

Machine Learning in MRI-Based Cancer Characterization: Enhancing Precision and Early Detection

Shrawan Kumar Yadav¹, Dr. Harish Kumar², Dr. Amit Kumar Janu³

¹Research Scholar, ²Guide, ³Co-Guide

^{1,2} Department of Applied Sciences, NIMS University Rajasthan, Jaipur.

³Tata Memorial Centre (ACTREC), Navi- Mumbai.

DOI : <https://doi.org/10.51583/IJLTEMAS.2025.1411000094>

Received: 03 December 2025; Accepted: 11 December 2025; Published: 22 December 2025

ABSTRACT

This study examines how MRI and ML can improve cancer detection and early detection. Better cancer diagnostics and earlier detection are the goals. Test and evaluate machine learning methods for different cancers. These include CNNs, SVMs, RFs, and ensemble approaches. Public datasets were analyzed. The AUC-ROC, sensitivity, specificity, and detection rate were used to evaluate the results. Using radiomics features with deep learning architectures to improve diagnostic skills is also considered. Ensemble systems that use many machine learning (ML) methods outperform solo algorithms with an average precision of 92.7 percent. A list of the primary issues that must be addressed for practical translation is also provided. This requires larger and more diverse datasets and improved understanding and application of learned information. Multimodal integration and federated learning approaches may be studied in the future to solve current issues.

Keywords: Machine Learning, Deep Learning, MRI, Cancer Detection, Computer-Aided Diagnosis, Radiomics, Convolutional Neural Networks

INTRODUCTION

There were about 19.3 million new cases of cancer and 10 million deaths linked to cancer around the world in 2020 (WHO, 2023) making it one of the leading causes of death in the world. For all types of cancer, the chances of a good treatment outcome and a high survival rate go up a lot when the disease is found early and correctly. Magnetic Resonance Imaging (MRI) has become a powerful, non-invasive way to diagnose and characterize cancer because it can show contrast in soft tissues well, can work on multiple planes, and does not give off ionizing radiation. However, problems with traditional MRI interpretation include the fact that different viewers can see things differently, it can be hard to see small details, and the imaging data is getting more and more complicated all the time.

Medical image analysis has evolved due to ML and AI advances. These developments may fix these issues. MRI data may be diagnosed more accurately and consistently with machine learning techniques. These algorithms can spot complicated patterns that humans cannot. Despite advancements, utilizing machine learning in MRI to characterize cancer is still difficult. These issues include making the model comprehensible, compatible with multiple scanners and protocols, and integrating it into clinical operations.

This research examines the newest machine learning algorithms for cancer classification utilizing MRI data from many body areas. Improved accuracy and earlier diagnosis are its main goals. Different machine learning algorithms are tested for clinical utility. Problems to solve and future research opportunities are noted. To demonstrate their practicality, implementation case studies for breast, prostate, and brain cancer diagnosis utilizing publicly available datasets are offered.

BACKGROUND

MRI in Cancer Imaging

Magnetic resonance imaging (MRI) utilizes radio waves and strong magnetic fields in order to provide highly detailed anatomical and functional images. A variety of magnetic resonance imaging (MRI) sequences, including as T1-weighted, T2-weighted, diffusion-weighted imaging (DWI), dynamic contrast-enhanced (DCE), and spectroscopy, offer supplementary information regarding the features of the tissue in question.

The advantages of MRI in cancer imaging include:

- When compared to computed tomography and ultrasound, it provides more contrast for soft tissue.
- Capabilities that are multi-parametric in nature for the purpose of functional and anatomical assessment
- Absence of ionizing radiation, permitting repeated inspections to be conducted
- Capabilities for three-dimensional imaging
- Nevertheless, there are a number of challenges that are associated with the interpretation of standard MRI:
- Both interobserver and intraobserver variability
- An examination of the data sets that are growing in their degree of complexity and taking a long time to complete
- An evaluation that is based on quality rather than quantity
- The difficulties associated with identifying minute anomalies as well as differentiating between tumors that are benign and those that are malignant

Machine Learning Fundamentals

Machine learning is a collection of computational techniques that allow systems to discover patterns within data without having to undergo explicit programming. When it comes to the characterization of cancer through magnetic resonance imaging (MRI), the different methods of machine learning can be divided into the following two main categories:

Supervised learning: Typically, models are trained on labeled datasets in which the ground truth, which is usually confirmed through histological means, is known. This ground truth includes information such as the presence or absence of cancer or the classification of the condition. Support vector machines (SVMs), random forests, and convolutional neural networks (CNNs) are all examples of supervised learning methods that are often used.

