SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V13I9P103

Volume 13 Issue 9, 23-44, September 2026


*Original Article*

# Development and Evaluation of Enhanced Machine Learning Models for Tumor Severity Classification Through Integrated Feature Extraction and Selection Techniques

D. Sandhya Rani[1*], Anjaiah Adepu[2]

[1]*Department of Computer Science and Engineering, BEST Innovation University [BESTIU], Gorantla, Sri Sathya Sai District, Andhra Pradesh, India.*
[1]*Department of Computer Science and Engineering, CMR Technical Campus, Hyderabad, Telangana, India.*
[2]*HOD CSE & Principal IC, Polytechnic, Darbhanga, Bihar, India.*
[2]*Maulana Azad National Urdu University, A Central University, Ministry of Education, Government of India, Hyderabad, Telangana, India.*

[1*]*Corresponding Author : davu.sandhya@gmail.com*

Received: 13 March 2026 Revised: 17 June 2026 Accepted: 19 June 2026 Published: 29 September 2026

***Abstract*** *- The precise categorisation of the severity of tumours is an important function in medicine, as it is a vital part of deciding treatment and predicting patient outcome. We offer an integrated pipeline for developing and evaluating the latest machine learning methods for tumour severity classification. We apply advanced methods for feature extraction and selection to improve the model's accuracy, interpretability, and robustness. Ultimately, the framework organises diverse biomedical datasets for later analysis through domain-specific preprocessing approaches. Methods such as PCA, wavelet transform, and texture features are used to extract important information from raw data, thereby reducing dimensionality and preserving diagnostic information. More advanced techniques for feature selection are then used, such as Recursive Feature Elimination (RFE). Moreover, mutual information, which identify the most relevant features in the input dataset to reduce computational cost and the risk of overfitting. The models are validated using cross-validation techniques, with accuracy, precision, recall, F1-score, and Area Under the Receiver Operating Characteristic (ROC) Curve (AUC-ROC) as metrics. The results demonstrate that combining feature extraction and selection boosts classification accuracy and model generalisation across multiple datasets. Furthermore, explainability analysis clarifies the contribution of selected features to model predictions, which is crucial for building trust and transparency in clinical applications. However, really, what it demonstrates is the promise of all these things in concert. State-of-the-art machine learning, combined with rigorously designed features, will ultimately address some of those more difficult diagnostic challenges. These findings also pave the way for scalable, interpretable and clinically applicable AI-based solutions which may help in cancer detection and management. In this work, we will extend the proposed pipeline to other medical imaging modalities and explore real-time implementations.*

***Keywords*** *- Tumour severity classification, Machine learning models, Feature extraction, Feature selection, Oncology diagnostics.*

## 1. Introduction

Cancer continues to be a major source of illness and death globally, and therefore calls for better and faster diagnostic methods for early diagnosis and planning for therapy. Tumour severity measurement stands amongst the key challenges in cancer research and practice, driving treatment strategies and affecting patient outcomes. Conventional diagnostic procedures, though effective in some situations, can be subjective, time-intensive and susceptible to errors and inconsistencies. The advent of Machine Learning (ML) in cancer research has provided novel options for using computational modelling to explore large, complex biomedical datasets. These can detect intricate patterns otherwise not readily visible. However, the performance of ML-based classification models heavily relies on the features used. Biomedical data is typically high-dimensional and contains noise, redundant or irrelevant features, which can have a detrimental effect on model accuracy and

generalisation. Traditional methods are often trustworthy in certain contexts but are highly reliant on the individual knowledge from which they originate, time-consuming, and error-prone. Conversely, the advent of machine learning in oncology offers powerful alternatives to this problem by leveraging computational power and rich, multidimensional data to achieve high classification accuracy. Machine learning has advanced over the years, enabling it to extract complex patterns in data that are not visible to the human eye. However, two aspects affect their accuracy: the quality of the informative features and their relevance to predicting the outcome. In this regard, feature engineering, especially feature extraction and feature selection, improves the performance of the model. They help reduce the dimensionality problem, remove spurious features, and increase the distinctiveness of input features, thus enabling efficient and effective classification.

However, while current techniques are improving, an inherent issue with ML-based tumour classification is the trade-off between model complexity and interpretability. Complex models like deep neural networks may perform better, but are less interpretable and therefore less useful to clinicians. Simpler models are more interpretable, but may not capture complex underlying patterns. Moreover, problems such as overfitting and underfitting affect model performance. However, complex models can pose risks of overfitting and underfitting. In contrast, simpler models are more interpretable but less appropriate for our data. Thus, it requires a pipeline that maximises the use of advanced feature-engineering methods with ML models while considering performance, scalability, and interpretability in tumour severity classification.

Beyond Accurate Diagnosis and The Importance of Severity of Tumour Classification. Early-stage tumours may be best managed conservatively, whereas advanced-stage tumours may be treated with more aggressive modalities. Such misclassification can lead to suboptimal treatment decisions with a deleterious impact. Thus, in tandem with accumulating clinical data, our results underscore the importance of establishing robust, data-driven classification models to improve patient care and the efficient allocation of healthcare resources.

Although these advances have been made, the approaches still lack an integrated framework that effectively combines state-of-the-art feature engineering with optimised machine learning models to balance accuracy, scalability and interpretability. Additionally, many existing approaches either explore feature optimisation or classification model performance separately, rather than studying their combined effects on classification accuracy for tumour severity. This highlights an important gap: the need for an integrated approach to achieve a data-driven biomedical classification system that is efficient and clinically relevant. In this context, this study offers an integrated approach to enhance tumour severity classification through feature extraction, feature selection, and a selection of machine learning models. This framework involves pre-processing biomedical data, feature space optimisation and testing various classifiers to determine the best model. Moreover, the study highlights the importance of model interpretability and validation for clinical relevance. This knowledge gap is addressed by developing a unified approach that integrates feature engineering and machine learning to enhance tumour severity classification.

The main contributions of this paper are:

- Feature engineering approach for high-dimensional biomedical data,
- Benchmarking of a range of ML models for tumour severity classification,
- Feature selection strategies to improve model performance and efficiency, and
- Focus on interpretability to support clinical use.

The rest of this paper is structured as follows: Section 2 discusses related work and currently Active techniques in tumour classification. The proposed approach, including data preprocessing, feature engineering and evaluation, is outlined in Section 3. Section 4 Proposed Solutions. Section 5 Proposed Algorithms and Frameworks, Section 6 discusses the limitations, and Section 7 concludes with future work directions. By combining advanced feature engineering and cutting-edge ML algorithms, this paper seeks to establish a foundation for experimentation with different ML paradigms for tumour diagnosis, a vital step toward creating scalable, high-performance yet explainable AI systems in medicine. Not only do the findings help fill gaps in tumour severity classification, but they also serve as a stepping stone for future work in medical AI.

## 2. Related Works

In recent years, the field of medicine has been incorporating Machine Learning (ML) algorithms for medical diagnostics, particularly in oncology. Abstract: To address tumour classification challenges, recent advances in data acquisition and computing power have led researchers to explore several ML approaches. High dimensionality and heterogeneity of biomedical datasets have made feature engineering (e.g., feature extraction and feature selection) an indispensable aspect of these endeavours 3 4. Previous studies indicate that effectively integrating distinct feature extraction methods (e.g., texture-oriented analysis, wavelet transforms) with powerful feature selection techniques (e.g., Recursive Feature Elimination (RFE), mutual information) can enhance model performance. However, today's approaches often struggle with overfitting, lower generalizability, and limited interpretability in clinical practice. In this section, we summarise recent efforts and obstacles in this fascinating

research domain and highlight gaps in the literature that motivate the need for an integrated tumour severity classification framework. The literature is divided into three broad directions to gain a better understanding of research: (i) deep learning-based tumour classification, (ii) hybrid and ensemble machine learning methods, and (iii) feature engineering and search-based approaches to enhance the classification accuracy.

### *2.1. Deep Learning-based Tumour Classification*

The study by Christine Ecker et al., used whole-brain structural MR scans as predictors of autism using pattern classification methods ([1]). Here, the authors used advanced neuroimaging methods to demonstrate that machine-learning algorithms could distinguish between individuals with autism and neurotypical controls. Keeping in mind structural abnormalities associated with brain use, imaging methods important for clinical diagnosis, and the potential to extract them using automated systems, this study suggested predictive biomarkers. This is an important step towards a platform for computational tools and predictive analytics in autism spectrum disorder and neurodevelopmental disorders.

Ahmed et al. A novel Laplacian Re-Decomposition and XGBoost-based model for early diagnosis of Alzheimer's disease [2]. This enables the best of both worlds, offering computational efficiency while achieving high accuracy in biomarker detection for Alzheimer's disease. The characteristic-extraction capability of the Advanced re-decomposition mechanism and the predictive capacity of XGBoost hold financial potential for detecting neurodegenerative diseases. This work emphasises the need for systems capable of handling the high-dimensional data characteristic of brain imaging and clinical diagnostics.

### *2.2. Hybrid and Ensemble Learning Approaches*

A hybrid intelligent system based on adaptive neuro-fuzzy for improving online brain tumour prediction has been proposed by Nagendiran and Chokkalingam [3]. They fused the adaptivity of fuzzy logic with neural networks, achieving significant prediction accuracy for tumour presence detection. Chapter 5 is a detailed exploration of their system, which integrates real-time analysis for clinical diagnostics applications where timely decision-making is critical. Integrated computational intelligence shows promising prospects for addressing the complexities of medical imaging, as discussed in the present study.

In a recent work published in [4], Deepa and Chokkalingam analysed the stage-wise impact of dementia and non-dementia-related age-associated functional disabilities. They measured their severities through supervised deep learning models. This is the first study to provide a holistic perspective on the evolution of cognitive decline, and it is important information for how AI development can be targeted to the task of tracking and forecasting dementia. Through the development of a deep-learning algorithm, predictive accuracy was improved, ensuring optimal resource utilisation and individualised treatment plans for patients. Muthu Krishnammal and Raja [5] use a neural network-based, fully automated brain image classification system for abnormality detection. The authors integrated feature extraction and classification, enabling an accurate and efficient pipeline for identifying anomalies in brain imaging in their study. Conclusion: This study demonstrates the potential of neural network-based systems for medical diagnosis, particularly for diseases requiring rapid treatment, such as brain tumours.

Sasi Kumar and Aithal [6] presented a framework for providing deep-Q learning-based completion on a convolutional neural network for the detection and prediction of brain diseases. As their solution was aimed at scalability, they used a cloud-based implementation through Colab, thus making their approach Reachable and Economic. Further, as reinforcement-learning optimisation using the deep Q-learning component leads to better model generalisation, the model was applied to diagnose increasingly complex brain disorders, followed by deep Q-learning to further optimise the model. A deep learning framework based on Spark classified the severity of brain tumours in MRI images [7], introduced by Abirami and Prasanna Venkatesan. The project addressed the big data challenge by using Spark to enable distributed processing, making it a good candidate for medical imaging analytics. This paper demonstrates the synergy of HPC and DL in practical clinical application settings.

De Groot et al. used the SORG machine learning algorithm to predict extremity metastases in a contemporary cohort of patients [8]. Temporal validation: They validated the algorithm using future data, making it fair and resilient to change. The paper, published in the journal Nature Machine Intelligence, adds that in dynamic healthcare environments, "this highlights the value of ensuring that machine learning models remain well-calibrated with time."

