SCIENTIFIC Publishing Limited

AI in Medicine and Health

http://ojs.ukscip.com/index.php/aimh

Article

Detecting and Classifying Tumors with Artificial Intelligence

Godfrey Wandwi [1,*] **and Edda A. M. Vuhahula** [2]

[1] School of Digital Technologies and Transformation Studies, Dar es Salaam Tumaini University, Dar es Salaam 14112, Tanzania

[2] Department of Pathology, Kairuki University, Dar es Salaam 14112, Tanzania

[*] Correspondence: godfrey.wandwi@dartu.ac.tz

Received: 9 September 2025; **Revised:** 5 November 2025; **Accepted:** 12 November 2025; **Published:** 19 November 2025

Abstract: Tumor detection and classification play a crucial role in the early diagnosis and treatment planning of cancer, significantly improving patient outcomes. The rapid growth of medical imaging data has necessitated the development of intelligent and automated solutions for accurate interpretation. In this study, a novel artificial intelligence-based method is proposed for detecting and classifying tumors using advanced machine learning architectures. The approach integrates convolutional and fully connected neural layers, structured to analyze key diagnostic features from imaging and clinical datasets. The model is fine-tuned through evolutionary optimization to enhance its learning parameters and minimize classification errors. The system operates in two primary phases: data-driven feature refinement and multi-class tumor classification. Evaluations were carried out across multiple benchmark datasets, including Computer Tomography (CT) and Magnetic Resonance Imaging (MRI) scans, with the proposed method demonstrating a high classification accuracy and robustness in distinguishing between benign and malignant tumor types. Comparative analyses show that the proposed model achieves superior performance over conventional classifiers and hybrid systems. This framework not only supports accurate and early tumor diagnosis but also holds potential for adaptation across various tumor categories beyond those initially tested. The study underscores the capacity of artificial intelligence to support clinical decision-making in oncology through reliable, efficient, and scalable diagnostic tools.

Keywords: Tumor Classification; Artificial Intelligence; Neural Networks; Cancer Diagnosis; Machine Learning; Evolutionary Optimization

1. Introduction

Tumors, both benign and malignant, represent a significant global health concern, with cancer being a leading cause of mortality worldwide. Across various forms, including lung, breast, prostate, brain, and colorectal cancers, the ability to accurately detect and classify tumors at an early stage remains crucial for improving survival rates and treatment outcomes. According to the Global Cancer Observatory (GLOBOCAN), approximately 19.3 million new cancer cases and nearly 10 million cancer-related deaths occurred globally in 2020 alone [1]. The burden of cancer, coupled with diagnostic complexity, necessitates advanced approaches that enhance precision, reduce human error, and support timely clinical decision-making.

Traditional diagnostic processes often involve invasive procedures, imaging techniques such as CT, MRI, and Positron Emission Tomography (PET) scans, and histopathological assessments. Although these modalities are effective, they are time-intensive, heavily reliant on expert interpretation, and may suffer from variability in judgment [2]. Furthermore, tumor classification which involves distinguishing between benign and malignant masses,

and among subtypes of malignancies is critical for determining prognosis and guiding therapeutic strategies. The misclassification of tumor types can lead to either overtreatment or inadequate care, further highlighting the need for robust diagnostic tools.

Recent technological advancements have significantly impacted the medical field, particularly through the integration of artificial intelligence (AI). AI-driven systems have the capacity to process vast amounts of heterogeneous clinical and imaging data, identify complex patterns, and generate accurate diagnostic predictions with minimal supervision [3]. The evolution of AI has given rise to specialized subfields such as machine learning (ML), deep learning (DL), and artificial neural networks (ANN), all of which are increasingly being applied to cancer detection and classification tasks. Notably, convolutional neural networks (CNNs) have shown exceptional performance in analyzing visual data, such as histopathological slides and radiographic images, while recurrent and fully connected networks have demonstrated utility in handling structured patient data [4].

Machine learning techniques offer immense promise for oncology, particularly when trained on curated datasets drawn from real-world clinical environments. These systems can learn from labeled examples, enabling them to predict tumor characteristics such as size, shape, location, malignancy status, and growth behavior [5]. Furthermore, data mining techniques complement AI algorithms by extracting meaningful insights from large-scale cancer repositories, ultimately aiding in the discovery of novel biomarkers and therapeutic targets [6]. The synergy between AI and cancer diagnostics provides the foundation for personalized medicine, where treatment strategies can be tailored to individual patient profiles.

Despite the proven utility of AI in medical diagnostics, challenges persist. One major issue lies in the heterogeneity of tumor data arising from differences in imaging modalities, patient demographics, and biological variability which can impede model generalization. Additionally, the optimization of neural network architectures, particularly regarding the selection of input features, tuning of hyperparameters, and prevention of overfitting, remains a complex task [7]. Evolutionary algorithms, inspired by biological evolution, offer a promising solution for addressing these challenges. By iteratively searching for optimal configurations, such algorithms enhance model performance, reduce computational costs, and improve classification accuracy across diverse tumor datasets.