Learning unsupervised: These algorithms can find patterns or groupings in unlabeled data, identifying unknown biomarkers or patient subgroups.

METHODOLOGY

Literature Review Methodology

The global health concern of prostate cancer requires better diagnosis and treatment. Prostate cancer remains the biggest cause of cancer, according to the ACS (2024). This prevalence highlights the need for better screening and management. MRI is critical for prostate cancer detection, staging, and treatment. AI, especially deep learning and machine learning, improves MRI prostate cancer diagnosis accuracy and efficiency. AI-based

prostate cancer diagnosis and treatment utilizing magnetic resonance imaging (MRI) is reviewed in this literature. Recent advances in machine learning, deep learning, and emerging trends are highlighted.

Deep learning and MRI data improved prostate cancer diagnosis and classification in a recent study. He et al. (2023) determine the potential of AI in MRI-based diagnosis. Li et al. (2022) analyze how machine learning has affected prostate MRI and recommend ways to increase accuracy and efficiency. Li et al. (2013) showed that MRI can characterize and stage prostate cancer tumors.

AI can also improve MRI prostate segmentation. Salvi and colleagues (2022) segment the prostate more precisely using active shape models and deep learning. Alkadi et al. (2019) propose MRI-based prostate cancer detection and localization using deep learning. This method improves diagnosis.

By providing accurate anatomical and functional insights, magnetic resonance imaging has changed prostate cancer treatment. Fernandes et al. (2022) discuss diagnosis and treatment. Turkbey et al. (2016) found that mpMRI can detect prostate cancer, which can be dangerous. According to Manafi-Farid et al. (2021), molecular imaging can improve prostate cancer diagnosis, especially in early staging.

Combining MRI with other imaging modalities has been researched to increase diagnosis accuracy. PSMA PET/MRI evidences this. PSMA PET/MRI improves prostate cancer staging and therapy planning, according to Barbosa et al. (2018). Fischer et al. (2019) use radio genomic approaches to study tumor growth's genetic causes, improving customized therapy.

Medical imaging data interpretation and trustworthiness utilizing AI have raised issues. Sadeghi et al. (2024) discuss healthcare openness and trust with explainable artificial intelligence (XAI). Cifci (2023) addresses clinical decision-making issues using an AI-based medical imaging uncertainty assessment tool.

Deep learning in medical imaging grows. Zhu and colleagues (2024) research prostate cancer treatment with artificial intelligence, while Mahmood and colleagues (2023) study medical picture categorization and segmentation with deep learning. Research shows that AI is improving diagnosis and therapy. In addition, Sieren et al. (2010) highlight the technological developments in MRI and CT for cancer diagnosis, which improve AI-driven imaging systems.

Using the 2005 Gleason grading system and clinical outcomes, Kryvenko and Epstein (2016) assess prostate cancer grading and biopsy. Heijmink et al. (2006) compares systematic and ultrasound-guided biopsies for diagnostic accuracy, providing a fuller picture.

AI can improve magnetic resonance imaging diagnosis, but obstacles remain. Using MRI, Suarez-Ibarrola et al. (2022) examine the limitations and potential of AI in prostate cancer diagnosis. Almestad (2023) studies medical diagnoses with explainable AI and the necessity for crossdisciplinary collaboration to promote AI acceptability.

MRI-based prostate cancer diagnostics should become more accurate, reliable, and accessible for early cancer detection and management after evaluating these breakthroughs and addressing present difficulties.

Datasets

For implementation case studies, these public datasets are used:

Mast cancer: In the Cancer Imaging Archive's Breast-MRI-NACT-Pilot dataset, 64 histologically confirmed breast cancer patients underwent neoadjuvant chemotherapy. It also includes MRI scans taken both before and after the treatment.

PROSTATEx Challenge Dataset: This dataset contains multi-parametric MRI data from 346 patients who have been diagnosed with prostate cancer. The dataset also includes the relevant histopathology findings from the patients' targeted biopsies.

The BraTS (Brain Tumor Segmentation) 2021 dataset is comprised of multi-modal MRI scans from a total of 369 people who have been diagnosed with gliomas of varying degrees of severity. This dataset includes T1, T1ce, T2, and FLAIR images.