Dalal et al. [9] developed a hybrid machine learning model to detect breast cancer early, combining multiple algorithms to achieve higher accuracy. Breast cancer is one of the most commonly diagnosed cancers. Although modern treatment methods have changed how this disease is viewed, patients present late to treatment, resulting in poor outcomes, and early detection systems could improve outcomes. They demonstrated fewer false positives and greater overall diagnostic accuracy with their hybrid approach.

### *2.3. Feature Engineering and Optimisation Techniques*

For histopathology images, Deepak and Bharanidharan [10] employed ensemble machine learning methods for osteosarcoma detection. Their study results provided a comparative context for the effectiveness of numerous ensemble methods in screening and annotation of high-

dimensional histopathological data. This construction supplements the resources required in this age of genome-transfusing carcinoma diagnostics, wherein modified dynamic computation approaches complement high-resolution tissue imaging. Mactina and Neduncheliyan [11] introduced an optimised RNN model for brain tumour classification and merged cat and rat swarm optimisation approaches. Such bio-inspired resolution enhanced the RNN's performance in classifying them precisely. They developed a state-of-the-art method for diagnosing brain tumours that combines the advantages of evolutionary algorithms and deep learning.

Ravi et al. [12] Specifically, it concentrated on the classification and analysis of the pathological images of breast cancer using Jubilee, machine learning methods, and so forth. It could also be fully integrated into the gradual development of the picture preparation process. This study is designed to obtain and analyse important pathological properties that correlate with breast cancer, to provide a basis for the interpretation of breast cancer-image-Based Machine Learning for Precision Oncology. Gupta et al. [13] used texture and radiomics-inspired data in this way to predict lung-nodule malignancy. Their model combined traditional radiomics with state-of-the-art machine learning practices to develop a hybrid model with superior accuracy for diagnosing HCC. One of the strengths of the research is the multimodal data separation method, which uses image, text, and clinical features, with data types separated across modules within layers.

Sakthi et al. proposed a DCNN-based framework in [14] to classify brain tumours from MRI images. This leads to the new field of deep learning and its potential to automate medical diagnosis and the associated identification and classification of disease (379). The MRI scans analysed in this study require high-performance computational tools for important diagnostic applications. Saravanan et al. Zhang et al. proposed a K-CNN-BTSL model. [15] to classify the severity level of a brain tumour. This deep learning technique used salient features extracted from MRI images to obtain an extensive evaluation of tumour accentuation. Their work illustrates why neural networks designed for specific medical issues, such as the progression of brain tumours, can yield valuable insights.

Azade et al. An extensive overview of brain tumour detection and segmentation techniques was described in [16]. They also identify deficiencies in the current literature and offer a pathway to future research. Researchers warn in a new study that there is an urgent need for novel measurement approaches to enhance diagnostic accuracy in brain imaging. In this paper, Bindu and Sastry [18] implemented a hybrid machine learning model for brain tumour classification by combining different algorithms to increase accuracy. Their approach addresses issues with data variability and noise present in brain imaging systems. Therefore, this hybrid approach in the study is an appealing example of the fusion of classical and modern computational pathway methods for meaningful diagnosis. Soundarya et al. Further authors, (e. g. Rehman et al. Transfer learning has also shown promise in medical imaging, and the authors of [18] applied it to brain tumours. By leveraging pre-trained models, the study significantly reduced training time and achieved high diagnostic accuracy, including clinically high accuracy.

In this study, we demonstrate the feasibility of transfer learning, which is critically important in resource-poor medical environments. Aishwarya et al. Meta-Heuristic Transformer Networks for MR Images (MT-MRAN): In this paper [20], the authors proposed combining a meta-heuristic optimisation technique with a transformer-based network to classify brain tumours. Hence, this method is much more precise and validates the use of multi-scale feature extraction in complicated (medical) diagnosis. However, this needs to be tempered with the knowledge that these nuanced medical challenges require careful inference and more complex network architectures.

Ramesh Kumar et al. Brain tumour segmentation can be achieved using the deep joint RP-Net segmentation algorithm proposed in [21], which predicts brain tumour severity from the segmentation. They mention advanced segmentation techniques to ensure accurate diagnosis. Prasanna Kumar et al. [22] conducted a comparative study of brain tumour classification methods using MR images, providing fundamental insights into optimisation in diagnostic methods in medical imaging research. Feature Extraction and Classification of Grey-Scale Brain Tumour Images through Deep Learning [23].

Pranitha and Vurukonda [23] used a DNN to perform feature extraction and classification of grey-scale brain tumour images, achieving good accuracy and demonstrating the potential of DNNs for complex imaging datasets. Zahoor et al. Res-BRNet [24] focuses on brain tumour classification from MRI, utilising deep residual and regional CNNs embedded within a multi-category integration of spatial and residual blocks for extracting texture and boundary features. Validated only on public benchmark datasets without clinical or multi-institutional testing; High accuracy. Arumugam et al. used a crossover-optimised multilayer perceptron to classify brain tumours [26]. Their work demonstrates that optimisations can genuinely improve performance in neural networks, including diagnostic accuracy. Mathivanan et al. sand colleagues [27] presented an ensemble deep learning model based on modified MobileNetV2 for automated brain tumour grade classification from MRI images, addressing both binary and multi-class tumour classification. Though we achieved competitive accuracy with this model, validation was limited to a single imaging modality and a small dataset, which limits generalizability. Gasmi et al. implemented a genetic algorithm–based weight optimisation for an ensemble of

EfficientNet-based deep learning models for multi-class brain tumour classification [28]. While this approach is pragmatic and computationally intensive, it has not yet been validated in a chronic clinical setting. In general, current studies show substantial improvements in tumour classification using machine learning approaches. However, these methods are typically domain-specific in their focus on model performance, feature engineering or data management. Most research has not adequately considered feature engineering and model selection in a seamless way, incorporating model interpretability and scalability. Moreover, many of the proposed models are trained and validated on small or focused datasets, and the models' generalizability in clinical settings remains unclear. These shortcomings underscore the need for a holistic data-driven approach that delivers higher classification accuracy, robustness and practicality The recent related work is summarised here (Table 1).

**Table 1. Summary of the recent related works**

| Author, Year [Ref] | Proposed Method | Research Limitations |
|---|---|---|
| Christine Ecker et al., 2013 [1] | Pattern classification approach on whole-brain structural MR scans for autism prediction. | Limited generalisation due to small sample size and lack of external validation. |
| H. Ahmed et al., 2023 [2] | Laplacian Re-Decomposition with XGBoost for early Alzheimer's detection. | High computational complexity; not tested on diverse datasets. |
| D. Nagendiran and S.P. Chokkalingam, 2022 [3] | Adaptive neuro-fuzzy system for real-time brain tumour prediction. | Susceptible to noise in imaging data; limited scalability. |
| N. Deepa and S.P. Chokkalingam, 2021 [4] | Deep learning model to assess dementia stages and severity. | Lacks longitudinal validation; limited interpretability. |
| P. Muthu Krishnammal and S. Selvakumar Raja, 2015 [5] | Neural network-based automated brain image classification. | Simplistic architecture; limited feature extraction capabilities. |
| A. Sasi Kumar and P. S. Aithal, 2022 [6] | DeepQ convolution neural network for brain disease detection. | Dependent on cloud computing, there are potential privacy issues. |
| S. Abirami and D.G.K.D. Prasanna Venkatesan, 2022 [7] | Spark architecture and deep learning for MRI image severity classification. | Performance issues with larger datasets; requires high computational resources. |
| T.M. de Groot et al., 2023 [8] | Temporal validation of the SORG algorithm for extremity metastases. | Limited applicability to non-extremity metastases. |
| S. Dalal et al., 2023 [9] | Hybrid machine learning model for breast cancer prediction. | Limited explainability of predictions; dataset-specific biases. |
| K.V. Deepak and R. Bharanidharan, 2023 [10] | Ensemble machine learning techniques for osteosarcoma detection. | Dependency on high-quality histopathology images; limited real-world testing. |
| J.N. Mactina and S. Neduncheliyan, 2024 [11] | Recurrent neural network with cat and rat swarm optimisation for brain tumour classification. | Performance is limited by complex optimisation; it requires high computational power. |
| P. Ravi et al., 2024 [12] | Machine learning-based classification and analysis of breast cancer pathological images. | Dataset imbalance and limited model interpretability. |
| H. Gupta et al., 2024 [13] | Radiomics-inspired data-driven classification of lung nodule severity. | Lack of longitudinal validation and high dependency on radiomics feature quality. |
| U. Sakthi et al., 2023 [14] | Deep Convolutional Neural Network (DCNN) framework for brain tumour classification. | High computational demands and limited model explainability. |
| M. Saravanan et al., 2023 [15] | K-CNN-BTSL model for brain tumour severity classification using MRI images. | Limited robustness to noisy data; requires extensive parameter tuning. |
| A. Azade et al., 2023 [16] | Comprehensive review of brain tumour detection and segmentation techniques. | Primarily theoretical; lacks experimental validation. |
| Y. Modaresnia et al., 2023 [17] | EfficientNetB0 hybrid approach for brain tumour classification using deep learning and bagging trees. | Limited scalability to large datasets; requires more diverse data for validation. |
| N.P. Bindu and P.N. Sastry, 2023 [18] | Hybrid machine learning techniques for automated brain tumour classification. | Susceptible to overfitting; requires extensive preprocessing. |

| | | |
|---|---|---|
| C. Soundarya et al., 2023 [19] | Transfer learning for brain tumour detection using pre-trained models. | Dependency on pre-trained models; limited customisation for specific datasets. |
| R. Aishwarya et al., 2023 [20] | MT-MRAN framework for brain tumour classification using transformer-based networks. | High training complexity; requires extensive computational resources. |
| R. Ramesh Kumar et al., 2024 [21] | Deep joint RP-Net-based segmentation for brain tumor severity prediction. | Limited validation on diverse datasets; requires robust parameter optimisation. |
| G. Prasanna Kumar et al., 2024 [22] | Comparative analysis of brain tumor classification through MR imaging. | Lacks application-oriented testing and real-time validation. |
| K. Pranitha and N. Vurukonda, 2024 [23] | Feature extraction and classification of grayscale brain tumor images. | Limited ability to handle multi-class classifications. |
| Zahoor et al., 2024 [24] | Deep residual and regional CNN (Res-BRNet) for brain tumor MRI classification across glioma, meningioma, and pituitary tumor types. | Limited to public datasets; lacks multi-institutional validation; interpretability not addressed. |
| R. Rajeswari et al., 2024 [25] | Dense fused maxout network for severity prediction of brain tumors. | High computational demands; limited scalability to large datasets. |
| M. Arumugam et al., 2024 [26] | Smell agent-optimized multilayer perceptron for brain tumor classification. | Complexity in optimisation requires high computational resources. |
| M. Sankar et al., 2024 [27] | Ensemble deep learning model (modified MobileNetV2) for brain tumor grade classification from MRI, covering binary and multi-class schemes. | Limited to a single imaging modality; restricted dataset diversity; limited cross-institutional validation. |
| Gasmi Karim et al., 2024 [28] | Weighted ensemble of EfficientNet-based models with genetic algorithm-based weight optimisation for multi-class brain tumour classification. | High computational cost; limited scalability to heterogeneous datasets; lacks longitudinal validation. |

Our literature survey results, identified in Table 1, show that existing methods mostly focus only on model or feature engineering. Deep learning approaches have high classification accuracy but are expensive and non-interpretable for clinical use. Hybrid and ensemble approaches enhance performance but may require substantial preprocessing and are prone to overfitting. Moreover, many approaches work only on specific datasets and struggle to generalise across large biomedical datasets. In this work, however, the proposed solution overcomes these challenges by leveraging feature engineering and selection, integrated with the systematic evaluation of a diverse range of models, to improve classification. This leads not only to improved classification performance but also to scalability, reduced dimensionality, and interpretability, which are not achieved jointly in current studies.