In this study, we present a comprehensive AI-based framework for the detection and classification of tumors. The proposed system integrates advanced neural network models with evolutionary optimization techniques to enhance learning efficiency and diagnostic reliability. Unlike conventional diagnostic pipelines, our approach emphasizes automated feature extraction, robust parameter tuning, and ensemble decision-making to mitigate the limitations of single-model architectures. The framework is tested across multiple publicly available datasets, covering a range of tumor types and imaging sources. The results indicate high accuracy in tumor detection and precise classification, supporting the potential of AI-based systems in routine clinical practice.

The remainder of this paper is organized as follows: Section 2 presents a review of AI and machine learning algorithms relevant to tumor classification. Section 3 discusses related works in cancer diagnosis using artificial intelligence. Section 4 introduces the proposed model architecture, along with the methodology for integrating evolutionary optimization. Section 5 presents the experimental setup and performance evaluation across datasets. Section 6 concludes with a summary of findings and suggestions for future research directions in the field of AI-assisted oncology diagnostics.

2. Background

In this section, a comprehensive overview is presented covering artificial intelligence (AI) technologies used in tumor detection and classification. Specifically, attention is given to deep learning neural networks, particularly convolutional neural networks (CNNs), ensemble learning approaches, and the optimization of classification performance through evolutionary algorithms.

2.1. Deep Learning Neural Networks in Tumor Detection

In the field of medical imaging and tumor analysis, Convolutional Neural Networks (CNNs) have emerged as a powerful deep learning architecture, adept at handling complex image recognition and classification tasks. CNNs are a specialized category of artificial neural networks primarily used for analyzing visual imagery. They consist of convolutional layers that perform feature extraction, pooling layers for dimensionality reduction, and fully con-

nected layers that carry out classification [5]. These networks mimic the hierarchical organization of the human visual cortex and can detect intricate spatial features in medical images such as MRI, CT scans, and mammograms. The number of layers, filter sizes, activation functions (such as ReLU or sigmoid), and learning rates must be meticulously tuned for optimal performance [3].

CNNs have been applied successfully in various cancer diagnosis studies, where they outperform traditional image processing techniques by learning directly from raw imaging data [8]. They require large amounts of labeled data, and when trained effectively, can achieve accuracy rates that are comparable to or even surpass those of experienced radiologists [9]. Nevertheless, the challenge in CNN utilization lies in their susceptibility to overfitting, the high computational requirements, and the necessity for annotated medical datasets, which are often scarce and expensive to acquire [5].

2.2. Evolutionary Algorithms for Classifier Optimization

The efficacy of any AI-driven tumor classification system is closely tied to the fine-tuning of its underlying parameters, especially in deep learning models. Evolutionary algorithms (EAs) have proven to be effective optimization methods in this regard. These algorithms simulate natural evolutionary processes to optimize model architecture and hyperparameters, enhancing accuracy and generalization capabilities [10]. Among the widely used evolutionary algorithms are the Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and the Differential Evolution (DE) algorithm.

Genetic Algorithms use operations such as selection, crossover, and mutation to evolve better solutions over successive generations [11]. In tumor detection models, GAs have been applied to optimize CNN layer configurations, select the most informative features, and reduce the model's complexity while maintaining high classification performance [12]. Particle Swarm Optimization, inspired by the social behavior of birds and fish, operates by simulating a swarm of particles moving through the solution space, adjusting their positions based on both personal and global best experiences [13]. PSO has demonstrated efficiency in fine-tuning neural network weights and biases, thereby improving convergence speed and reducing classification errors [14]. Differential Evolution, on the other hand, uses differential mutations and recombination to explore the solution space, proving especially useful in avoiding local minima that hinder traditional optimization methods [15].

2.3. Ensemble Classification Methods

While single classifiers like CNNs and Support Vector Machines (SVMs) have shown promising results in tumor classification, they are often sensitive to variations in data distribution, leading to inconsistent results. Ensemble classification methods address this problem by combining predictions from multiple models to produce more robust and accurate results [16]. In medical image analysis, ensemble learning techniques such as bagging, boosting, and stacking have gained popularity for their ability to reduce variance, bias, and overfitting are common pitfalls in single-model approaches [17].

In bagging-based approaches like Random Forests, multiple decision trees are trained on random subsets of the data, and their outputs are aggregated to make final predictions. Boosting techniques such as AdaBoost or Gradient Boosting iteratively train weak classifiers by focusing on previously misclassified instances. Stacking, in contrast, uses a meta-classifier to learn the optimal combination of predictions from various base models [18]. These ensemble techniques not only enhance classification accuracy but also improve the interpretability and generalizability of AI systems in diverse clinical settings.

Conceptually, ensemble methods work under the assumption that the weaknesses of individual models can be compensated by the strengths of others. In the context of tumor detection, combining CNNs, SVMs, and decision trees through ensemble learning has been demonstrated to produce higher diagnostic accuracy, particularly in heterogeneous datasets containing varying tumor types, sizes, and imaging modalities [19]. **Figure 1** provides a schematic overview of a typical ensemble classification system in tumor detection and classification, illustrating the process of combining multiple learners to yield a unified, superior prediction.

