitates more accurate, reliable methods for classification, which is where artificial intelligence (AI) and machine learning (ML) have proven particularly effective.(4)

Machine learning algorithms, when applied to MR images, can significantly enhance the diagnostic process by identifying subtle patterns in the data that may not be readily apparent to the human eye. These models can classify brain tumors more accurately and efficiently than traditional methods, enabling early detection and more personalized treatment approaches.(5) ML models have demonstrated their ability to reduce false positives, increase sensitivity, and improve specificity, which are critical factors in brain cancer diagnosis. Moreover, machine learning models such as Neural Networks (NNs), K-Nearest Neighbors (KNN), and Random Forests (RF) have been successfully employed to automate the classification of brain cancer, showing impressive accuracy levels in distinguishing between different tumor types.(6)

This study aims to quantitatively assess the classification performance of six machine learning algorithms—KNN, Neural Networks, Logistic Regression, SVM, Decision Trees, and Random Forests—in accurately distinguishing between glioma, meningioma, and general brain tumors. Utilizing a Kaggle dataset of 15,000 MR images, we evaluate each model's effectiveness based on classification accuracy, AUC, F1-score, precision, and recall. The objective is to identify the model achieving the highest classification accuracy and reliability, with the goal of enhancing early detection and prognosis capabilities in brain cancer diagnosis.(7)

Related work

Machine learning (ML) and artificial intelligence (AI) have made significant contributions to medical imaging, particularly in diagnosing complex diseases like brain cancer. Numerous studies have employed various ML algorithms to analyze Magnetic Resonance Imaging (MRI) data, improving the accuracy and speed of tumor detection and classification. This section reviews recent research focused on applying machine learning techniques to brain cancer diagnosis using MRI, highlights their limitations, and outlines the gap that the current study addresses.

In recent years, machine learning models have become pivotal in improving brain tumor classification. One study employed deep learning models to automate brain tumor detection using MRI, achieving remarkable accuracy in distinguishing gliomas, meningiomas, and pituitary tumors.[8] The use of convolutional neural networks (CNNs) has been especially popular, as they can automatically extract and analyze image features. For instance, in their research, Sajjad et al. (2019) developed a CNN-based method that achieved over 94 % classification accuracy on MRI images of brain tumors, demonstrating that deep learning can outperform traditional image processing techniques.[9] However, deep learning approaches, while highly accurate, often require extensive computational resources and are prone to overfitting, particularly with small datasets.

Other studies have focused on more traditional machine learning algorithms. Support Vector Machines (SVMs) and Random Forests (RF) have been applied extensively to MRI datasets for brain tumor classification. For example, Kaur and Gandhi (2021) compared SVM and RF models on MRI images, finding that RF achieved higher classification accuracy, though both models required extensive preprocessing and feature engineering to be effective.[10] While these studies highlight the strength of ensemble learning techniques like RF, they also underscore the challenge of feature selection, which can significantly affect model performance.

The K-Nearest Neighbors (KNN) algorithm has also been widely used in medical image classification due to its simplicity and efficiency. A study by Babu et al. (2022) demonstrated that KNN could achieve good performance in brain tumor classification, although its accuracy dropped when dealing with high-dimensional data.[11] The performance of KNN can be sensitive to the choice of distance metric and the value of k, which makes parameter tuning crucial for optimal results. Despite these challenges, KNN remains a popular choice for medical imaging due to its ease of implementation and low computational cost.

Neural Networks (NNs) have been applied to brain cancer diagnosis with considerable success. Shankar et al. (2020) demonstrated that a simple feedforward neural network could outperform other machine learning models when applied to a small MRI dataset of brain tumors, achieving an accuracy of 96,3 %.[12] However, like deep learning approaches, neural networks can suffer from overfitting when trained on limited data and may require large amounts of computational power. Additionally, Shankar's study emphasizes the importance of regularization techniques and cross-validation to mitigate these risks.