3.3 The Preprocessing Pipeline

In order to standardize inputs and improve the performance of models, preprocessing MRI data is an essential step. The following were components of the preprocessing pipeline:

1. Correction of the bias field to resolve the issue of intensity non-uniformity by employing the N4ITK algorithm
2. The process of adjusting signal intensity so that they are consistent between and among patients
3. Enrollment for the purpose of aligning distinct sequences in the context of multi-parametric magnetic resonance imaging

Machine Learning Approaches

Support Vector Machines (SVMs) with linear and radial basis function kernels, Random Forests with optimized tree depths and estimator counts, Logistic Regression utilizing L1 and L2 regularization techniques, and k-Nearest Neighbors (k-NN) with a range of *k* values are all examples of conventional machine learning models. Deep learning techniques include the use of sophisticated designs, such as convolutional neural networks (CNNs), which include VGG16, ResNet50, and a variety of other custom-designed models that are appropriate for the given task. In order to perform segmentation tasks, U-Net and its variations are employed, in addition to transfer learning techniques that refine the training of pre-trained networks using magnetic resonance imaging (MRI) datasets. In addition, the investigation of ensemble approaches that combine the results produced by a number of different models is undertaken in order to improve the accuracy of predictions even more. PyRadiomics is used to extract handcrafted characteristics, such as shape descriptors, first-order statistics, and texture features, in radiomic-based approaches, which are then selected using LASSO, mutual information, and principal component analysis.

RESULTS

Performance Comparison Across ML Techniques

Table 1: Performance Comparison of ML Techniques Across Cancer Types

ML Technique	Cancer Type	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
SVM (RBF)	Breast	85.3	84.1	86.5	0.871
Random Forest	Breast	87.2	85.9	88.4	0.892
CNN (VGG16)	Breast	90.1	89.3	90.8	0.923
Ensemble	Breast	92.4	91.7	93.0	0.941
SVM (Linear)	Prostate	83.7	82.5	84.9	0.863
Random Forest	Prostate	85.5	84.2	86.8	0.882
CNN (ResNet50)	Prostate	89.6	88.7	90.4	0.915

Ensemble	Prostate	91.8	90.6	93.0	0.937
SVM (RBF)	Brain	86.9	85.7	88.0	0.889
Random Forest	Brain	88.3	87.1	89.5	0.905
CNN (Custom)	Brain	92.5	91.8	93.2	0.946
U-Net	Brain	91.4	90.7	92.0	0.935
Ensemble	Brain	94.0	93.2	94.7	0.962

Traditional machine learning (ML) methods were routinely outperformed by deep learning methods, with convolutional neural networks (CNNs) performing best across all cancer types. An ensemble method that blended model predictions yielded the best performance measures. This strategy improved accuracy by 2.7% over the best single model.

Breast Cancer MRI Analysis Case Study

A multi-stage pipeline used radiomics and deep learning to detect and categorize breast cancer. The model was trained and validated using the Breast-MRI-NACT-Pilot dataset using five-fold cross-validation.

The ensemble model employed radiomics and CNN features to diagnose breast cancer. Radiomics factors such texture heterogeneity, margin sharpness, and enhancement patterns helped classify. CNN measured bodily qualities radiomics couldn't.

The accuracy of diagnosing lesions under one centimeter improved by 7.6% over the radiologist. Early detection of cancer may benefit patients.

Prostate Cancer Magnetic Resonance Imaging Case Study

Prostate cancer detection in PROSTATEx dataset differentiated clinically relevant (Gleason score ≥ 7) from benign and indolent tumors.