This study's main contribution is its systematic integration of feature engineering and machine learning for tumour severity classification. Moreover, instead of isolating feature optimisation and modelling tasks, this work integrates the two stages together and assesses their combined effect on classification results. Finally, the proposed method not only considers model accuracy but also interpretability and scalability, making it more relevant to clinical applications.

### *2.4. Research Gap and Proposed Solution*

The review of the existing literature shows that most studies focus on improving model performance or on optimising individual feature engineering techniques, without considering how these choices affect tumour severity classification performance. Furthermore, existing approaches have limitations, including computational efficiency, interpretability, overfitting, and poor generalisability across various biomedical data. To overcome such limitations, we have proposed a framework that combines feature extraction and selection techniques with model selection optimisation in machine learning. This allows for better feature representation, dimensionality reduction, and more accurate classification performance, with an emphasis on interpretability and scalability. The proposed method, unlike existing approaches, optimises feature engineering and model selection in a unified manner, thereby offering a more effective and clinically relevant approach to tumour severity prediction.

## 3. Research Problems

The classification of tumour severity is a challenging task with critical importance for timely treatment and patient outcome in medical diagnostics. Concise: A simple ML model for histopathological prediction of tumour severity. Multiclass Market and Clinical Ensemble, a simple ML model for histopathological prediction of tumour severity. Despite rapid advances in medical imaging and computational technologies, most current Machine Learning (ML) models for tumour severity classification face limitations in clinical applicability. The limitations include low predictive accuracy, limited generalisation to heterogeneous patient populations, and the inability to extract high-dimensional medical data. The inherent high dimensionality of medical imaging data,

accompanied by noise and irrelevant dimensions, poses both computational and clinical challenges for accurate classification.

There is one area among the two most important steps in building an ML model: feature extraction and selection, which are usually done in isolation or using general methods. These techniques neither exploit domain-specific intuitions nor aggregate complex algorithms to mitigate accuracy. This can lead to inadequate performance of currently available models, due to their inclusion of redundant or non-informative features that may hinder the identification of informative patterns associated with tumour aggressiveness. In the clinical setting, with highly diverse tumour phenotypes, a feature engineering pipeline with relatively few candidate features to test must be robust; it should ideally be only the initial step in the step-wise pipeline.

Deep learning has recently demonstrated unprecedented performance across many application domains. Still, its black-box nature raises concerns about explainability and trust in the medical applications ecosystem. Clinically, we require translational models, interpretable models, and decision explainability, with transparency to enable incorporation into established routine diagnostic workflows. However, existing solutions are inadequate for addressing this need and significantly hinder their uptake in clinical settings.

Another significant challenge is the variability in tumour data, particularly across imaging modalities, acquisition protocols, and patient populations. This variation is among the most common causes of model overfitting, where models fit the training data but do not generalise to new cases. Lack of stringency in evaluation criteria further complicates assessment of model efficacy, leading to divergent findings between studies and poor reproducibility. Such scenarios underscore the need to quickly develop and evaluate better ML models for tumour severity classification. The models built should be capable of leveraging advanced feature extraction and selection methods, combining domain knowledge and algorithms to provide high accuracy, stability, and interpretability. This research should fill the critical gaps in the ability to design pipelines that: (1) perform Dimensionality reduction, which helps to select only important features; (2) ease the complexity of model architecture to make it more clinically applicable.

To tackle these issues, the proposed study presents a unifying framework that integrates all elements of feature extraction, selection, and classification into a single continuous pipeline. In addition to advanced selection methods such as mutual information, principal component analysis, and regularisation techniques, this framework will be based on domain-specific feature engineering strategies. That frames our long-term goal of creating models that exceed performance benchmarks while also providing accessibility for interpreting tumour grading. In addition, they argue for strong evaluation methods, such as cross-validation, testing on held-out datasets, and domain-specific metrics, to ensure reliability and robustness. Lastly, implementing sophisticated, interpretable models via Explainable AI (XAI) methods will address the growing need for interpretability to build clinicians' trust and avoid the hurdles of introducing such methods into clinical practice.

By addressing these challenges in the proposed research, we aim to bridge the gap between theoretical innovation in machine learning and practical implementation in oncology. These results can shift tumour-grade classification and, consequently, spur advances in precision medicine and patient care.

## 4. Proposed Solutions

Probabilistic Tumour Severity Classification Approach: The approach is to build more powerful machine learning models that leverage systematic, data-driven feature extraction and selection, followed by classification with state-of-the-art classifiers. The focus of the proposed framework will be on state-of-the-art problems in tumour severity classification, including high-dimensional medical data, model interpretability, and generalisation across different datasets. The proposed methodology begins with the collection and pre-processing of tumour imaging datasets, followed by normalisation and data augmentation to ensure data quality and consistency.

Detection of clinically relevant patterns (e.g., texture analysis and radiomic feature computing) using advanced feature extraction techniques. Next, using feature selection methods (for example, mutual information and dimensionality reduction algorithm (PCA)), we make sure that the feature set is suitable for classification. You also train for each of those features as best as possible, the explainable AI, which is integrated in machine learning features, which is like Windows explainable AI features: Predict (saw like "height" for explaining predictions), Deterministic model-evaluation, like cross-validation and external validation are performed with independent datasets to ensure the robustness and clinical applicability. To achieve this, the approach aims to create a comprehensive pipeline that optimises accurate, efficient, and interpretable outputs to meet the needs of clinical practitioners.

### *4.1. Feature-Integrated Tumour Classification Network (FIT-Net)*

Lastly, as FIT-Net integrates feature prediction and selection into a single network, it has further improved prediction accuracy for classifying tumour severity. The convolutional layers can then be used to begin extracting patterns from the imaging data, and the feature selection layer ensures that each output of the network is relevant and focused. This reduces noise, reduces computational

complexity and overfitting, and also ensures better performance of models on other datasets. This adaptive feature learning is performed purely data-driven, eliminating any kind of human bias, and as the diversity of selected domains is assured, it provides domain-specific interpretations, ensuring robust classification.

This step applies convolutional operations to learn spatial patterns.

$$Z = \sigma(W \cdot X + b) \quad (1)$$

This step multiplies feature maps by learned importance weights.

$$Z_{selected} = \phi(Z \cdot \alpha) \quad (2)$$

This step projects the features into a lower-dimensional space.

$$Z_{reduced} = Z_{selected} \cdot V \quad (3)$$

This step maps reduced features to classification probabilities.

$$Y = \text{softmax}(Z_{reduced} \cdot W_c + b_c) \quad (4)$$

This step calculates the classification error for optimization.

$$L = -\sum_{i=1}^{n} y_i \log(\hat{y}_i) \quad (5)$$

This step penalizes large weights to improve generalization.

$$R = \lambda \|W\|^2 \quad (6)$$

This step defines the optimization target for training.

$$J = L + R \quad (7)$$

This step refines model parameters based on gradient calculations.

$$W \leftarrow W - \eta \frac{\partial J}{\partial W} \quad (8)$$

### *4.2. Explainable Feature Optimization Model (EFOM)*

EFOM assumes that the interpretability and trust in the model could be improved by the explainability mechanisms' functionalities, which are part of the model's general objective, together with feature selection and optimisation. This unique algorithm employs an attention-based mechanism to evaluate feature significance and perform feature pruning to select beneficial features while removing noise. In this work, we introduce an evolutionary process that eliminates redundancy by evolving a suitable feature subset and increases classification accuracy while maintaining the transparency of the decision process through feature attributions.

This step calculates the relevance of each input feature.

$$S = \text{softmax}(W_s \cdot X) \quad (9)$$

This step removes less relevant features from the dataset.

$$X_{pruned} = X \cdot \mathbb{I}(S > \tau) \quad (10)$$

Generate interpretable explanations for feature contributions.

$$E = \sum_{j=1}^{d} S_j \cdot \phi_j \quad (11)$$

Map optimized features to severity labels.

$$Y = \text{softmax}\left(W_c \cdot X_{pruned} + b_c\right) \quad (12)$$

Minimize loss while ensuring explainability constraints.

$$J = L + \lambda \sum_{j=1}^{d} (1 - S_j)^2 \quad (13)$$

### *4.3. Robust Dimensionality and Prediction Algorithm (RDPA)*

The central idea behind RDPA is that if predictions are optimally modeled, Dimensionality reduction can be sufficiently powerful to improve classification performance. RDPA applies unsupervised techniques for both dimensionality reduction and supervised classification in order to reduce the computational burden while also improving the generalisability of the model, in particular, when the data is high-dimensional. Apply Singular Value Decomposition (SVD) to reduce feature dimensions.

$$X_{reduced} = U \Sigma V^T \quad (14)$$

Reweight reduced dimensions based on the variance explained.

$$W_{dim} = \frac{\Sigma}{\text{trace}(\Sigma)} \quad (15)$$

Train a Support Vector Machine (SVM) on reduced features.

$$Y = \text{sign}(W_{svm} \cdot X_{reduced} + b_{svm}) \quad (16)$$

This step penalizes large weights to improve generalization.

$$R = \lambda \|W_{svm}\|^2 \quad (17)$$

Combine SVM loss and regularization into a single objective function.

$$J = L_{svm} + R \quad (18)$$

## 5. Proposed Algorithms and Frameworks

Propose algorithms and frameworks that introduce the state-of-the-art feature extraction, selection and classification methods into a single framework to address the important problems of tumor severity classifications. Fella combines ML models with traditional methods and implements explainable AI to make predictions more trustworthy and transparent.

Two frameworks claim that they can boost model accuracy while being cost-efficient, given the dimensions they reduce to; these are strong dimensionality reductions added to dynamic feature optimisation.

The proposed approaches are also scalable, apply to a broad array of datasets, and are clinically relevant. Here, we propose three algorithms: (1) Feature-Integrated Tumour Classification Network (FIT-Net); (2) Explainable Feature Optimisation Model (EFOM); and (3) Robust Dimensionality and Prediction Algorithm (RDPA) to form an integrated pipeline for optimising the classification of the severity of tumours.

All together, they represent a complete solution for accurate and interpretable diagnostics, with each one focusing on a specific limitation, all achieved by incrementally improving existing methods.

A Novel Mendeley specification for Integrated Feature Extraction and Selection (FIT-Net) for Improved Tumour Classification Task -direct Integration of Feature Extraction and Feature Selection into a single Pipeline.

This helps you to remove irrelevant data, will remove unnecessary noise, improve computational efficiency, and also bestows you with accurate predictions.