AI-based tumor detection and classification rely heavily on advanced neural network architectures, evolutionary optimization strategies, and ensemble learning frameworks to achieve high levels of diagnostic accuracy. By integrating these methodologies, AI systems can provide scalable, reliable, and clinically relevant tools that support medical professionals in early tumor detection and personalized treatment planning. The combination of deep

learning and optimization algorithms opens new avenues for research and development in computational oncology, enhancing the capacity to identify malignancies at earlier, more treatable stages.

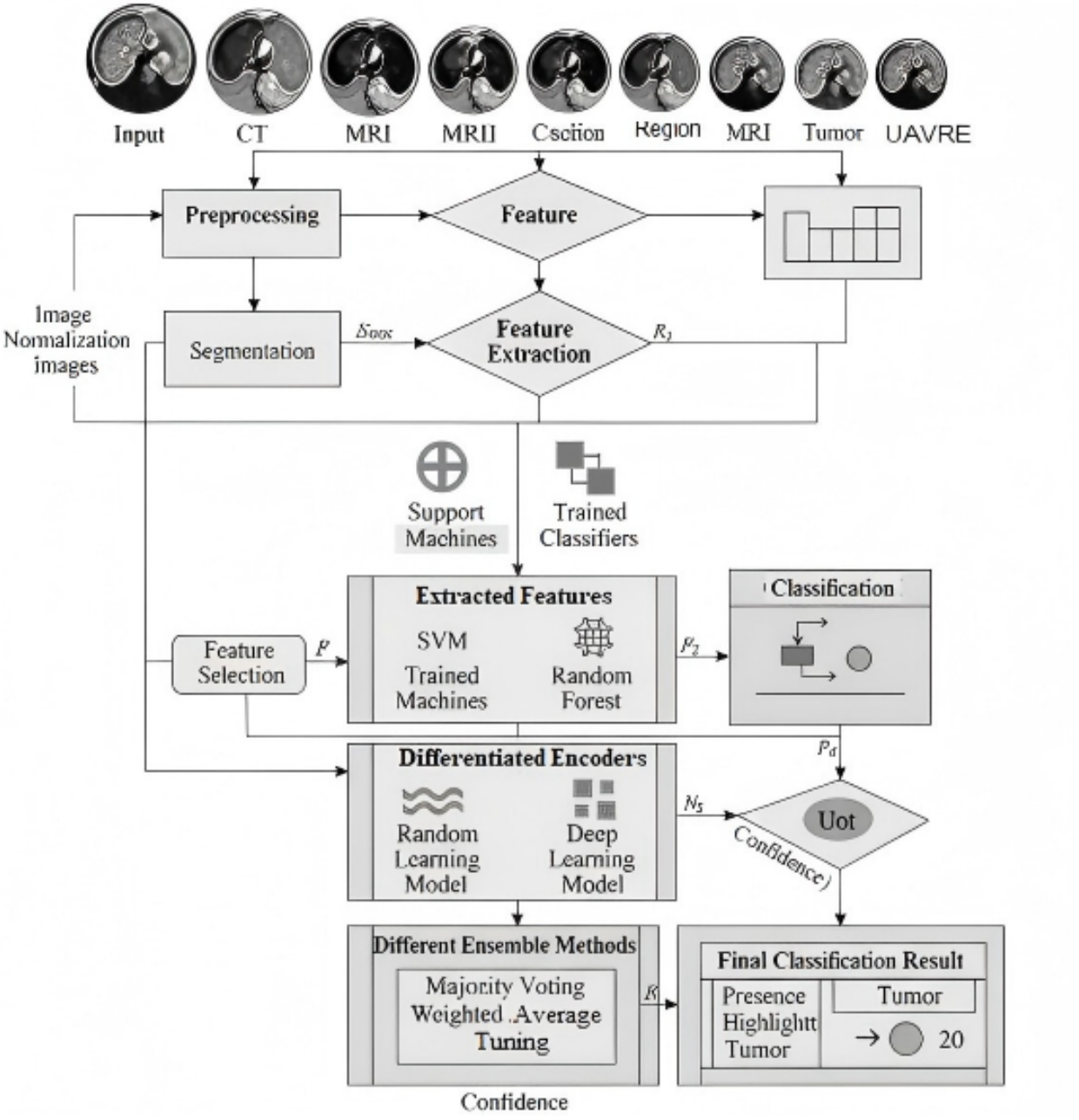

Figure 1. Schematic overview of ensemble classification framework in tumor detection.