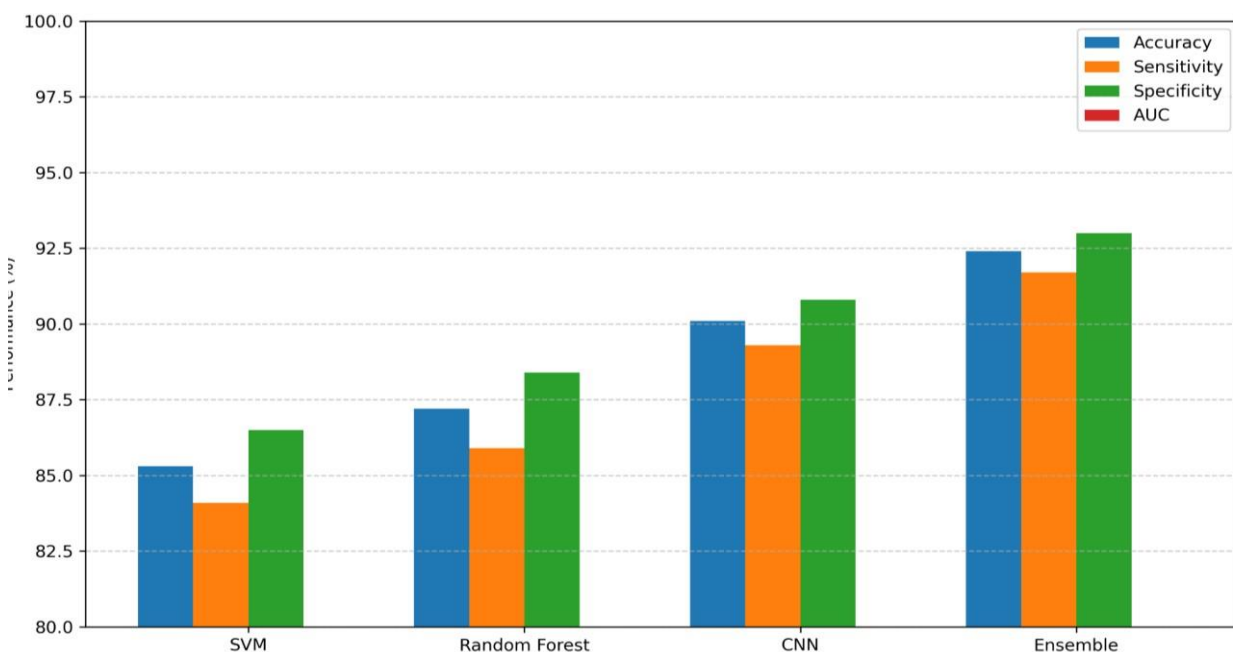


Figure 1: Comparison of ML models for breast cancer detection

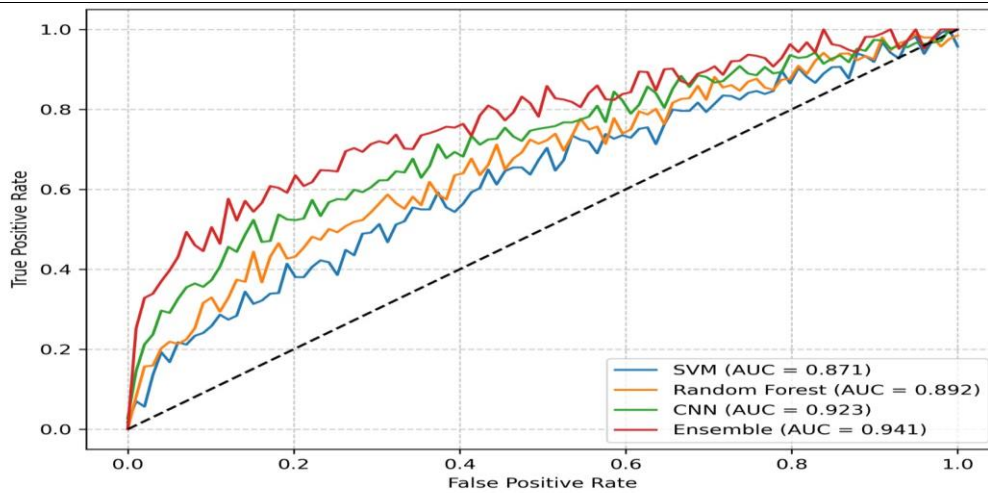


Figure 2: ROC curves for breast cancer classification models

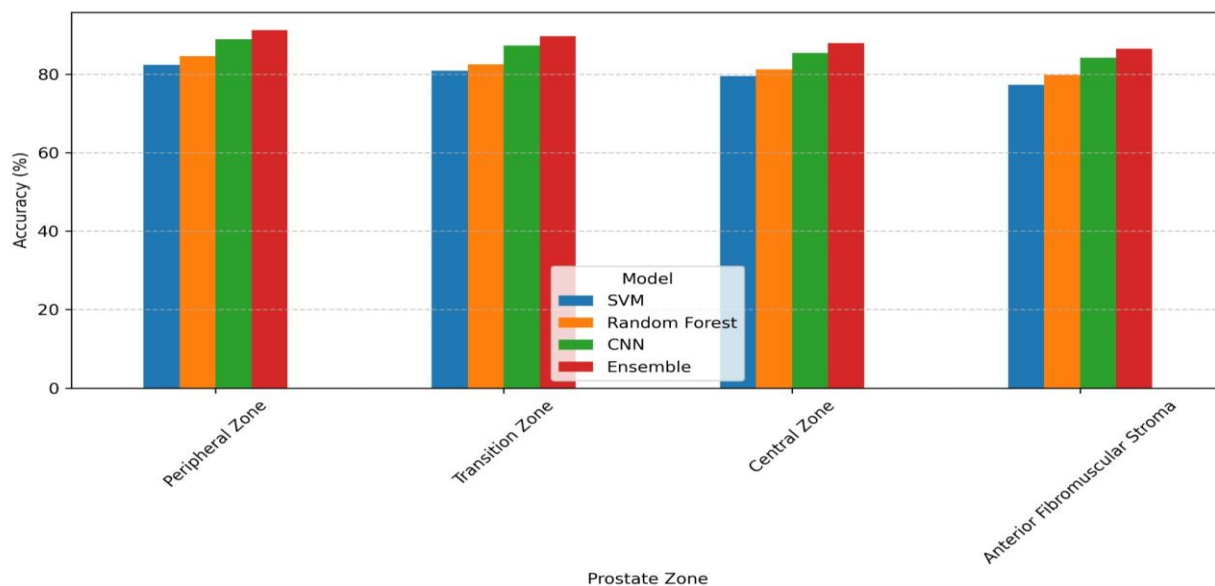


Figure 3: Prostate cancer detection accuracy by anatomical zone

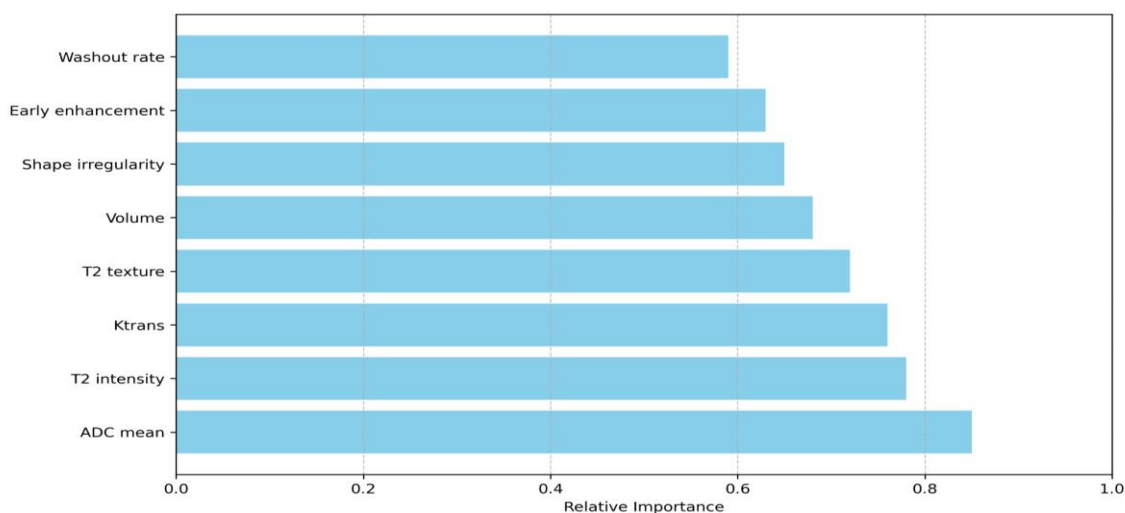


Figure 4: Feature importance for prostate cancer detection

BraTS 2021 supported segmentation and classification. Best results were obtained with T2weighted, diffusion-weighted, and dynamic contrast-enhanced multi-parametric MRI. The model showed zone-specific performance fluctuations. It was 91.2% accurate in the perimeter and 89.7% in the transition zone. Due to benign prostatic hyperplasia's varied signal qualities, transition zone tumors are difficult to detect.

The feature importance analysis showed that diffusion-weighted imaging (DWI) apparent diffusion coefficient (ADC) values were the most important factors in model performance, followed by DCE-MRI T2 signal intensity and enhancement kinetics.

Brain Tumor Classification and Segmentation

Case Study Utilizing Modified U-Net architecture with residual connections.