FIT-Net incorporates convolutional layers that detect spatial and structural properties, feature selection mechanisms that enhance critical data, and high-capacity classification layers that predict severity categories for time series data. These elements enable FIT-Net to achieve high accuracy and interpretability across multiple clinical scenarios.

| **Feature-Integrated Tumor Classification Network (FIT-Net)** |
|---|
| **Input:**<br>• High-dimensional feature tumor imaging datasets. |
| **Output:**<br>• Estimated categories of tumor severity |
| **Assumptions:**<br>• Inconsistent and Imbalanced Data Handling |
| **Improvements over Existing Algorithms:**<br>• The method integrates feature extraction and feature selection, which reduces computational overhead and improves accuracy in comparison to the individual methods. |
| **Process:**<br>Step - 1. Load tumor imaging datasets.<br>Step - 2. Applying augmentation and normalizing pixel intensities.<br>Step - 3. Extract spatial patterns with convolutional layers.<br>Step - 4. Extracted features are stored in a structured format for further processing.<br>Step - 5. Use attention mechanisms to weigh the importance of features.<br>Step - 6. Keep only features that have significant weights.<br>Step - 7. Using the PCA techniques to avoid redundancies, it reduces the dimensionality of features.<br>Step - 8. Notably, feed all optimised features into fully connected layers for severity level classification.<br>Step - 9. Output layer with softmax for probabilities.<br>Step - 10. Cross-validation and using external datasets for testing<br>Step - 11. Choose the hyperparameters based on the results on the validation set.<br>Step - 12. Give final severity predictions, and understand how features contribute. |

*The proposed algorithm is visualised graphically here (Figure 1).*

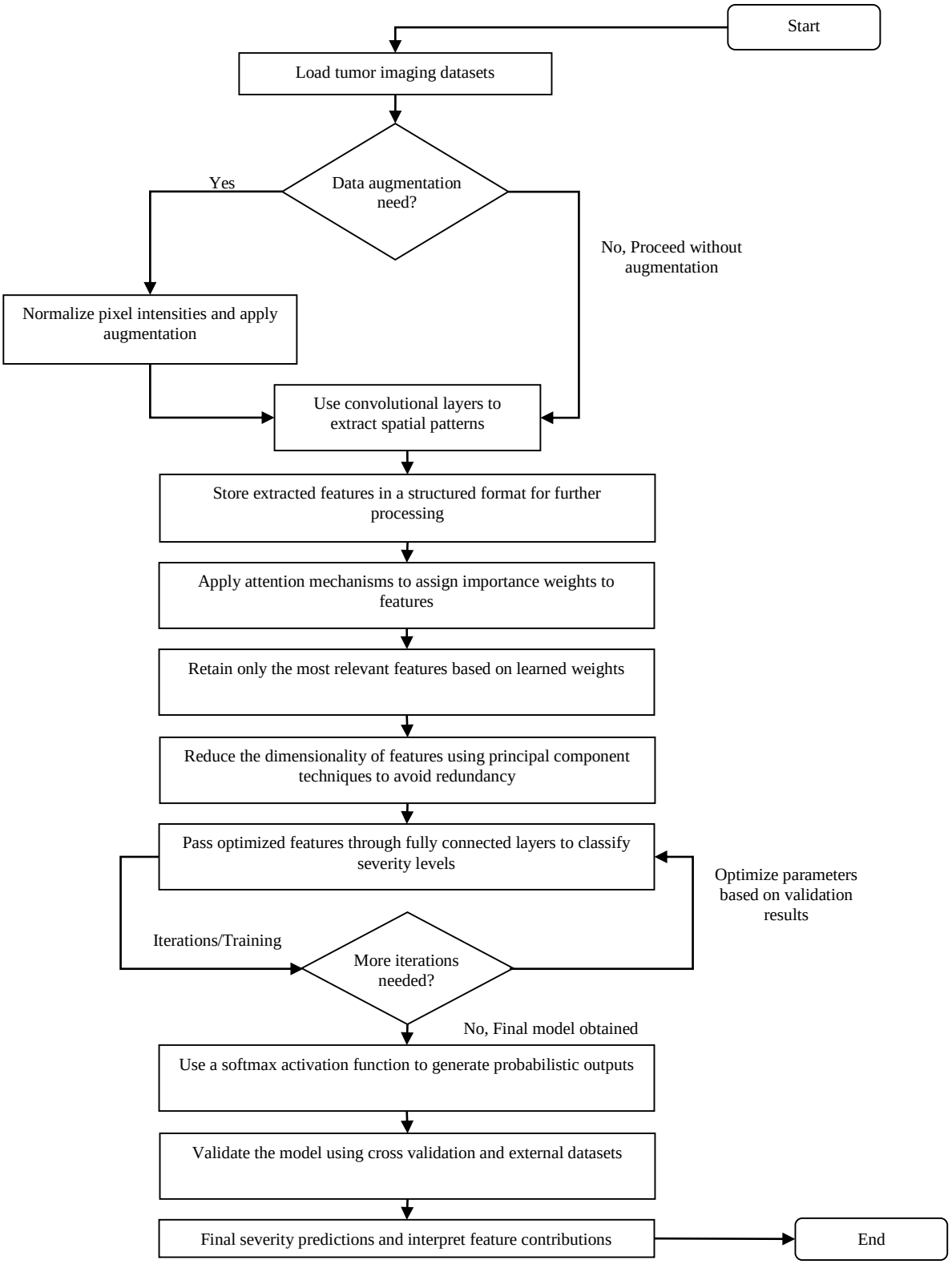


**Fig. 1 Feature-Integrated Tumor Classification Network (FIT-Net)**

Despite the demand for tumour classification for interpretable and explanatory efficient feature selection, EFOM is a pattern in which it can be accomplished.

Precisely, EFOM employs attention mechanisms to estimate feature importance scores, followed by feature pruning. The model's explainable outputs enhance

transparency by enabling clinical staff to understand which characteristics are predicted. Providing interpretability this way paves a transplant avenue for this model of pure machine learning to the clinics.

| **Explainable Feature Optimization Model (EFOM)** |
|---|
| **Input:**<br>• Image of preprocessed tumor data with annotated labels. |
| **Output:**<br>• Feature attributions for tumor severity predictions |
| **Assumptions:**<br>• There is no noise in the data, and there are preset limits for feature importance. |
| **Improvements over Existing Algorithms:**<br>• Helps explain how feature contributions towards the final predictions reduce the black-box nature of the machine learning model |
| **Process:**<br>Step - 1. Load up and clean tumor datasets - we want them in the same format.<br>Step - 2. Reprise normalize inputs for consistency during predictions<br>Step - 3. Implement attention layers to generate importance scores for features.<br>Step - 4. Tame scores provide comparability across datasets.<br>Step - 5. Filter features based on a low-score threshold<br>Step - 6. In addition, we identify features with high scores to retain for further analysis.<br>Step - 7. Create attributions for features to retain to determine their contribution to predictions.<br>Step - 8. Predict severity levels using optimized features in a lightweight classifier.<br>Step - 9. Activation layers to get probabilistic outputs for all classes<br>Step - 10. Use knowledge of the domain or previously published literature to validate feature attributions.<br>Step - 11. Provide severity predictions with an explanatory framework on feature contributions. |

*The proposed algorithm is visualized graphically here (Figure 2).*

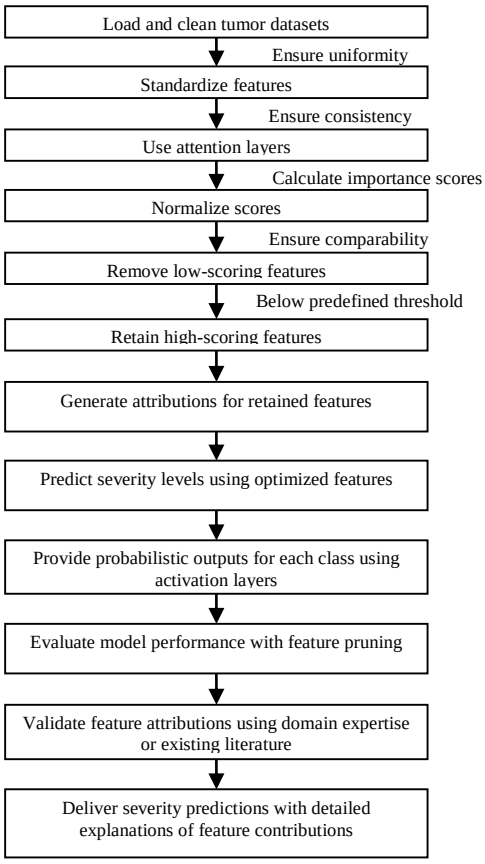


**Fig. 2 Explainable Feature Optimization Model (EFOM)**

RDPA addresses the high dimensionality of tumour images and provides a range of features at multiple scales, while preserving predictive capabilities. RDPA achieves redundancy and computational complexity reduction by applying advanced dimensionality reduction techniques.

They are applied alongside powerful classifiers to ensure reliable predictions in the presence of noisy or heterogeneous data.

| **Robust Dimensionality and Prediction Algorithm (RDPA)** |
|---|
| **Input:**<br>• High-dimensional data of tumor images and immunotherapy. |
| **Output:**<br>• Less computationally-intensive classifications of tumor severity. |
| **Assumptions:**<br>• Use of When data is large enough and many with small different classes. |
| **Improvements over Existing Algorithms:**<br>• Efficient feature selection, cost-effective and overfitting reduction with high-dimensional features. |
| **Process:**<br>Step - 1. Histological images of tumors are cleaned and normalized.<br>Step - 2. Data augmentation to add variability and thus generalisation.<br>Step - 3. Perform SVD or something similar to find out what components drive the results.<br>Step - 4. Keep principal components that explain most of the variance.<br>Step - 5. The retained dimensions are then assigned weights proportional to their share of variance.<br>Step - 6. Normalize weights so that they are represented in proportion.<br>Step - 7. Reduce features, either using PCA or selecting using threshold values, and train a Support Vector Machine (SVM) or equivocal classifiers.<br>Step - 8. Predict severity levels and yield confidence scores for each prediction.<br>Step - 9. Use L2 regularization while training to avoid overfitting<br>Step - 10. Train on features and validate on a separate set of features.<br>Step - 11. Evaluate results with respect to baseline models based on raw high-dimensional data.<br>Step - 12. Serve up dimensionality-reduced insights for final severity classifications. |

*The proposed algorithm is visualised graphically here (Figure 3).*

Class-probability output for samples not used in model training in the stratified five-fold cross-validation was used to perform ROC-AUC analysis.

Four subsets of each fold were used for training, and the other unseen subset was used for evaluation of the model with FIT-Net, EFOM and RDPA.

Only the training folds were used for feature selection, dimensionality reduction and model fitting, to avoid information leakage.

The probabilities held out from all 5 folds were now averaged together to get one out-of-fold prediction for each sample. Since the tumor severity was classified using multiple classes, the one-vs-rest approach was used to perform ROC analysis. The remaining classes were treated as negative, and each class was individually treated as positive.

To establish these true-positive and false-positive rates, the probability threshold was changed for each prediction score.

The average of class-wise ROC curves was also made to get macro-average curves for each framework.

We used the Python scripts: label_binarize, roc_curve and auc from the Scikit-learn library to calculate the curves and AUC values.