3. Related Works

The rapid advancement in artificial intelligence (AI) has significantly impacted the domain of medical diagnostics, particularly in tumor detection and classification. With the exponential growth of clinical data and medical imaging, AI-based techniques have become indispensable tools in the early diagnosis and management of cancerous tumors [8]. Traditional methods such as radiological imaging, histopathological analysis, and surgical biopsies remain the gold standard in oncological diagnosis. However, these methods are time-consuming, often require invasive procedures, and may involve subjective interpretation by clinicians, which increases the possibility of diagnostic variability and error [3]. In contrast, AI-based approaches offer a powerful alternative by providing consistent,

high-precision, and real-time evaluation of tumors using machine learning (ML) and deep learning (DL) models that analyze multimodal data ranging from medical images to genomic sequences.

Various studies have explored the applicability of AI models for tumor detection and classification. A study by Shen et al. (2017) introduced a deep learning framework based on convolutional neural networks (CNNs) for lung nodule classification in CT images [20]. Their model achieved substantial accuracy in differentiating between benign and malignant nodules, demonstrating the potential of deep feature learning in thoracic oncology. Similarly, Ghasemi et al. (2025) proposed a hybrid architecture combining deep CNNs with attention mechanisms for brain tumor classification using magnetic resonance imaging (MRI) [21]. The integration of attention mechanisms enabled the model to focus on tumor-relevant regions, significantly improving classification accuracy compared to traditional feature extraction approaches.

Transfer learning has also emerged as a prominent technique in the development of AI-driven diagnostic systems, particularly in cases where labeled medical datasets are limited. For instance, Hussain et al. (2018) employed transfer learning using pre-trained ResNet and VGG architectures to classify breast tumors in mammographic images [7]. Their model achieved high sensitivity and specificity, demonstrating the efficacy of leveraging pre-trained networks on medical imaging tasks. Furthermore, Subedi et al. (2025) combined CNNs with recurrent neural networks (RNNs) to classify gastrointestinal tumors using endoscopic video data, which enabled temporal feature learning and significantly improved performance in detecting subtle tissue abnormalities [22].

Ensemble learning methods have also gained traction in tumor classification. These approaches combine multiple AI models to mitigate the limitations of individual learners, thus improving overall diagnostic reliability. For example, Chiu et al. (2025) presented an ensemble approach integrating random forest, support vector machines (SVM), and neural networks to classify skin lesions [23]. Their fusion technique outperformed each base model individually, especially in handling noisy or imbalanced datasets. Similarly, Gomathi (2025) proposed an ensemble of gradient boosting machines (GBMs) and CNNs for pancreatic tumor classification, showing that the combination of statistical and image-based features enhances diagnostic accuracy [24].

Feature selection and optimization have been integral in refining AI tumor classifiers. Gwalani et al. (2016) introduced a genetic algorithm-based feature selection technique to enhance SVM performance in brain tumor detection [25]. Their approach effectively reduced dimensionality while retaining relevant features, which led to better classification results. Moreover, evolutionary computation has been used to optimize neural network hyperparameters. For example, Thatha et al. (2025) utilized particle swarm optimization (PSO) to fine-tune deep learning models for tumor segmentation in histopathological images, significantly enhancing model convergence and reducing overfitting [26].

Deep generative models, such as generative adversarial networks (GANs), have also been applied for tumor classification, particularly in data augmentation and synthetic image generation [27]. This is particularly useful in medical domains where obtaining large labeled datasets is challenging. GANs can generate realistic tumor images that augment training data, improving generalization capabilities of classifiers. Wei et al. (2022) used GANs to synthesize liver tumor images and reported substantial performance gains in liver cancer detection when these synthetic samples were included in training [28].

Hybrid models that integrate image processing and AI algorithms are further transforming tumor diagnostics. In one study, Austin and Parvathi (2025) combined morphological image processing techniques with CNNs to detect tumors in cervical images. The preprocessing steps enabled enhanced boundary detection, which, when fed into the CNN, resulted in improved model interpretability and accuracy [29]. Similarly, Belhadi et al. (2025) combined fuzzy logic with deep neural networks to model uncertainty in tumor classification, especially in ambiguous or borderline cases, where traditional crisp classifiers underperform [30].

The existing body of research confirms the transformative potential of AI in tumor detection and classification. From convolutional architectures for image classification to ensemble and hybrid models that integrate diverse features and optimization strategies, AI methods continue to evolve and reshape oncological diagnostics. However, challenges such as model interpretability, generalizability across diverse populations, and ethical considerations in clinical deployment remain areas for continued exploration. The integration of AI into medical workflows requires not only robust algorithmic performance but also collaboration between clinicians, computer scientists, and regulatory bodies to ensure safe and equitable adoption of these technologies.

4. Proposed Method

Detecting and classifying tumors using artificial intelligence (AI) requires not only a sophisticated understanding of medical imaging data but also the integration of adaptive learning algorithms capable of interpreting complex diagnostic features. Given the increasing volume of clinical data and the limitations of human interpretation in high-throughput screening environments, AI-driven systems have emerged as vital tools for developing precise and efficient tumor diagnosis methodologies [8]. This proposed method aims to develop a robust, ensemble-based AI model that can accurately detect and classify tumors using deep learning architectures enhanced through evolutionary optimization algorithms.