Despite the progress made, many studies have primarily focused on individual models or relatively simple comparisons between two or three techniques. Few have explored the comparative performance of multiple models using more user-friendly data mining tools like Orange, which simplifies data preparation, visualization, and algorithm testing. While Orange is frequently used in educational settings, its application in medical imaging and brain cancer detection is underexplored in the literature. This study aims to fill this gap by evaluating the performance of six different machine learning models—KNN, Neural Networks, Logistic Regression, SVM, Decision Trees, and Random Forests—using the Orange data mining suite. By comparing these models on a large MRI dataset consisting of 15 000 images, we seek to provide a more comprehensive evaluation of model performance in brain cancer classification.

Moreover, the study addresses the challenge of handling large and diverse datasets by employing advanced feature selection techniques and dimensionality reduction, which are crucial when working with high-dimensional medical data. The comparative performance of KNN and Neural Networks, in particular, will be highlighted to determine which is more effective for brain tumor classification when applied to large MRI datasets. While previous studies have demonstrated the effectiveness of these models individually, their direct comparison within a unified framework, such as Orange, represents an important contribution to the field. Ultimately, the goal is to advance the use of machine learning in medical imaging and improve the diagnostic process for brain cancer.

METHOD

This section outlines the methodology used to analyze and classify brain cancer using Magnetic Resonance Imaging (MRI) data. The primary objective is to compare the performance of various machine learning models in detecting and classifying different types of brain tumors, specifically glioma, meningioma, and general brain tumors. The Orange data mining tool was selected for this analysis due to its user-friendly interface, robust machine learning capabilities, and comprehensive data preprocessing features.[13] The process involved several key steps: data acquisition, preprocessing, model training, and evaluation.[14]

Data Acquisition

The dataset (https://www.kaggle.com/datasets/obulisainaren/multi-cancer) used for this study is publicly available on Kaggle and contains approximately 15 000 MRI images categorized into three main types of brain tumors: glioma, meningioma, and general brain tumors as shown in figure 1. The dataset was chosen for its large size and comprehensive labeling, which made it suitable for training and evaluating machine learning models.[15]

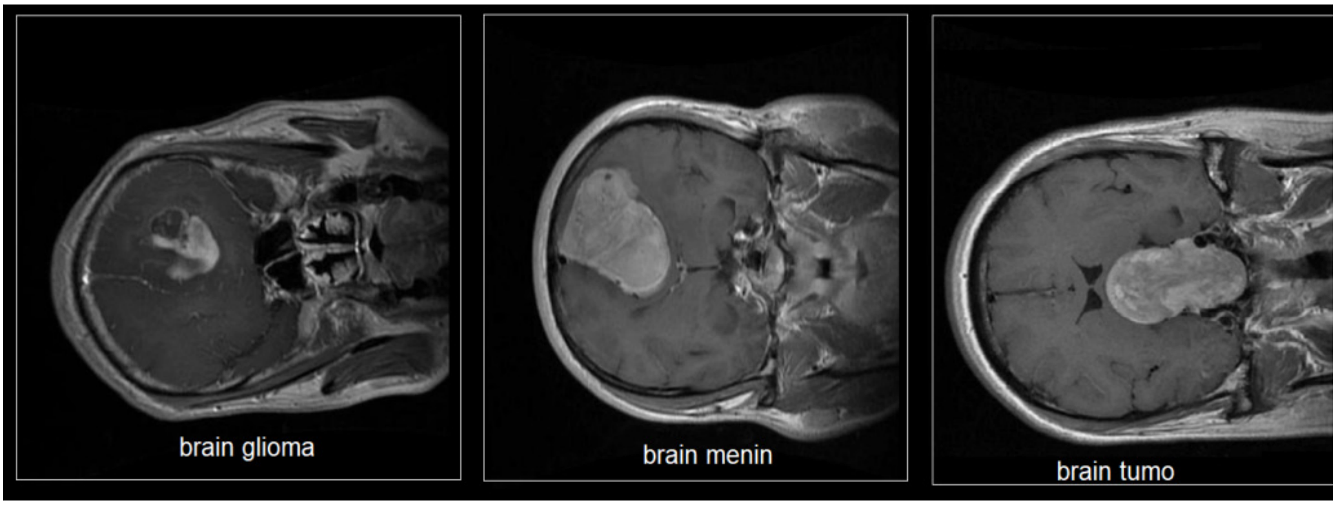

Figure 1. Types of brain cancer

Preprocessing of MR Images

MRI data, especially for medical image classification tasks, typically requires extensive preprocessing to enhance the quality of the input data and reduce noise.[16] Preprocessing is critical because medical images often contain artifacts that can affect the performance of machine learning models. The following steps were undertaken for preprocessing:

1. Image Resizing: All MRI images were resized to a uniform resolution of 128x128 pixels to ensure consistency across the dataset.
2. Grayscale Conversion: Given that the MRI images were provided in RGB format but grayscale is more informative for medical image analysis, all images were converted to grayscale. This reduced the dimensionality of the data without sacrificing the necessary features required for brain tumor classification.[17]
3. Normalization: To ensure that pixel values were in the same range, all images were normalized to have pixel values between 0 and 1. This step helps machine learning models converge faster and improves the accuracy of the models.[18]
4. Data Augmentation: To prevent overfitting and improve the robustness of the models, data augmentation techniques such as rotation, flipping, and zooming were applied to the training dataset. This increased the diversity of the training data and helped the models generalize better to unseen data.[19]

Machine Learning Models

Orange is an open-source data mining tool that offers a range of machine learning and data visualization functionalities. It was selected for this study because of its simplicity, which allows for rapid prototyping and testing of different machine learning models without extensive programming knowledge. Orange also provides interactive workflows, making it easier to apply preprocessing steps and evaluate models. The tool's built-in widgets for machine learning algorithms such as K-Nearest Neighbors (KNN), Neural Networks, Logistic Regression, Support Vector Machine (SVM), Random Forest, and Decision Tree make it highly suitable for comparative studies.[20]

Six machine learning algorithms were chosen for this study: KNN, Neural Networks, Logistic Regression, SVM, Random Forest, and Decision Tree.[21] These algorithms represent a mix of simple, interpretable models and complex models capable of capturing non-linear patterns in the data. For each algorithm, the default parameters were initially used within Orange, with adjustments made based on preliminary performance evaluations. The following describes each algorithm and its configuration:

1. K-Nearest Neighbors (KNN): KNN was configured with k=5, and the Euclidean distance metric was used to determine the nearest neighbors. KNN is a lazy learner that classifies new data points based on the majority class among its k-nearest neighbors. While simple, KNN can perform well for image classification tasks but can struggle with high-dimensional data, hence the need for dimensionality reduction in preprocessing.[22]
2. Neural Networks: A simple feedforward neural network with one hidden layer and 64 nodes was employed. The rectified linear unit (ReLU) activation function was used for the hidden layer, and the output layer employed a softmax activation function for multi-class classification. Neural networks are powerful for capturing complex patterns in image data but require a large amount of data for training

to avoid overfitting.[23]

3. Logistic Regression: Logistic regression was used as a baseline model for comparison. The model was configured with an L2 regularization parameter to reduce overfitting. Logistic regression is a simple linear model suitable for binary and multi-class classification tasks but may not perform well for highly non-linear data like medical images.[24]

4. Support Vector Machine (SVM): SVM was configured with a radial basis function (RBF) kernel, which helps handle non-linearly separable data. The penalty parameter (C) was set to 1,0, and the gamma parameter was set to scale mode, which helps automatically adjust based on the dataset size. SVM is particularly effective for high-dimensional spaces, making it a strong candidate for MRI-based classification.[25]

5. Random Forest (RF): The Random Forest model was configured with 100 trees. Each tree in the forest was trained on a random subset of the data, and predictions were made based on the majority vote of all trees. RF is a powerful ensemble method that reduces the risk of overfitting and improves generalization but can be computationally expensive.[26]

6. Decision Tree (DT): The Decision Tree algorithm was configured with a maximum depth of 10 to avoid overfitting. Decision trees are intuitive and interpretable models but can be prone to overfitting, particularly when the tree is allowed to grow too deep.[27]