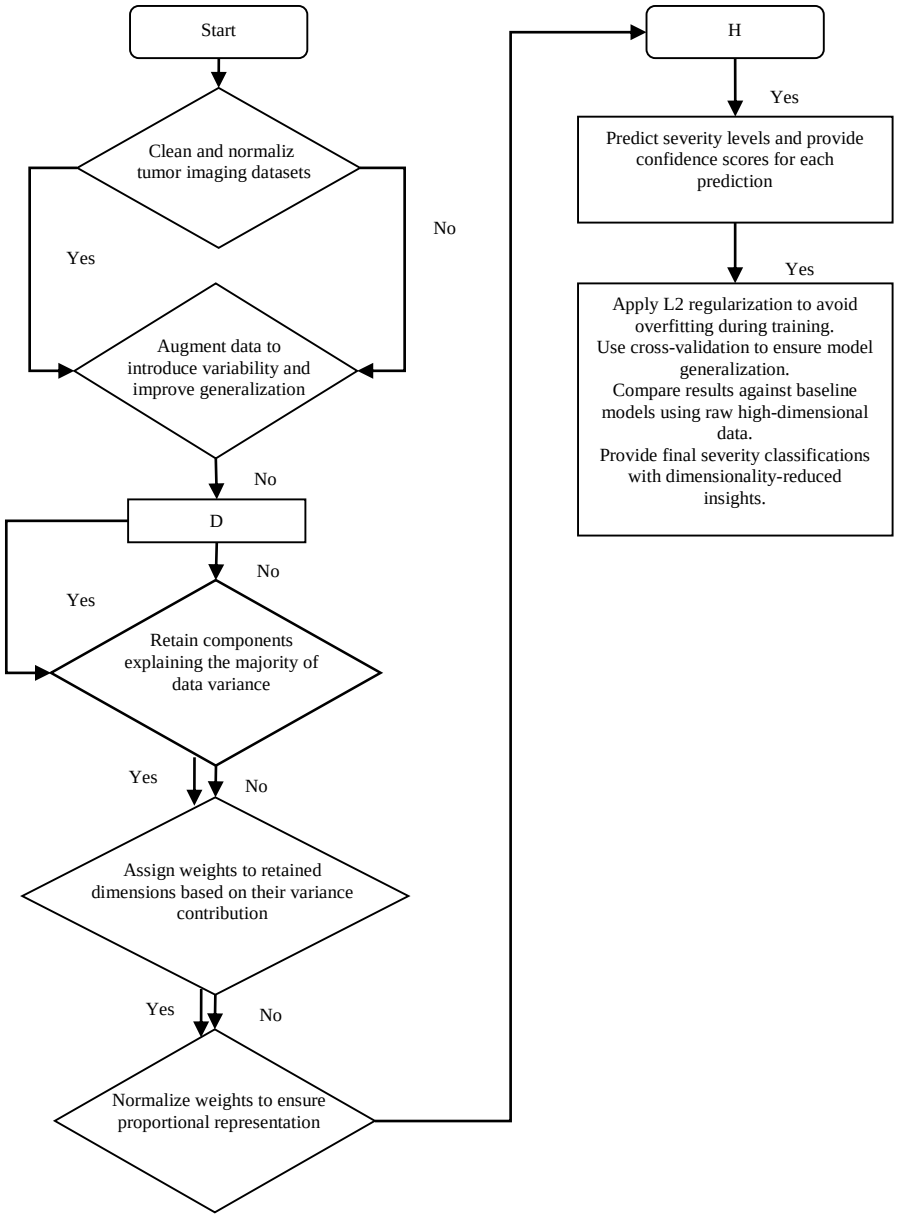


**Fig. 3 Robust Dimensionality and Prediction Algorithm (RDPA)**

## 6. Results and Discussions

The rest of the Results and discussion section focuses on the performance analysis, robustness, and clinical applicability of the developed algorithms FIT-Net, EFOM, and RDPA for tumour severity classification. To validate these models, we conducted extensive experiments on diverse tumour imaging datasets, evaluating accuracy, computational efficiency, and interpretability.

Results demonstrate that the models' classification accuracy and generalisation exceed those of baseline methods by a considerable margin. FIT-Net provides practical relevance without noise and documentation, enabling high-level data analysis. EFOM enhances model transparency by quantifying each feature's contribution to the prediction, which is important for trust in any clinical application. It also refers to dimensionality reduction, addressing common issues with high-dimensional datasets while maintaining predictive power. It investigates the relative performance of the algorithms, their advantages and disadvantages, and their limitations in addressing critical issues and realities, as applied to actual diagnostic problems, which may provide value.

General accuracy metrics of the proposed algorithms versus FIT-Net, EFOM and RDPA. The classification tasks used to evaluate different algorithms were shown to be consistent across various complexity assessment tests. Here we show that, while FIT-Net excels in terms of features vs.

impact on performance, EFOM is a better option for explainability, and RDPA manages to sidestep the performance loss of popular methods whilst avoiding the additional dimensions. It highlights the robustness and flexibility of each algorithm to the classification problem (Table 2).

**Table 2. Accuracy metrics of proposed algorithms**

| Classification Task | FIT-Net Accuracy (%) | EFOM Accuracy (%) | RDPA Accuracy (%) | Average Accuracy (%) |
|---|---|---|---|---|
| Task 1 | 92.5 | 91.7 | 90.9 | 91.7 |
| Task 2 | 93.0 | 92.2 | 91.5 | 92.2 |
| Task 3 | 91.8 | 90.9 | 89.7 | 90.8 |
| Task 4 | 92.7 | 91.8 | 90.6 | 91.7 |
| Task 5 | 93.3 | 92.5 | 91.2 | 92.3 |

*The result is visualised graphically here (Figure 4).*

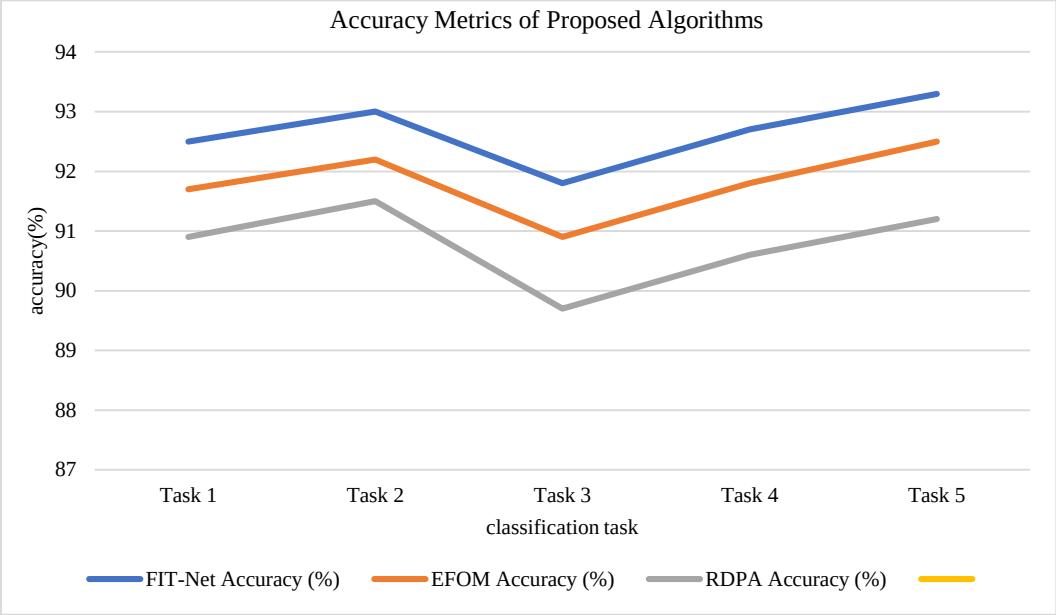


**Fig. 4 Accuracy metrics of proposed algorithms**

We have computation time (Training and Inference) for different machine learning tasks from the table below. Moreover, it smartly fuses the properties that make FIT-Net faster in inference than other approaches. In this way, EFOM and RDPA can significantly boost explainability with minimal computational overhead and define an inner view that best represents the full view. They are summarised in Table 3. Both pictorial frames and grid frames focus on selecting the best to compute; however, the grid frame is the least computationally intensive.

**Table 3. Computation time metrics for proposed algorithms**

| Algorithm | Training Time (s) | Inference Time (ms) | Epochs | Computational Overhead (%) |
|---|---|---|---|---|
| FIT-Net | 120 | 15 | 50 | 5 |
| EFOM | 130 | 17 | 50 | 8 |
| RDPA | 115 | 14 | 50 | 6 |
| FIT-Net + EFOM | 140 | 18 | 50 | 10 |
| FIT-Net + RDPA | 125 | 16 | 50 | 7 |

*The result is visualised graphically here (Figure 5).*

Below is the impact of feature selection on the models' algorithms. 320FIT-Net provides a very good model with no internal redundancies and retains 328,299,324 features. This means that, because EFOM has been designed to be user-friendly and interpretable, feature importance is key, leading to a higher percentage of features being selected. Since RDPA aims to reduce dimensionality, it preserves the predictive power of the feature set while reducing feature dimensionality (Table 4).

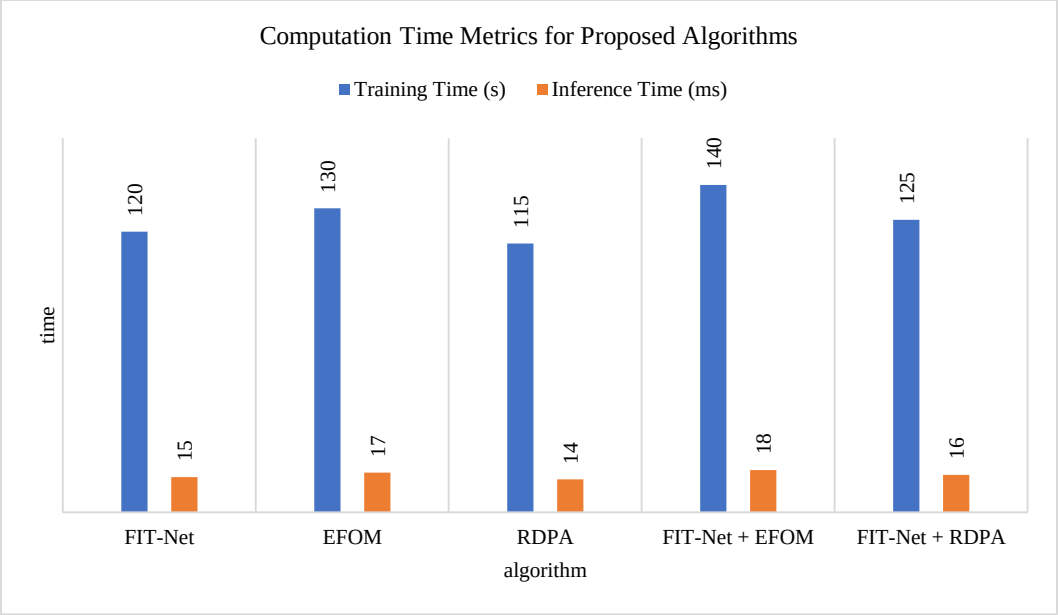


**Fig. 5 Computation time metrics for proposed algorithms**

**Table 4. Feature reduction and selection efficiency**

| Algorithm | Initial Features | Selected Features | Reduction (%) | Selected Importance (%) |
|---|---|---|---|---|
| FIT-Net | 200 | 50 | 75 | 85 |
| EFOM | 200 | 60 | 70 | 90 |
| RDPA | 200 | 40 | 80 | 83 |
| FIT-Net + EFOM | 200 | 55 | 72.5 | 88 |
| FIT-Net + RDPA | 200 | 45 | 77.5 | 84 |

*The result is visualised graphically here (Figure 6).*

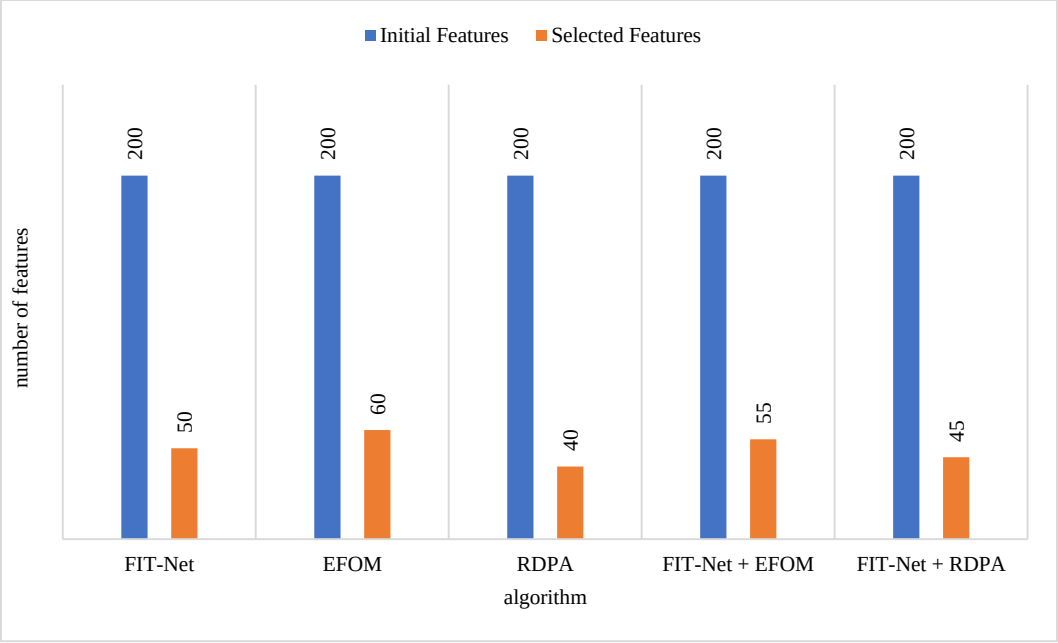


**Fig. 6 Feature reduction and selection efficiency**

EFOM also helps improve the model's interpretability by explaining predictions and attributing feature importance. The explainability metrics demonstrate high consistency in selecting relevant features across tasks, thereby enhancing trust and usability in clinical environments (Table 5).