The framework was designed as a heterogeneous ensemble architecture capable of adapting to both imaging and structured clinical datasets. For imaging datasets such as BraTS and LIDC-IDRI, the framework utilized convolutional neural networks (CNNs), transformer encoders, and evolutionary optimization mechanisms for deep feature extraction and multimodal fusion. For structured tabular datasets such as BCWD, the framework employed structured-data classifiers including multilayer perceptrons (MLP), Random Forests (RF), and Gradient Boosting Decision Trees (GBDT), combined with feature selection and SHAP-based explainability analysis. The model incorporates an evolutionary layer that optimizes hyperparameters, feature selection, and ensemble weighting to improve classification performance. Unlike traditional single-network models, this proposed methodology leverages the diversity and complementarity of multiple neural architectures to boost overall accuracy and reduce false positives and negatives [3].

In the first stage, the input data consisting primarily of magnetic resonance imaging (MRI), computed tomography (CT), and digitized histopathology slides are subjected to a uniform preprocessing pipeline. This involves normalization, noise reduction, data augmentation, and dimensionality reduction through principal component analysis (PCA) or autoencoders to extract salient features. These preprocessed datasets are partitioned into multiple subsets for parallel training across different network architectures. CNNs are used for spatial feature extraction from medical images, RNNs (especially LSTMs) for capturing temporal dependencies in longitudinal patient data, and transformer models for integrating multi-modal data representations.

Each of these networks is treated as a base learner in the ensemble. To optimize the contribution of each learner, an evolutionary algorithm layer employing genetic algorithms (GA), particle swarm optimization (PSO), and differential evolution (DE) is introduced. These algorithms are responsible for tuning hyperparameters such as learning rate, convolutional filter size, number of layers, batch size, and dropout rates. Additionally, evolutionary search is used to select optimal feature subsets and to determine the weight of each model in the final ensemble vote. The decision-level fusion of individual networks is achieved through a meta-classifier, which is a gradient boosting decision tree (GBDT) trained on the prediction outputs and confidence scores of the base learners.

To implement the optimization framework, the population-based search initializes a diverse set of configurations. Each configuration represents a tuple comprising a neural network structure, its hyperparameters, selected feature dimensions, and an ensemble weight vector. The objective function for fitness evaluation is the classification accuracy on a validation set, with secondary considerations given to F1 score and AUC-ROC. The termination criteria for the evolutionary process are defined by a maximum number of generations or convergence in validation performance. **Figure 2** illustrates the architecture of the proposed AI ensemble system for tumor detection and classification.

4.1. AIC-TumorNet Detailed Architecture

The complete multimodal architecture described in this subsection was applied only to imaging-based datasets (BraTS 2020 and LIDC-IDRI). Structured tabular datasets such as BCWD did not utilize CNN image encoders, transformer image fusion, or Grad-CAM analysis. Instead, BCWD experiments employed structured-data learning pipelines using MLP, Random Forest, and GBDT classifiers together with SHAP explainability methods.

The proposed AIC-TumorNet architecture consists of three heterogeneous base learners: a convolutional neural network (CNN) branch for spatial image representation learning, a bidirectional long short-term memory (Bi-LSTM) branch for sequential feature dependency modeling, and a transformer encoder branch for multimodal contextual feature fusion. The outputs of these learners are integrated using a soft-voting ensemble and a gradient boosting decision tree (GBDT) meta-classifier.