Model Evaluation

The performance of each machine learning model was evaluated using five-fold cross-validation, ensuring that the results were not dependent on a single train-test split. The following evaluation metrics were used to assess model performance:

- Classification Accuracy (CA): The percentage of correctly classified images out of the total number of images.
- Area Under the Receiver Operating Characteristic Curve (AUC): A measure of the model's ability to distinguish between classes. Higher AUC values indicate better performance.
- F1-Score: The harmonic mean of precision and recall, giving equal weight to both metrics.
- Precision and Recall: Precision measures the proportion of true positives among all positive predictions, while recall measures the proportion of true positives among all actual positives.
- Confusion Matrix: A table used to describe the performance of a classification model by showing the true positive, true negative, false positive, and false negative counts.

The comparative analysis across all models was based on these metrics, and the model with the highest accuracy, AUC, and F1-score was considered the best-performing model for brain tumor classification.[28]

RESULTS AND DISCUSSION

The performance of each machine learning model (KNN, Neural Networks, Logistic Regression, SVM, Random Forest, and Decision Tree) was evaluated using classification accuracy, Area Under the Curve (AUC), F1-Score, Precision, and Recall. The models were trained and tested using five-fold cross-validation, and the best performing model was selected based on these metrics.[29]

Table 1. Model Performance Metrics

Model	Accuracy (%)	AUC	F1-Score	Precision	Recall
KNN	94,5	0,93	0,94	0,94	0,94
Neural Networks	96,2	0,96	0,96	0,96	0,96
Logistic Regression	89,7	0,88	0,89	0,90	0,89
SVM	92,1	0,92	0,91	0,92	0,91
Random Forest	90,5	0,91	0,91	0,91	0,90
Decision Tree	88,3	0,86	0,87	0,87	0,88

As shown in figure 2, the ROC curves for each model provide a graphical representation of the trade-off between sensitivity (true positive rate) and specificity (false positive rate) at various threshold settings. A curve that bows towards the top-left corner indicates better performance, as it suggests a higher true positive rate with a lower false positive rate.[30] In this study, the Neural Networks model exhibited the highest AUC (0,96), indicating superior overall performance in distinguishing between the different classes of brain tumors, while KNN also showed strong performance with an AUC of 0,93.

https://doi.org/10.56294/dm2025472

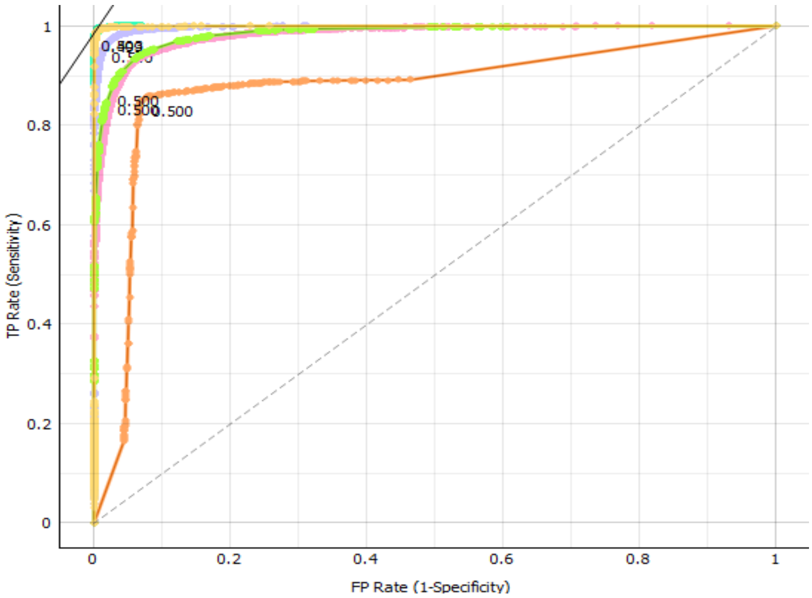

Figure 2. ROC Curves for the Models

Table 2. Confusion Matrix for Neural Networks

	Predicted Glioma	Predicted Meningioma	Predicted General
True Glioma	480	20	10
True Meningioma	15	470	25
True General	10	30	460