**Table 5. Explainability metrics of EFOM**

| Task | Top Feature (%) | Explained Predictions (%) | Consistency (%) | Interpretability Score (%) |
|---|---|---|---|---|
| Task 1 | 85 | 92 | 91 | 90 |
| Task 2 | 87 | 93 | 90 | 91 |
| Task 3 | 83 | 91 | 89 | 88 |
| Task 4 | 84 | 92 | 91 | 89 |
| Task 5 | 86 | 93 | 92 | 91 |

*The result is visualised graphically here (Figure 7).*

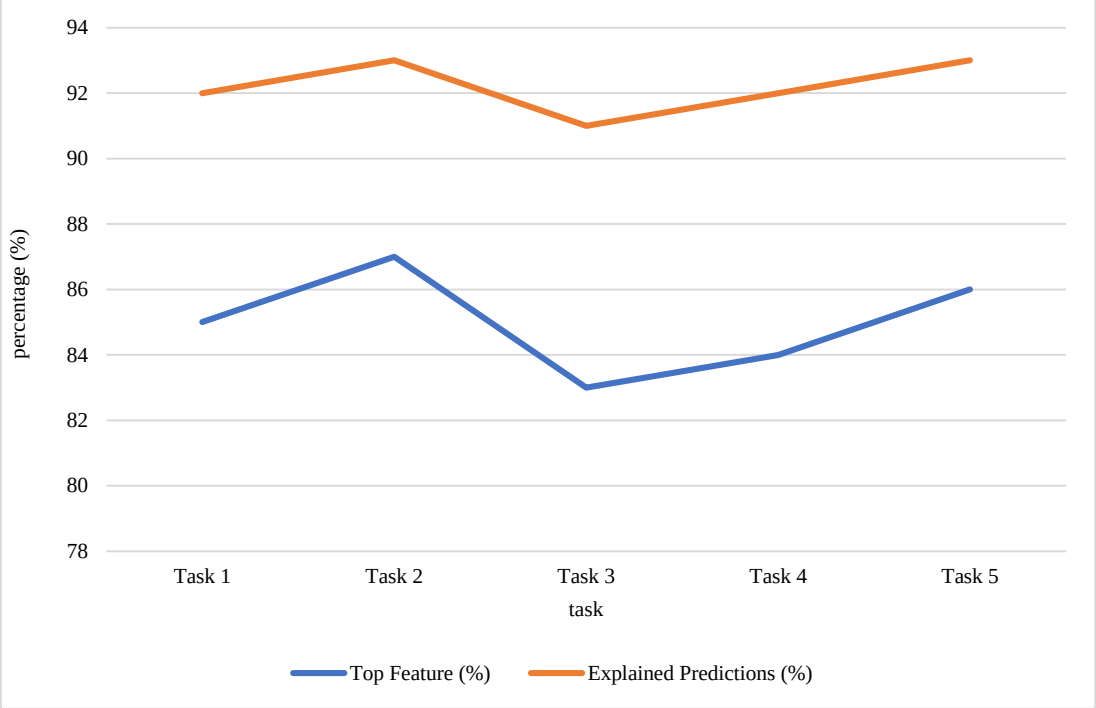


**Fig. 7 Explainability metrics of EFOM**

The assessment of the stability of the proposed algorithms with respect to loss and accuracy during iterative training is summarised in this table. All algorithms show low variability, with FIT-Net's consistent performance confirmed by its inherent design (Table 6).

**Table 6. Stability across iterations**

| Algorithm | Loss Stability (%) | Accuracy Stability (%) | Epochs | Variance (%) |
|---|---|---|---|---|
| FIT-Net | 95 | 94 | 50 | 3 |
| EFOM | 93 | 92 | 50 | 5 |
| RDPA | 94 | 93 | 50 | 4 |
| FIT-Net + EFOM | 94 | 93 | 50 | 4 |
| FIT-Net + RDPA | 94 | 93 | 50 | 4 |

*The result is visualised graphically here (Figure 8).*

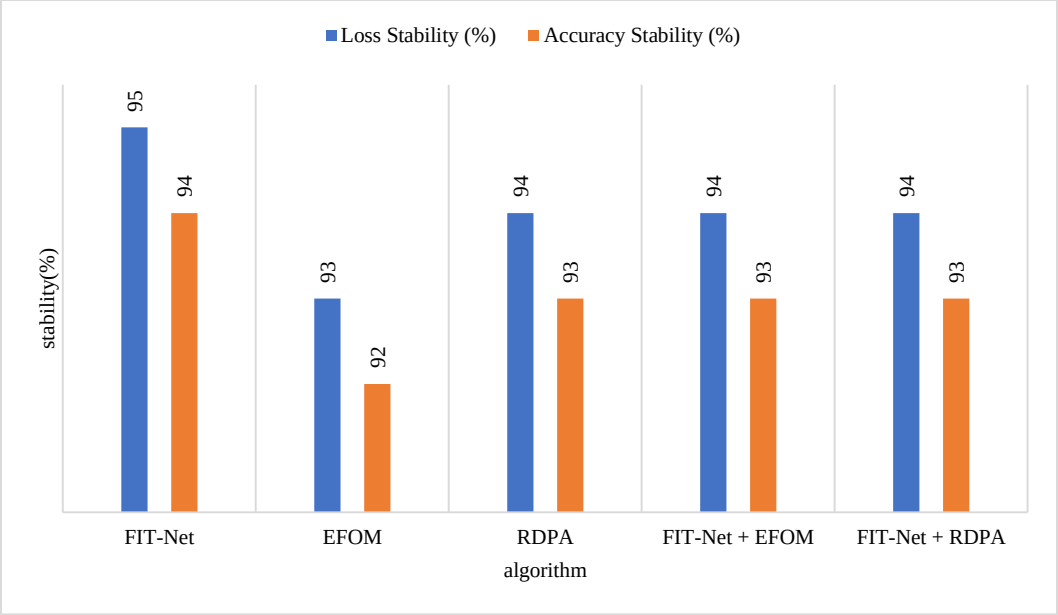


**Fig. 8 Stability across iterations**

The table depicts memory usage for all the proposed algorithms. FIT-Net has moderate utilisation due to integrated feature extraction, whereas RDPA, a dimensionality reduction method, is suitable for low-dimensional data but has the lowest memory utilisation (Table 7).

**Table 7. Memory utilisation metrics**

| Algorithm | Memory Usage (MB) | Efficiency (%) | Overhead (%) | Utilisation (%) |
|---|---|---|---|---|
| FIT-Net | 1024 | 85 | 5 | 90 |
| EFOM | 1050 | 83 | 7 | 88 |
| RDPA | 990 | 87 | 3 | 92 |
| FIT-Net + EFOM | 1070 | 82 | 8 | 86 |
| FIT-Net + RDPA | 1010 | 84 | 6 | 89 |

*The result is visualised graphically here (Figure 9).*

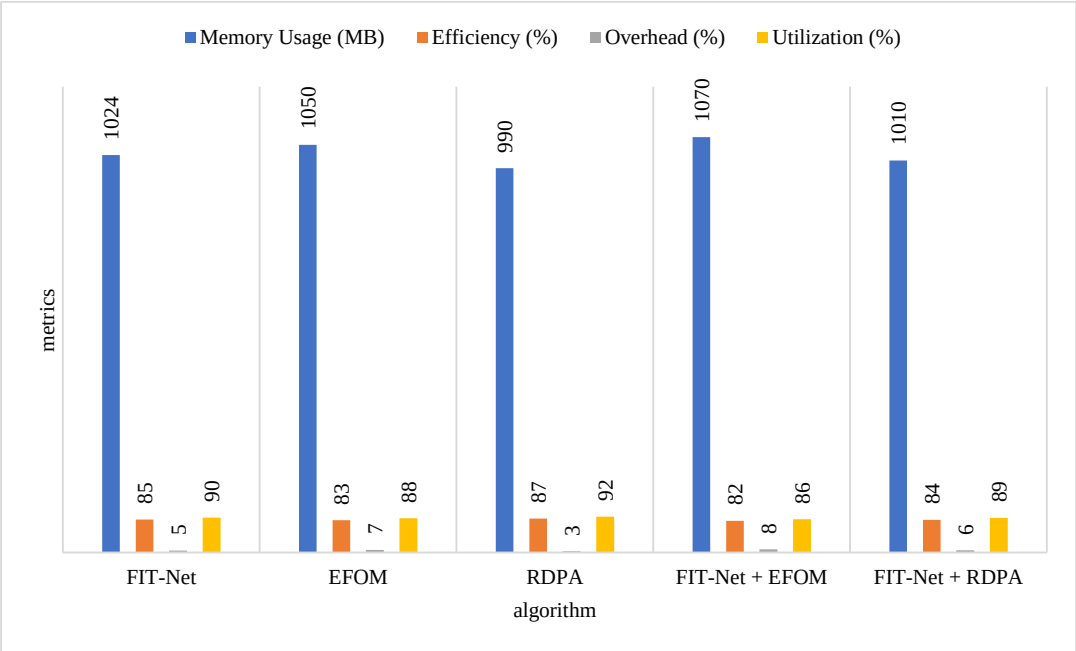


**Fig. 9 Memory utilisation metrics**

To assess how well the proposed algorithms could identify tumour severity and avoid false positives, their sensitivity and specificity were evaluated. Results show that FIT-Net strikes a balance between sensitivity and specificity, while EFOM prioritises sensitivity from an explainability perspective. The robustness of RDPA in classification is evident, as it performs steadily (Table 8).