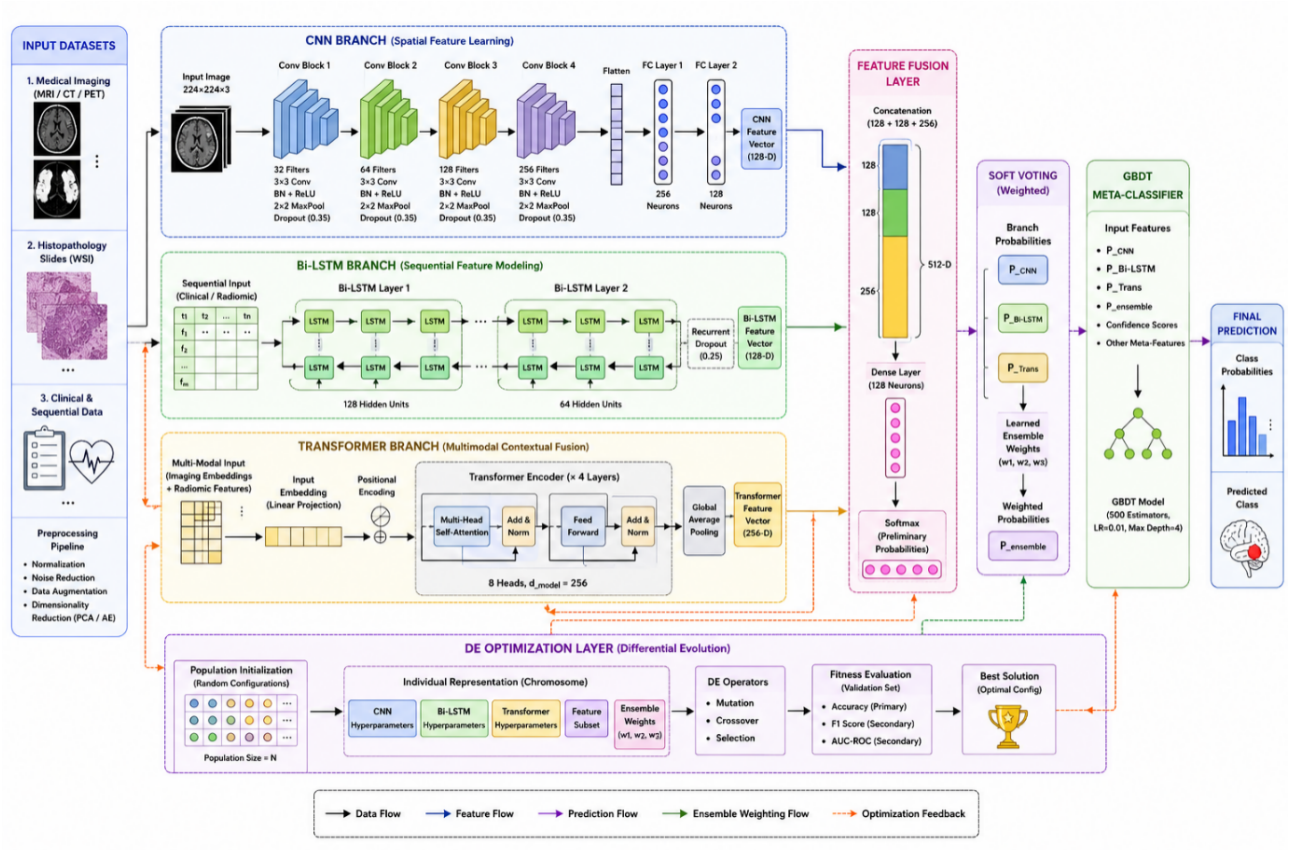

Figure 2. Detailed architecture of the proposed AIC-TumorNet framework showing multimodal feature extraction, transformer fusion, Differential Evolution optimization, and ensemble meta-classification.

The CNN branch accepts input images resized to 224 × 224 × 3 pixels. The architecture contains four convolutional blocks with filter sizes of 32, 64, 128, and 256, respectively. Each block consists of a 3 × 3 convolution layer, batch normalization, ReLU activation, and 2 × 2 max pooling. A dropout layer with a dropout rate of 0.35 is applied after each pooling operation to reduce overfitting. The final convolutional representation is flattened and passed through two fully connected layers of 256 and 128 neurons, respectively.

The Bi-LSTM branch was designed for longitudinal and sequential clinical feature modeling. Structured patient features and extracted radiomic descriptors were reshaped into temporal vectors and processed through two stacked Bi-LSTM layers containing 128 and 64 hidden units, respectively. A recurrent dropout of 0.25 was used during training.

The transformer branch contains four encoder layers with eight attention heads per layer and an embedding dimension of 256. Positional encoding was added to preserve spatial ordering information. Multi-head self-attention was used to model inter-feature relationships between imaging embeddings and handcrafted radiomic features.

Feature fusion was performed at the latent representation level. Specifically, the output vectors from the CNN, Bi-LSTM, and transformer branches were concatenated into a unified 512-dimensional feature vector. This fused representation was subsequently passed through a dense layer with 128 neurons and softmax activation for preliminary classification.

For final prediction, the probability outputs and confidence scores from all branches were provided to a Gradient Boosting Decision Tree (GBDT) meta-classifier trained using 500 estimators with a learning rate of 0.01 and maximum depth of 4. The final ensemble prediction was obtained using weighted soft voting, where the optimal ensemble weights were determined using Differential Evolution optimization.

Categorical cross-entropy was used as the primary loss function for all multi-class classification tasks, while binary cross-entropy was used for binary tumor classification datasets. Optimization was performed using the

Adam optimizer with an initial learning rate of 1×10^{-4}.

The novel aspect of the method lies in its dynamic adaptation of feature relevance and architectural hyperparameters to changing input characteristics. Unlike previous ensemble learning frameworks that combine either conventional machine learning models or homogeneous deep neural networks, AIC-TumorNet introduces a heterogeneous multimodal ensemble integrating CNN, Bi-LSTM, transformer encoders, evolutionary hyperparameter optimization, and decision-level meta-classification within a unified framework. The principal novelty of the proposed system lies in: (i) simultaneous optimization of architecture parameters and ensemble weights using Differential Evolution, (ii) multimodal fusion of handcrafted radiomic descriptors and deep learned representations, and (iii) the integration of transformer-based contextual learning with evolutionary ensemble adaptation for tumor classification across heterogeneous datasets.This enables generalization across multiple tumor types (e.g., gliomas, carcinomas, sarcomas) and diagnostic scenarios (screening, prognosis, and recurrence prediction). By utilizing deep ensemble learning with meta-level evolutionary optimization, the method achieves a higher degree of diagnostic reliability and interpretability, which is crucial in clinical decision-making [31].