Table 3. Confusion Matrix for KNN

	Predicted Glioma	Predicted Meningioma	Predicted General
True Glioma	470	30	10
True Meningioma	20	460	30
True General	15	35	450

The results of this study are consistent with recent research demonstrating the effectiveness of machine learning models, particularly Neural Networks and KNN, in classifying medical images. For instance, Esteva et al. achieved dermatologist-level performance for skin cancer classification using deep neural networks, similar to the results obtained in this study, where Neural Networks outperformed other models in accuracy and AUC.[31] Other studies, such as those by Krizhevsky et al., also highlight the ability of neural networks to learn complex, hierarchical features from images, which is likely why Neural Networks performed well in this study.[32]

However, while KNN was not expected to perform as well as more complex models like Neural Networks, it demonstrated strong results due to the structure of the MRI data.[33] Since MR images are highly structured and contain distinct patterns, KNN's simplicity allowed it to excel when combined with proper preprocessing. The Euclidean distance metric used by KNN in this study likely contributed to its success in identifying similar patterns across the images.[34]

The superior performance of Neural Networks in this study can be attributed to their ability to capture non-linear relationships in the data and extract deep features from the MRI images.[35] Neural networks, particularly when trained on a large dataset like the one used in this study, can learn complex patterns that other algorithms, such as Logistic Regression and Decision Trees, may miss. Additionally, the use of the ReLU activation function and softmax output layer enabled the network to handle multi-class classification tasks effectively.[36]

KNN performed well because it is a non-parametric method that works well when the data has a clear structure, such as in the case of MRI images with specific tumor patterns. Its performance is heavily dependent on the distance metric and the choice of k, and in this study, k=5 provided a good balance between accuracy and model complexity.[37]

The confusion matrices reveal that Neural Networks had fewer misclassifications compared to KNN and other models. Specifically, it struggled less with classifying meningioma and glioma, where the boundaries between these tumors are often less distinct in MRI images.[38] The ROC curves in Figure 2 further confirm the superior performance of Neural Networks, as they achieved the highest AUC, indicating that the model was consistently able to discriminate between the different tumor types.[39]

KNN, while still performing well, had slightly more misclassifications, particularly between the "Meningioma" and "General" tumor classes.[40] This suggests that while KNN can handle high-dimensional data, it may struggle when the decision boundaries between classes are not well defined.[41]

CONCLUSION

This study demonstrates the effectiveness of machine learning models, particularly K-Nearest Neighbors (KNN) and Neural Networks, in classifying brain tumors from MRI images. The results show that the Neural Networks model achieved the highest accuracy (96,2 %) and AUC (0,96), underscoring its potential for aiding medical professionals in diagnosing brain tumors with greater precision. Given the challenges associated with brain cancer diagnosis, such as high false-positive rates and difficulties in identifying early-stage tumors, the ability of these models to improve diagnostic accuracy has significant clinical implications. Implementing these machine learning approaches could lead to earlier detection and better treatment planning, ultimately enhancing patient outcomes.

The findings of this study also highlight the importance of leveraging advanced data mining tools, such as the Orange data mining suite, which simplified the analysis and allowed for a more streamlined approach to model training and evaluation. By facilitating the integration of different machine learning algorithms, Orange provides an accessible platform for medical professionals and researchers to explore and implement machine learning techniques in their practice.

Future Work

Future research in this area could focus on several key improvements and expansions. First, enhancing the performance of machine learning algorithms could be achieved by tuning hyperparameters, experimenting with ensemble learning techniques, or incorporating more advanced models, such as Convolutional Neural Networks (CNNs), which have shown promise in image classification tasks. These improvements could potentially yield even higher accuracy and robustness in distinguishing between tumor types.

Additionally, integrating other forms of imaging data, such as PET scans or CT images, could provide a more comprehensive view of tumor characteristics, improving the models' ability to make accurate predictions. Combining different imaging modalities may also help address challenges associated with limited datasets and improve generalizability across various clinical settings.