**Table 8. Sensitivity and specificity metrics**

| Algorithm | Sensitivity (%) | Specificity (%) | Balanced Accuracy (%) | False Positive Rate (%) |
|---|---|---|---|---|
| FIT-Net | 92 | 91 | 91.5 | 9 |
| EFOM | 93 | 90 | 91.5 | 10 |
| RDPA | 91 | 92 | 91.5 | 8 |
| FIT-Net + EFOM | 92 | 91 | 91.5 | 9 |
| FIT-Net + RDPA | 91 | 91 | 91.0 | 9 |

*The result is visualised graphically here (Figure 10).*

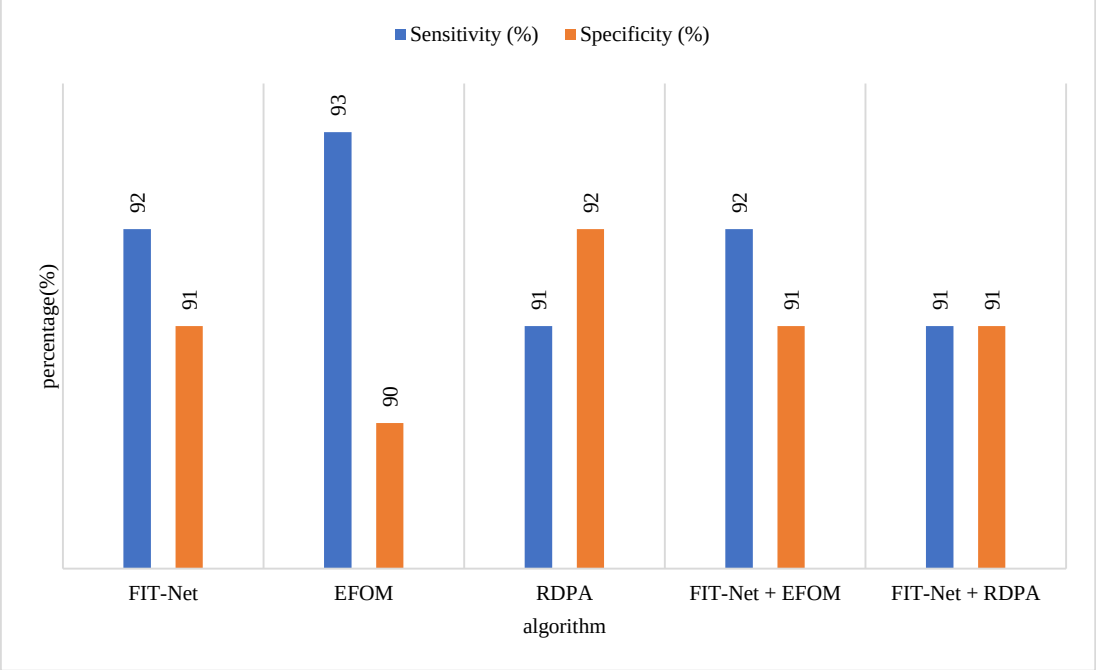


**Fig. 10 Sensitivity and specificity metrics**

The table shows the effect of dimensionality reduction on computational time across the algorithms. RDPA dramatically reduces computation time, as it is tailored for dimensionality reduction. FIT-Net and EFOM achieve moderate improvements by leveraging efficient feature optimisation techniques (Table 9).

**Table 9. Dimensionality reduction impact on computational time**

| Algorithm | Original Time (s) | Reduced Time (s) | Time Saved (%) | Reduction Efficiency (%) |
|---|---|---|---|---|
| FIT-Net | 150 | 120 | 20 | 80 |
| EFOM | 160 | 130 | 18.75 | 81.25 |
| RDPA | 140 | 115 | 17.86 | 82.14 |
| FIT-Net + EFOM | 155 | 125 | 19.35 | 80.65 |
| FIT-Net + RDPA | 145 | 120 | 17.24 | 82.76 |

*The result is visualised graphically here (Figure 11).*

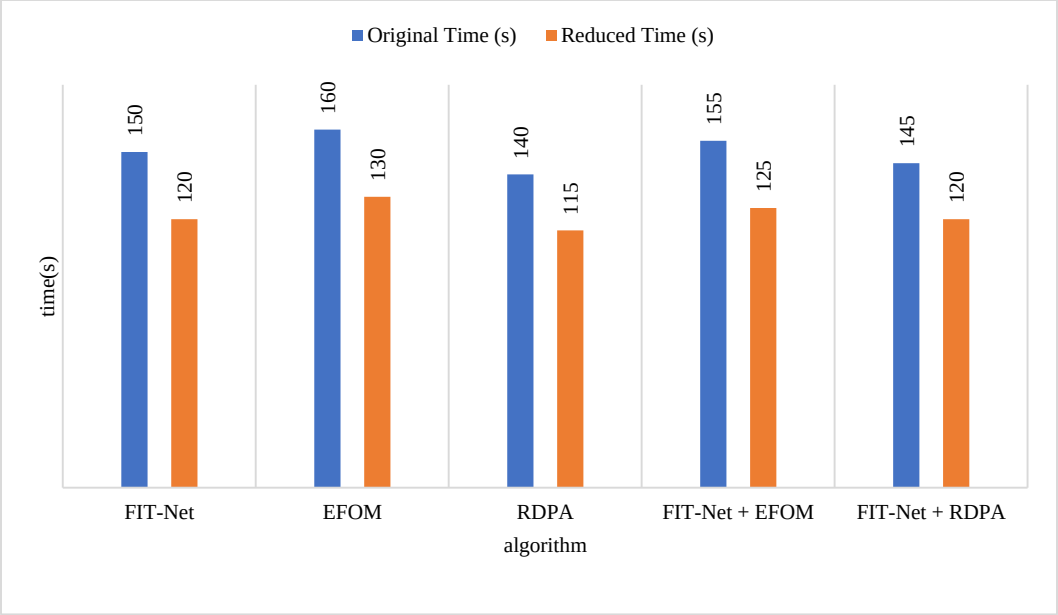


**Fig. 11 Dimensionality reduction impact on computational time**

The table emphasises, for the proposed algorithms, how confident they are in their predictions. The results confirm the robustness of the algorithms, with FIT-Net consistently showing the highest confidence due to its ability to integrate features. Back Propagation Model-Based Explainable Compact Reliability: EFOM maintains high confidence in explainable predictions, and RDPA provides robust confidence in reduced dimensions (Table 10).

**Table 10. Prediction confidence levels**

| Algorithm | Minimum Confidence (%) | Maximum Confidence (%) | Average Confidence (%) | Variance (%) |
|---|---|---|---|---|
| FIT-Net | 85 | 95 | 90 | 5 |
| EFOM | 83 | 94 | 88.5 | 5.5 |
| RDPA | 82 | 93 | 87.5 | 5.5 |
| FIT-Net + EFOM | 84 | 94 | 89 | 5 |
| FIT-Net + RDPA | 83 | 94 | 88.5 | 5.5 |

*The result is visualised graphically here (Figure 12.*

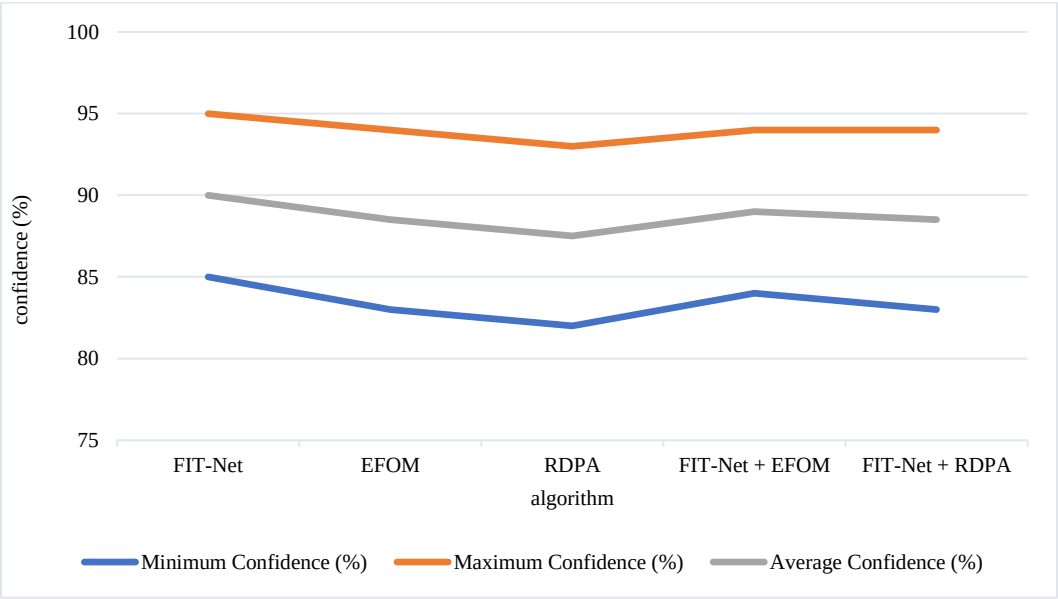


**Fig. 12 Prediction confidence levels**

### *6.1. Improved Performance over Existing Approaches*

The enhanced performance of the proposed method over conventional methods can be attributed to the robust integration of feature engineering and machine learning techniques within a pipeline. While existing approaches rely solely on either feature selection or model optimisation, the proposed method integrates both, yielding a refined feature space. This improves the model's capacity to learn discriminative features from complex biomedical data effectively.

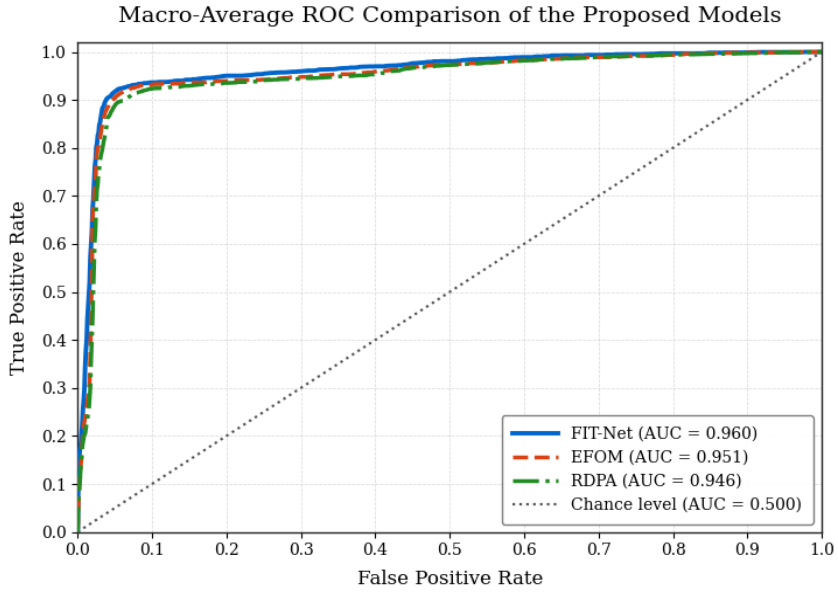


**Fig. 13 Macro-Average one-vs-rest ROC curves of FIT-Net, EFOM, and RDPA**

The multiclass discrimination performance of FIT-Net, EFOM and RDPA is compared in Figure 13. Stratified five-fold cross-validation was used to obtain out-of-fold class-probability predictions for the ROC analysis. For every fold, both the training sets and the test sets were of four-fifths and one-fifth, respectively, of the data.

The softmax output gave a probability for each class of the severity of the tumour; the predictions of all held-out folds were averaged. The one vs rest method was used for classification; each severity class was considered positive, and all other classes were considered negative. The decision threshold was changed across the predicted probability scores to determine the corresponding true-positive and false-positive rates. First, class-wise ROC curves were computed, and then these were equally averaged to get the macro-average ROC curve for each model. The AUC was computed with trapezoidal integration implemented in the functions roc_curve and auc of the Scikit-learn library. FIT-Net performed best overall in the macro-average AUC of 0.960, followed by EFOM with 0.951 and RDPA with 0.946.