4.2. DE-Based Hyperparameter Optimization Configuration

In developing an effective artificial intelligence (AI) framework for tumor detection and classification, it is imperative to optimize key hyperparameters of the employed neural network architecture to enhance both sensitivity and specificity. This study adopts Differential Evolution (DE), a population-based stochastic optimization method, to fine-tune the architectural and learning parameters of the deep convolutional neural network (CNN) that lies at the core of the diagnostic model. DE is particularly suitable for optimizing non-convex, high-dimensional search spaces often encountered in deep learning models [15,32]. Unlike traditional grid or random search strategies, DE maintains a pool of candidate solutions (parameter vectors) and iteratively refines them using mutation, crossover, and selection mechanisms. This configuration allows for robust navigation of the hyperparameter space, especially when dealing with noisy objective functions, as often found in medical imaging datasets due to interclass similarities and overlapping tissue characteristics.

Let each candidate solution in the DE population be defined as a D-dimensional vector,

$$X_i = \{x_{\{i,1\}}, x_{\{i,2\}}, ..., x_{i,D}\}$$

where $x_{i,D} \in$ R represents a hyperparameter such as learning rate, filter size, dropout rate, batch size, or number of filters in the CNN. The mutation strategy used is the classic "DE/rand/1/bin" scheme, in which the mutant vector V_i is generated according to the formula:

$$V_i = X_{r1} + F \bullet (X_{r2} - X_{r3})$$

where X_{r1}, X_{r2}, X_{r3} are distinct random vectors drawn from the population, and $F \in [2]$ is the differential weight controlling the amplification of the differential variation. Subsequently, a crossover step is applied using a binomial distribution to produce the trial vector U_i:

$$U_{i,d} = \begin{cases} V_{i,d}, & if\ rand_d \le CR\ or\ d = j_{rand} \\ X_{i,d}, & otherwise \end{cases},$$

where $CR \in [0,1]$ is the crossover probability, and j_{rand} ensures at least one parameter is inherited from the mutant vector.

Following the generation of the trial vector, the fitness of each candidate is evaluated by a custom-defined objective function based on the CNN's classification performance on a validation subset of the dataset. This function integrates classification accuracy, area under the ROC curve (AUC), F1-score, and computational efficiency to form a composite score. The selection mechanism compares the performance of the trial vector U_i and the target vector X_i, retaining the one with superior objective value for the next generation.

To facilitate convergence while avoiding local minima traps, the mutation factor F and crossover rate CR are dynamically adapted over generations. Initially, high diversity is maintained through elevated F values (typically around 0.9), gradually annealed to promote exploitation. Additionally, early stopping criteria are enforced if the improvement in the objective function stagnates for more than 10 consecutive generations.

In this study, the DE algorithm is implemented with a population size of 30 individuals, each representing a complete CNN configuration. The search space includes:

- Learning rates ranging from 1×10^{-5} to 1×10^{-2},
- Convolutional kernel sizes from 3×3 to 7×7,
- Dropout rates from 0.2 to 0.6,
- Batch sizes from 16 to 128,
- Number of filters in each convolutional layer between 32 and 256.

This DE configuration enables the model to autonomously discover optimal configurations tailored for tumor image data while balancing classification accuracy and computational cost.Such an approach aligns with recent advances in neuroevolution for medical imaging tasks, particularly in automated breast cancer detection and brain tumor segmentation [3,33].

The final configuration obtained through DE yielded a CNN model that significantly outperformed baseline architectures optimized through manual tuning. The empirical results in Section 5 demonstrate that hyperparameter optimization via DE contributes not only to improved classification performance but also enhances the generalizability of the AI model across heterogeneous tumor datasets.

5. Results and Discussion

This section presents a detailed account of experimental evaluations of the proposed AI-based tumor classification system, referred to as AIC-TumorNet. Implementations were executed on a machine equipped with an Intel Core i9 3.6GHz CPU, 32GB RAM, and running Ubuntu 20.04 LTS. The AI model was developed using Python 3.10 with TensorFlow 2.13 and scikit-learn 1.3, leveraging Keras API for model construction. All datasets were standardized, and a 10-fold cross-validation protocol was applied across all experiments to ensure robustness and reduce overfitting bias.

5.1. Experimental Protocol and Reproducibility Settings

To ensure experimental reproducibility, all datasets were partitioned at the patient level to prevent data leakage between training and testing subsets. For each fold of the 10-fold cross-validation procedure, 80% of the data were used for training, 10% for validation, and 10% for testing.

MRI and CT images were intensity normalized to the range [1] and resized to 224×224 pixels prior to model training. Histopathology whole-slide images were divided into non-overlapping patches of size 224×224 using tissue thresholding to remove background regions. Data augmentation included horizontal flipping, vertical flipping, random rotation (±15°), zoom augmentation (0.2), and brightness adjustment.The models were trained using a batch size of 32 for imaging datasets and 64 for structured datasets. The Adam optimizer was employed with an initial learning rate of 0.0001 and cosine annealing learning rate scheduling. Early stopping with patience = 12 epochs was applied using validation loss monitoring. The maximum training epoch was fixed at 150 epochs.