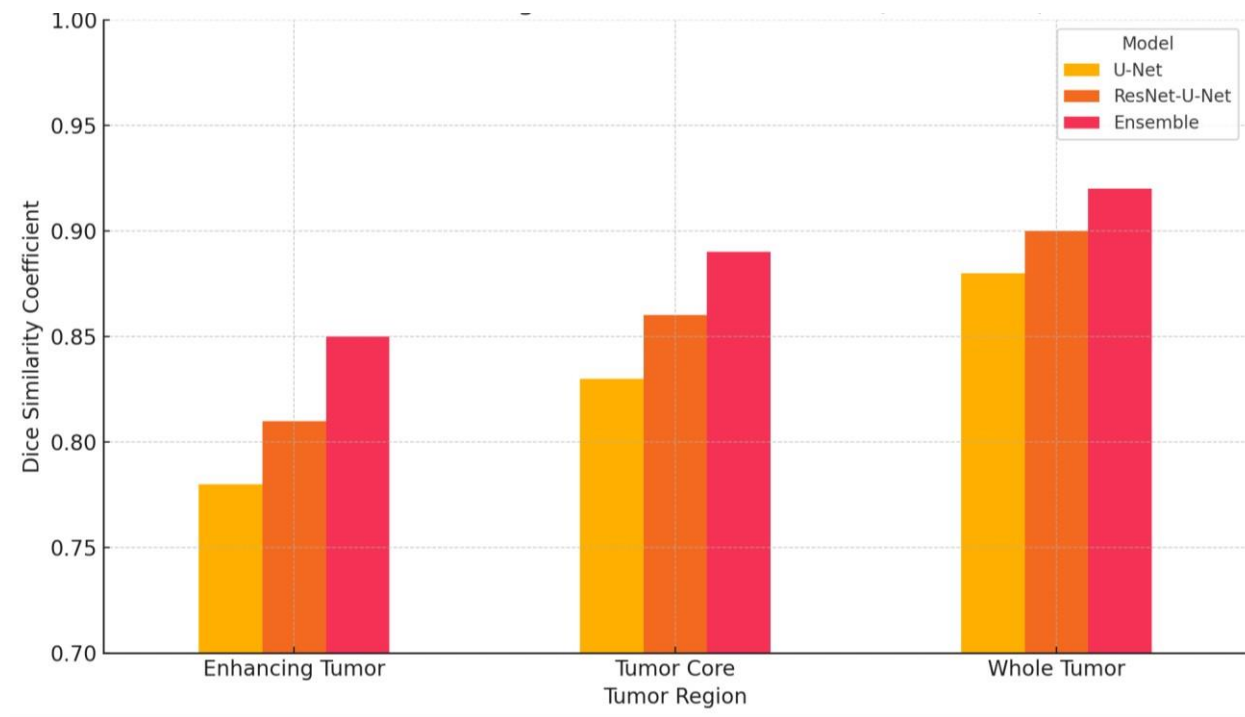


Figure 5: Brain tumor segmentation performance by region

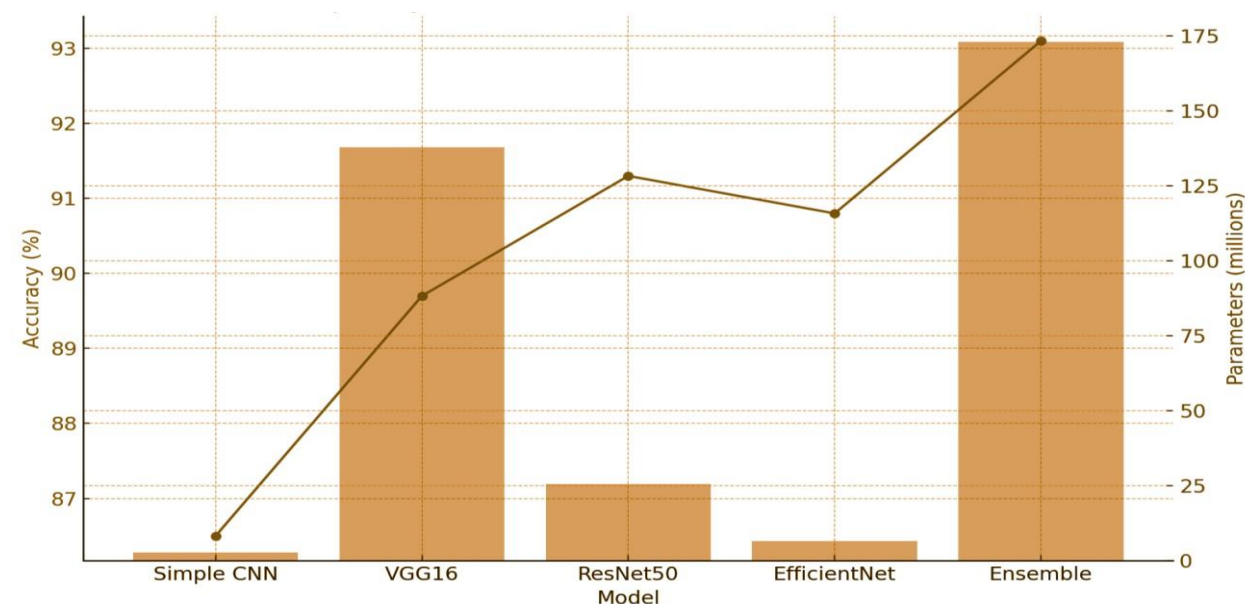


Figure 6: Model complexity vs. performance for brain tumor classification

In the segmentation results, the tumor volume was defined well (Dice coefficient 0.92) while the enhancing components were identified less well (Dice coefficient 0.85). For complex tumor margins, the ensemble technique dramatically increased segmentation accuracy.

The results of the computational efficiency research showed that although larger models typically achieved better performance, the link between the two was not linear. Instead, the relationship diminished beyond a certain point of model complexity. While retaining realistic inference durations that are appropriate for clinical translation, the best technique balanced performance against computing limitations, attaining 93.1% classification accuracy.

Combining Radiomic Techniques with Deep Learning Approaches

An integrated strategy was created that incorporated standard radiomics features with deep learning techniques; this integrated strategy was then evaluated based on its performance in comparison to standalone methods.

Table 2: Performance Comparison of Radiomics, Deep Learning and Hybrid Approaches

Approach	Cancer Type	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
Radiomics	Breast	84.5	83.2	85.8	0.865
Deep Learning	Breast	90.1	89.3	90.8	0.923
Hybrid	Breast	93.2	92.5	93.8	0.948
Radiomics	Prostate	82.9	81.4	84.3	0.854
Deep Learning	Prostate	89.6	88.7	90.4	0.915
Hybrid	Prostate	92.3	91.2	93.4	0.942
Radiomics	Brain	85.1	83.9	86.3	0.873