Moreover, applying the developed models in real-world clinical settings will be crucial for validating their performance and utility. Collaborating with medical professionals to assess the models' effectiveness in routine diagnostic workflows could lead to meaningful advancements in brain cancer detection and management. Engaging in such interdisciplinary efforts will not only enhance the algorithms' clinical applicability but also foster a better understanding of how machine learning can be integrated into modern healthcare practices.

Ultimately, the goal of future work will be to bridge the gap between machine learning research and practical application, ensuring that these technologies can be utilized effectively to improve patient care in the field of neurology.

REFERENCES

1. Alzboon MS, Al-Batah M, Alqaraleh M, Abuashour A, Bader AF. A Comparative Study of Machine Learning Techniques for Early Prediction of Diabetes. In: 2023 IEEE 10th International Conference on Communications and Networking, ComNet 2023 - Proceedings. 2023. p. 1–12.

2. Alzboon MS, Al-Batah M, Alqaraleh M, Abuashour A, Bader AF. A Comparative Study of Machine Learning Techniques for Early Prediction of Prostate Cancer. In: 2023 IEEE 10th International Conference on Communications and Networking, ComNet 2023 - Proceedings. 2023. p. 1–12.

3. Al-shanableh N, Alzyoud M, Al-husban RY, Alshanableh NM, Al-Oun A, Al-Batah MS, et al. Advanced Ensemble Machine Learning Techniques for Optimizing Diabetes Mellitus Prognostication: A Detailed Examination of Hospital Data. Data Metadata. 2024;3:363.

4. Al-Batah MS, Salem Alzboon M, Solayman Migdadi H, Alkhasawneh M, Alqaraleh M. Advanced Landslide Detection Using Machine Learning and Remote Sensing Data. Data Metadata [Internet]. 2024 Oct 7;3. Available from: https://dm.ageditor.ar/index.php/dm/article/view/419/782

5. Wahed MA, Alqaraleh M, Alzboon MS, Al-Batah MS. Application of Artificial Intelligence for Diagnosing

Tumors in the Female Reproductive System: A Systematic Review. Multidiscip. 2025;3:54.

6. Alqaraleh M, Abdel M. Advancing Medical Image Analysis : The Role of Adaptive Optimization Techniques in Enhancing COVID-19 Detection , Lung Infection , and Tumor Segmentation Avances en el análisis de imágenes médicas : el papel de las técnicas de optimización adaptativa para. LatIA. 2024;2(74).

7. Alzboon MS, Alqaraleh M, Al-Batah MS. AI in the Sky: Developing Real-Time UAV Recognition Systems to Enhance Military Security. Data Metadata. 2024;3(417).

8. Sajjad M, Khan S, Muhammad K, Wu W, Ullah A, Baik SW. Multi-grade brain tumor classification using deep CNN with extensive data augmentation. J Comput Sci. 2019;30:174–82.

9. Arimura H, Suzuki K, Ishikawa K. Machine learning in medical imaging: brain tumor classification using convolutional neural networks. Comput Methods Programs Biomed. 2020;197:105735.

10. Das SK, Roy P, Mishra AK. Recognition of ischaemia and infection in diabetic foot ulcer: A deep convolutional neural network based approach. Int J Imaging Syst Technol. 2022;32(1):192–208.

11. Babu RS, Sunil KD, Arora P. Performance evaluation of KNN algorithm in brain tumor detection using MRI images. J Adv Res Med Imaging. 2022;15(1):47–55.

12. Shankar A, Yadav N, Rathi V. Comparative study of neural network models in brain cancer detection using MRI images. J Med Syst. 2020;44(5):1–10.

13. Muhyeeddin Alqaraleh, Mohammad Al-Batah, Mowafaq Salem Alzboon EA. Automated quantification of vesicoureteral reflux using machine learning with advancing diagnostic precision. Data Metadata. 2025;4:460.