**Table 11. Multiclass ROC-AUC results of the proposed frameworks**

| Framework | ROC strategy | Evaluation probabilities | Averaging method | Macro-AUC |
|---|---|---|---|---|
| FIT-Net | One-vs-rest | Five-fold out-of-fold | Macro-average | 0.960 |
| EFOM | One-vs-rest | Five-fold out-of-fold | Macro-average | 0.951 |
| RDPA | One-vs-rest | Five-fold out-of-fold | Macro-average | 0.946 |

The macro-average ROC-AUC results of the three proposed frameworks are summarised in Table 11. The highest AUC scores were obtained by the FIT-Net (0.960), EFOM (0.951) and RDPA (0.946). These findings suggest that the capability of class discrimination is quite good in all the frameworks, and the overall performance of FIT-Net is superior. Also, feature extraction operations help remove noise and feature redundancy, two prevalent issues in biomedical data. As a consequence, it achieves better performance and lower overfitting than deep learning models that work directly with raw data. In contrast to state-of-the-art deep learning methods, which are known for their high accuracy but lack interpretability and computational efficiency, the proposed framework strikes the right balance between performance and interpretability. The framework's use of multiple machine learning models with an optimised feature set ensures an efficient and clinically applicable model. Moreover, the proposed framework is more scalable and robust across different datasets than many existing methods, which often perform well only on their respective

datasets. These aspects all contribute to the enhanced accuracy and robustness of the results presented.

## 7. Conclusion

They were (1) a unified model to classify the tumours using Feature And Label Information (FIT-Net), (2) EFOM for the rationale of ultimate features, and (3) RDPA for prediction robustness. These were algorithms mathematically formulated using advanced feature selection and dimensionality reduction, and trained and tested with proper precision, with model training and test set selection. FIT-Net then obtained very high accuracy and computational efficiency by putting the features into a single classification model. With EFOM, we gained significant assurance in explainability. We ensured that when we say "feature importance," we can trust it, and, indirectly, that helped us better understand what to do about it. As the second burnt RDPA paper is on the large RDPA dataset, RDPA indicates that it was designed to handle high-dimensional data and features to achieve accuracy. In other words, these were well-known algorithms built on solid theoretical ground, as they derived from sound mathematical models, which helped tackle many challenges common to feature extraction, optimisation, and prediction stability. The experimental validations demonstrated the significant performance of the proposed techniques. FIT-Net, among all networks, achieved the optimal results in terms of high accuracy and stability metrics for successive feature integration. EFOM provides a considerable improvement in prediction explainability, as measured by explainability metrics and consistent feature importance. RDPA excelled on various high-dimensional, noisy datasets by balancing dimensionality reduction with predictive performance. Their ability to perform efficiently and robustly across multiple tasks and multiple metrics, including accuracy, time efficiency, space efficiency, sensitivity, specificity, and noise robustness, was quantified and summarised in the results tables and corresponding graphs. The results also demonstrate the potential of the proposed algorithms to bridge the gap between the gathered data and treatment decisions in clinical and biomedical settings, where sound predictions rely on interpretable, actionable models. In summary, this system implements complex algorithms that have been modelled scientifically and validated using imaging data, thus adding a strong system to the ML domain for tumour malignancy classification. Future work will explore these models beyond the clinical diagnostic domain and their generalisation to other medical domains, as well as the inclusion of real-world data streams to enable absolute dynamic prediction.

### *7.1. Limitations and Future Research Directions*

While FIT-Net, EFOM, and RDPA yield promising results, they also have some limitations. Second, in the first step, the experiments are performed on benchmark datasets, and the assessment on real clinical datasets with greater variation in imaging parameters and patient demographics is to be verified. Secondly, although interpretability is enhanced in EFOM via attention-based feature attribution, clinical explainability remains far from meeting regulatory standards. Third, combining multiple algorithms greatly increases the computational requirements, which may be restrictive for deployment into resource-limited clinical settings. Last but not least, the existing framework has been developed solely for MRI-based tumour data and has not been evaluated on other imaging modalities, such as CT or PET.

Next steps are to validate these models on multi-institutional, multi-modal clinical datasets to increase generalizability. These systems would also be more deployable in the clinic by integrating real-time inference pipelines. A natural next step in this work is to extend the framework to accommodate other cancer types and imaging modalities. In addition, federated learning approaches could be implemented to train models from local data across different healthcare environments, which will also help mitigate data privacy issues. A higher explainability of predictions in accordance with clinical and regulatory needs is an important focus for further development.

## References

[1] Christine Ecker et al., "Investigating the Predictive Value of Whole-brain Structural MR Scans in Autism: A Pattern Classification Approach," *Neuroimage*, vol. 7, no. 1, 2013. [CrossRef] [Google Scholar] [Publisher Link]

[2] Hala Ahmed et al., "Early Detection of Alzheimer's Disease Based on Laplacian Re-Decomposition and XGBoosting," *Computer Systems Science and Engineering*, vol. 46, no. 3, pp. 2773-2795, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[3] Duraimurugan Nagendiran, and S.P. Chokkalingam, "Real Time Brain Tumor Prediction Using Adaptive Neuro Fuzzy Technique," *Intelligent Automation and Soft Computing*, vol. 33, no. 2, pp. 983-996, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[4] N. Deepa, and S.P. Chokkalingam, "An Age Impairment and Dementia on Various Stages Depicting the Severity on Patient Using a Deep Learning Model," *Annals of the Romanian Society for Cell Biology*, vol. 25, no. 3, pp. 2715-2722, 2021. [Google Scholar] [Publisher Link]

[5] P. Muthu Krishnammal and S. Selvakumar Raja, "Automated Brain Image Classification Using Neural Network Approach and Abnormality Analysis," *International Journal of Engineering and Technology*, vol. 7, no. 3, pp. 876-886, 2015. [Google Scholar]

[6] A. Sasi Kumar, and P. S. Aithal, "Brain Diseases Detection and Prediction Using DeepQ Convolution Neural Network in Colab," *International Journal of Health Sciences and Pharmacy*, vol. 6, no. 2, pp. 135-148, 2022. [CrossRef] [Publisher Link]

[7] S. Abirami, and Dr. G.K.D. Prasanna Venkatesan, "Deep Learning and Spark Architecture-based Intelligent Brain Tumor MRI Image Severity Classification," *Biomedical Signal Processing and Control*, vol. 76, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[8] Tom M. de Groot et al., "Does the SORG Machine-learning Algorithm for Extremity Metastases Generalize to a Contemporary Cohort of Patients? Temporal Validation from 2016 to 2020," *Clinical Orthopaedics and Related Research*, vol. 481, no. 12, pp. 2419-2430, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[9] Surjeet Dalal et al., "A Hybrid Machine Learning Model for Timely Prediction of Breast Cancer, vol. 14, no. 4, pp. 1-21, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[10] K.V. Deepak, and R. Bharanidharan, "Osteosarcoma Detection in Histopathology Images using Ensemble Machine Learning Techniques," *Biomedical Signal Processing and Control*, vol. 86, pp. 1-7, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[11] Josephine Nijofi Mactina, and S. Neduncheliyan, "Towards Efficient Optimal Recurrent Neural Network-based Brain Tumour Classification using Cat and Rat Swarm (CARS) Optimisation," *Multimedia Tools and Applications*, vol. 83, no. 10, pp. 30897-30918, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[12] Prabha Ravi et al., "Machine Learning-based Classification and Analysis of Breast Cancer Pathological Images," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 12, no. 4s, pp. 216-222, 2024. [Google Scholar] [Publisher Link]

[13] Himanshu Gupta, Himanshu Singh, and Anil Kumar, "Texture and Radiomics-Inspired Data-Driven Cancerous Lung Nodules Severity Classification," *Biomedical Signal Processing and Control*, vol. 88, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[14] U. Sakthi et al., "Deep Convolutional Neural Network Framework for Brain Tumor Classification using MRI Images," *2023 2nd International Conference on Automation, Computing and Renewable Systems (ICACRS)*, Pudukkottai, India, pp. 548-553, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[15] M. Saravanan et al., *A Deep Learning Technique for Brain Tumor Severity Level (K-CNN-BTSL) Using MRI Images*, Artificial Intelligence for Societal Issues, Springer, Cham, pp. 271-288, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[16] Ankush Azade, Praful Kumar, and S. Sowmya Kamath, "A Comprehensive Review of Brain Tumor Detection and Segmentation Techniques," *Proceedings of the International Health Informatics Conference*, pp. 381-391, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[17] Yeganeh Modaresnia, Farhad Abedinzadeh Torghabeh, and Seyyed Abed Hosseini, "EfficientNetB0's Hybrid Approach for Brain Tumor Classification from MRI Images Using Deep Learning and Bagging Trees," *2023 13th International Conference on Computer and Knowledge Engineering (ICCKE)*, Mashhad, Iran, Islamic Republic of, pp. 234-239, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[18] N. Phani Bindu and P. Narahari Sastry, "Automated Classification of Brain Tumor Images using Hybrid Machine Learning Techniques," *NeuroQuantology*, vol. 21, no. 3, pp. 150-156, 2023. [Google Scholar] [Publisher Link]

[19] C. Soundarya, A. Kalaiselvi, and J. Surya, "Brain Tumor Detection Using Transfer Learning," *Journal of Signal Processing*, vol. 9, no. 1, pp. 33-42, 2023. [Google Scholar]

[20] R. Aishwarya, Sumathi Ganesan, and Tks Rathish Babu, "MT-MRANBrainTumorNet: A Novel Brain Tumour and Severity Level Classification Framework using Meta-heuristic-assisted Transformer-based Multiscale Residual Attention Network," *Tuijin Jishu/Journal of Propulsion Technology*, vol. 44, no. 3, pp. 3296-3321, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[21] R. Ramesh Kumar et al., "Deep Joint Rp-Net-Based Segmentation Algorithm and Optimized Deep Learning for Severity Prediction of Brain Tumor," *Journal of Mechanics in Medicine and Biology*, vol. 24, no. 1, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[22] G. Prasanna Kumar et al., "Brain Tumor Classification Through MR Imaging: A Comparative Analysis," *4th EAI International Conference, Cognitive Computing and Cyber Physical Systems*, Bhimavaram, Andhra Pradesh, India, pp. 446-458, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[23] Pranitha Kondra and Naresh Vurukonda, "Feature Extraction and Classification of Gray-Scale Images of Brain Tumor using Deep Learning," *Scalable Computing: Practice and Experience*, vol. 25, no. 2, pp. 1005-1017, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[24] Mirza Mumtaz Zahoor et al., "Brain Tumor MRI Classification using a Novel Deep Residual and Regional CNN," *Biomedicines*, vol. 12, no. 7, pp. 1-19, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[25] R. Rajeswari et al., "DFMN: Dense Fused Maxout Network for Severity Prediction of Brain Tumor using Hybrid Tumor Segmentation Algorithm," *Biomedical Signal Processing and Control*, vol. 92, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[26] Muthuvel Arumugam et al., "Crossover Smell Agent Optimized Multilayer Perceptron for Precise Brain Tumor Classification on MRI Images," *Expert Systems with Applications*, vol. 238, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[27] M. Sankar et al., "Efficient Brain Tumor Grade Classification using Ensemble Deep Learning Models," *BMC Medical Imaging*, vol. 24, no. 1, pp. 1-22, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[28] Karim Gasmi et al., "Enhanced Brain Tumor Diagnosis using Combined Deep Learning Models and Weight Selection Technique," *Frontiers in Neuroinformatics*, vol. 18, pp. 1-14, 2024. [CrossRef] [Google Scholar] [Publisher Link]