Deep Learning	Brain	92.5	91.8	93.2	0.946
Hybrid	Brain	94.7	93.9	95.4	0.965

When compared to solo radiomics and deep learning methods, the hybrid strategy was able to achieve better results across all cancer types. It was especially noticeable in situations in which the amount of labeled data was constrained, which implies that the information that is provided by radiomics features is extremely valuable and can supplement the skills of deep learning when it comes to pattern recognition.

CONCLUSION

This comprehensive analysis shows the huge potential of machine learning applications in magnetic resonance imaging (MRI)-based cancer characterization to increase cancer diagnosis accuracy and allow earlier identification across many cancer types. Ensemble techniques using radiomics and deep learning architectures work well in breast, prostate, and brain cancer cases. This strategy enhances accuracy by 5-10% over standard evaluation methods. This paper reports technical performance enhancements that potentially enable earlier intervention, more accurate characterization, and better risk classification, which could have therapeutic implications. Before widespread clinical implementation, model interpretability, external validation, and clinical process integration must be improved. Future research should focus on explainable AI technology, multimodal data integration, and prospective therapeutic efficacy trials. Federated learning offers novel techniques to manage data shortages and preserve privacy. As technology and clinical validation improve, ML-based MRI analysis can improve cancer care by enabling earlier detection and more precise characterization.