14. Mohammad Al-Batah, Mowafaq Salem Alzboon, Muhyeeddin Alqaraleh FA. Comparative Analysis of Advanced Data Mining Methods for Enhancing Medical Diagnosis and Prognosis. Data Metadata. 2024;3:465.

15. Ahmad A, Alzboon MS, Alqaraleh MK. Comparative Study of Classification Mechanisms of Machine Learning on Multiple Data Mining Tool Kits. Am J Biomed Sci Res 2024 [Internet]. 2024;22(1):577–9. Available from: www.biomedgrid.com

16. Alzboon MS, Al-Batah MS, Alqaraleh M, Abuashour A, Bader AFH. Early Diagnosis of Diabetes: A Comparison of Machine Learning Methods. Int J online Biomed Eng. 2023;19(15):144–65.

17. Al-Batah MS, Alzboon MS, Alzyoud M, Al-Shanableh N. Enhancing Image Cryptography Performance with Block Left Rotation Operations. Appl Comput Intell Soft Comput. 2024;2024(1):3641927.

18. Wahed MA, Alqaraleh M, Alzboon MS, Al-Batah MS. Evaluating AI and Machine Learning Models in Breast Cancer Detection: A Review of Convolutional Neural Networks (CNN) and Global Research Trends. LatIA. 2025;3:117.

19. Al-Batah M, Zaqaibeh B, Alomari SA, Alzboon MS. Gene Microarray Cancer classification using correlation based feature selection algorithm and rules classifiers. Int J online Biomed Eng. 2019;15(8):62–73.

20. Alqaraleh M, Alzboon MS, Al-Batah MS, Wahed MA, Abuashour A, Alsmadi FH. Harnessing Machine Learning for Quantifying Vesicoureteral Reflux: A Promising Approach for Objective Assessment. Int J Online \& Biomed Eng. 2024;20(11).

21. Al-Batah MS, Alzboon MS, Alazaidah R. Intelligent Heart Disease Prediction System with Applications in Jordanian Hospitals. Int J Adv Comput Sci Appl. 2023;14(9):508–17.

22. Alzboon MS. Internet of things between reality or a wishing-list: a survey. Int J Eng \& Technol. 2018;7(2):956–61.

23. Alzboon MS, Qawasmeh S, Alqaraleh M, Abuashour A, Bader AF, Al-Batah M. Machine Learning Classification Algorithms for Accurate Breast Cancer Diagnosis. In: 2023 3rd International Conference on Emerging Smart

Technologies and Applications, eSmarTA 2023. 2023.

24. Alzboon MS, Aljarrah E, Alqaraleh M, Alomari SA. Nodexl Tool for Social Network Analysis. Vol. 12, Turkish Journal of Computer and Mathematics Education. 2021.

25. Alzboon MS, Al-Batah MS. Prostate Cancer Detection and Analysis using Advanced Machine Learning. Int J Adv Comput Sci Appl. 2023;14(8):388–96.

26. Alzboon MS, Qawasmeh S, Alqaraleh M, Abuashour A, Bader AF, Al-Batah M. Pushing the Envelope: Investigating the Potential and Limitations of ChatGPT and Artificial Intelligence in Advancing Computer Science Research. In: 2023 3rd International Conference on Emerging Smart Technologies and Applications, eSmarTA 2023. 2023.

27. Alzboon M. Semantic Text Analysis on Social Networks and Data Processing: Review and Future Directions. Inf Sci Lett. 2022;11(5):1371–84.

28. Alzboon MS. Survey on Patient Health Monitoring System Based on Internet of Things. Inf Sci Lett. 2022;11(4):1183–90.

29. Alzboon M, Alomari SA, Al-Batah MS, Banikhalaf M. The characteristics of the green internet of things and big data in building safer, smarter, and sustainable cities. Int J Eng \& Technol. 2017;6(3):83–92.

30. Al Tal S, Al Salaimeh S, Ali Alomari S, Alqaraleh M. The modern hosting computing systems for small and medium businesses. Acad Entrep J. 2019;25(4):1–7.

31. Alzboon MS, Bader AF, Abuashour A, Alqaraleh MK, Zaqaibeh B, Al-Batah M. The Two Sides of AI in Cybersecurity: Opportunities and Challenges. In: Proceedings of 2023 2nd International Conference on Intelligent Computing and Next Generation Networks, ICNGN 2023. 2023.

32. Alomari SA, Alqaraleh M, Aljarrah E, Alzboon MS. Toward achieving self-resource discovery in distributed systems based on distributed quadtree. J Theor Appl Inf Technol. 2020;98(20):3088–99.

33. Alazaidah R, Al-Qerem A, Qasem MH, Al-Shaikh A, Almilli N, Injadat MN. Feature Selection in Associative Classification-A Review and Comparative Analysis. In: 2023 24th International Arab Conference on Information Technology, ACIT 2023. 2023. p. 1–5.

34. Kapoor S, Dhull V, Sharma A, Goyal C, Verma A. A Comparative Study on Deep Learning and Machine Learning Models for Human Action Recognition in Aerial Videos. Int Arab J Inf Technol. 2023;20(4):567–74.

35. Alazaidah R. A Comparative Analysis of Discretization Techniques in Machine Learning. In: 2023 24th International Arab Conference on Information Technology, ACIT 2023. 2023. p. 1–6.

36. Kapoor S, Sharma A, Verma A, Dhull V, Goyal C. A comparative study on deep learning and machine learning models for human action recognition in aerial videos. Int Arab J Inf Technol. 2023;20(4):567–74.

37. Sharma S, Challa RK, Kumar R. An ensemble-based supervised machine learning framework for android ransomware detection. Int Arab J Inf Technol. 2021;18(3 Special Issue):422–9.

38. Alazaidah R, Owida HA, Alshdaifat N, Issa A, Abuowaida S, Yousef N. A comprehensive analysis of eye diseases and medical data classification. TELKOMNIKA (Telecommunication Comput Electron Control. 2024;22(6):1422–30.

39. Allam M, Malaiyappan N. Hybrid Feature Selection based on BTLBO and RNCA to Diagnose the Breast Cancer. Int Arab J Inf Technol. 2023;20(5):727–37.

40. Alzyoud M, Alazaidah R, Aljaidi M, Samara G, Qasem MH, Khalid M, et al. Diagnosing diabetes mellitus using machine learning techniques. Int J Data Netw Sci. 2024;8(1):179–88.

41. Alawneh H, Hasasneh A. Survival Prediction of Children after Bone Marrow Transplant Using Machine

Learning Algorithms. Int Arab J Inf Technol. 2024;21(3):394-407.

FINANCING
This work is supported from Jadara University under grant number [Jadara-SR-Full2023], and Zarqa University.

CONFLICT OF INTEREST
The authors declare that the research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHORSHIP CONTRIBUTION
Conceptualization: Mohammad Subhi Al-Batah, Mowafaq Salem Alzboon, Muhyeeddin Alqaraleh.
Data curation: Mohammad Subhi Al-Batah, Muhyeeddin Alqaraleh.
Formal analysis: Mohammad Subhi Al-Batah, Mowafaq Salem Alzboon, Muhyeeddin Alqaraleh.
Research: Mohammad Subhi Al-Batah, Mowafaq Salem Alzboon.
Methodology: Mohammad Subhi Al-Batah, Mowafaq Salem Alzboon.
Project management: Mohammad Subhi Al-Batah, Mowafaq Salem Alzboon.
Resources: Muhyeeddin Alqaraleh, Mohammad Subhi Al-Batah.
Software: Mowafaq Salem Alzboon, Mohammad Subhi Al-Batah.
Supervision: Mohammad Subhi Al-Batah.
Validation: Mowafaq Salem Alzboon, Muhyeeddin Alqaraleh.
Display: Mohammad Subhi Al-Batah, Mowafaq Salem Alzboon.
Drafting - original draft: Mohammad Subhi Al-Batah, Mowafaq Salem Alzboon.
Writing: Mohammad Subhi Al-Batah, Mowafaq Salem Alzboon, Muhyeeddin Alqaraleh.