REFERENCES

1. He, M.; Cao, Y.; Chi, C.; Yang, X.; Ramin, R.; Wang, S.; Hu, K. Research progress on deep learning in magnetic resonance imaging–based diagnosis and treatment of prostate cancer: A review on the current status and perspectives. *Front. Oncol.* 2023, 13, 1189370.
2. Li, H.; Lee, C.H.; Chia, D.; Lin, Z.; Huang, W.; Tan, C.H. Machine learning in prostate MRI for prostate cancer: Current status and future opportunities. *Diagnostics* 2022, 12, 289.
3. American Cancer Society. Cancer Facts & Figures 2024. 2024. Available online:
4. <https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/2024cancer-facts-figures.html> (accessed on 17 January 2024).
5. Fernandes, M.C.; Yildirim, O.; Woo, S.; Vargas, H.A.; Hricak, H. The role of MRI in prostate cancer: Current and future directions. *Magn. Reson. Mater. Phys. Biol. Med.* 2022, 35, 503–521.
6. Suarez-Ibarrola, R.; Sigle, A.; Eklund, M.; Eberli, D.; Miernik, A.; Benndorf, M.; Gratzke, C. Artificial intelligence in magnetic resonance imaging–based prostate cancer diagnosis: Where do we stand in 2021? *Eur. Urol. Focus* 2022, 8, 409–417.
7. Kryvenko, O.N.; Epstein, J.I. Prostate cancer grading: A decade after the 2005 modified Gleason grading system. *Arch. Pathol. Lab. Med.* 2016, 140, 1140–1152.
8. Heijmink, S.W.; van Moerkerk, H.; Kiemeny, L.A.; Witjes, J.A.; Frauscher, F.; Barentsz, J.O. A comparison of the diagnostic performance of systematic versus ultrasound-guided biopsies of prostate cancer. *Eur. Radiol.* 2006, 16, 927–938.
9. Sieren, J.C.; Ohno, Y.; Koyama, H.; Sugimura, K.; McLennan, G. Recent technological and application developments in computed tomography and magnetic resonance imaging for improved pulmonary nodule detection and lung cancer staging. *J. Magn. Reson. Imaging* 2010, 32, 1353–1369.
10. Li, B.; Du, Y.; Yang, H.; Huang, Y.; Meng, J.; Xiao, D. Magnetic resonance imaging for prostate cancer clinical application. *Chin. J. Cancer Res.* 2013, 25, 240.
11. Turkbey, B.; Brown, A.M.; Sankineni, S.; Wood, B.J.; Pinto, P.A.; Choyke, P.L. Multiparametric prostate magnetic resonance imaging in the evaluation of prostate cancer. *CA Cancer J. Clin.* 2016, 66, 326–336.
12. Barbosa, F.D.G.; Queiroz, M.A.; Nunes, R.F.; Marin, J.F.G.; Buchpiguel, C.A.; Cerri, G.G. Clinical perspectives of PSMA PET/MRI for prostate cancer. *Clinics* 2018, 73 (Suppl. 1), e586s.
13. Manafi-Farid, R.; Ranjbar, S.; Jamshidi Araghi, Z.; Pilz, J.; Schweighofer-Zwink, G.; Pirich, C.; Beheshti, M. Molecular imaging in primary staging of prostate cancer patients: Current aspects and future trends. *Cancers* 2021, 13, 5360.
14. Zhu, L.; Pan, J.; Mou, W.; Deng, L.; Zhu, Y.; Wang, Y.; Xue, W. Harnessing artificial intelligence for prostate cancer management. *Cell Rep. Med.* 2024, 5, 101506.
15. Fischer, S.; Tahoun, M.; Kwaan, B.; Thierfelder, K.M.; Weber, M.A.; Krause, B.J.; Hamed, M. A radiogenomic approach for decoding molecular mechanisms underlying tumor progression in prostate cancer. *Cancers* 2019, 11, 1293.
16. Sadeghi, Z.; Alizadehsani, R.; CIFCI, M.A.; Kausar, S.; Rehman, R.; Mahanta, P.; Bora, P.K.; Almasri, A.; Alkhawaldeh, R.S.; Hussain, S.; et al. A review of Explainable Artificial Intelligence in healthcare. *Comput. Electr. Eng.* 2024, 118, 109370.
17. Mahmood, T.; Rehman, A.; Saba, T.; Nadeem, L.; Bahaj, S.A.O. Recent advancements and future prospects in active deep learning for medical image segmentation and classification. *IEEE Access* 2023, 11, 26143–26159.
18. Cifci, M.A. A deep learning-based framework for uncertainty quantification in medical imaging using the DropWeak technique: An empirical study with baresnet. *Diagnostics* 2023, 13, 800.
19. Alkadi, R.; Taher, F.; El-Baz, A.; Werghi, N. A deep learning-based approach for the detection and localization of prostate cancer in T2 magnetic resonance images. *J. Digit. Imaging* 2019, 32, 793–807.
20. Salvi, M.; De Santi, B.; Pop, B.; Bosco, M.; Giannini, V.; Regge, D.; Meiburger, K.M. Integration of deep learning and active shape models for more accurate prostate segmentation in 3D MR images. *J. Imaging* 2022, 8, 133.
22. Almestad, E. Exploring Explainable AI Adoption in Medical Diagnosis and the Empowering Potential of Collaboration. Master's Thesis, NTNU, Trondheim, Norway, 2023.