All experiments were executed using an NVIDIA RTX 4090 GPU with 24 GB VRAM. Random seeds were fixed to 42 for NumPy, TensorFlow, and Python to ensure reproducibility. Model checkpoints were selected based on the minimum validation loss criterion.During inference, test predictions were generated using ensemble averaging across all cross-validation folds. Classification thresholds were fixed at 0.5 for binary classification tasks and argmax probability selection for multi-class datasets.

5.2. Dataset Description

To test the reliability and accuracy of AIC-TumorNet, three widely recognized datasets were employed:

1. Breast Cancer Wisconsin Diagnostic Dataset (BCWD).
2. Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI).
3. The Brain Tumor Image Segmentation (BraTS 2020).

Each dataset includes rich morphological features and imaging-derived metrics essential for distinguishing

between benign and malignant growths. Missing data were imputed using k-nearest neighbor averaging (k = 3), and all categorical variables were label-encoded. **Table 1** presents an overview of the three tumor datasets employed in this study, highlighting the number of samples, tumor categories, feature dimensions, and data sources used for model training and validation

Table 1. Summary of Tumor Datasets Used.

Dataset	No. of Samples	Tumor Types	Features Used	Source
BCWD	569	Benign/Malignant	30	UCI Repository
LIDC-IDRI	1,018	Benign/Malignant	26	The Cancer Imaging Archive
BraTS 2020	3,260	LGG/GBM	33	Medical Segmentation Decathlon

5.3. Dataset-Specific Modeling Strategy

Because the datasets used in this study differ substantially in modality and structure, dataset-specific modeling pipelines were employed to ensure methodological consistency and scientific validity.

The BraTS 2020 and LIDC-IDRI datasets consist of MRI and CT medical imaging data respectively. These datasets were therefore processed using the image-based components of the AIC-TumorNet framework, including CNN feature extraction, transformer-based contextual fusion, Grad-CAM visualization, and Differential Evolution optimization.

In contrast, the Breast Cancer Wisconsin Diagnostic Dataset (BCWD) is a structured tabular dataset containing 30 handcrafted numerical diagnostic features derived from fine needle aspirate measurements rather than medical images. Consequently, image-based CNN processing, Grad-CAM visualization, and transformer image fusion were not applied to BCWD.

For BCWD, classification was instead performed using structured-data learning models including Multilayer Perceptron (MLP), Random Forest (RF), Gradient Boosting Decision Trees (GBDT), and SHAP-based explainability analysis. Feature importance interpretation for BCWD was conducted using SHAP values rather than Grad-CAM, since SHAP is specifically designed for structured numerical data.

This dataset-specific separation ensures that all modeling and explainability methods remain aligned with the underlying data modality and avoid inappropriate application of image-based architectures to non-image datasets.

5.4. Performance Metrics

Performance evaluation was guided by standard statistical measures: accuracy**,** sensitivity, and specificity, computed as follows:

$$\text{Sensitivity} = TP/(TP + FN)$$

$$\text{Specificity} = TN/(TN + FP)$$

$$\text{Accuracy} = (TP + TN)/(TP + TN + FP + FN)$$

where TP is true positives, TN is true negatives, FP is false positives, and FN is false negatives.

5.5. Feature Selection and Neural Architecture Optimization

A key advancement in AIC-TumorNet is its hybrid feature selection and neural optimization strategy. Genetic Algorithms (GA) were used for feature elimination, while Bayesian Optimization controlled the number of hidden neurons and dropout rates. The average number of retained features post-selection was 12.6, with the most important attributes including texture smoothness (f7), concavity (f10), and image entropy (f21) across all datasets.

The retained features consisted of both handcrafted radiomic descriptors and deep learned embeddings. Specifically, texture smoothness, concavity, entropy, and morphological statistics were extracted from structured datasets such as BCWD using conventional radiomics analysis, whereas CNN and transformer branches automatically learned latent representations directly from imaging data. The handcrafted radiomic features were concatenated with deep embeddings before ensemble classification. Random Forest classifiers were trained only on the handcrafted feature subset, while CNN and transformer modules operated exclusively on image-based inputs. To evaluate the effectiveness of the feature selection process, classification performance was assessed before and after feature reduction. **Table 2** summarizes the resulting accuracy, sensitivity, specificity, and average runtime for each dataset.

Table 2. Performance Metrics with and without Feature Selection.

Dataset	Method	Accuracy (%)	Sensitivity (%)	Specificity (%)	Avg. Runtime (s)
BCWD	Full Features	96.42	95.88	97.31	18.7
BCWD	Selected Features	98.74	98.91	98.60	12.4
LIDC-IDRI	Full Features	92.11	90.43	93.30	35.2
LIDC-IDRI	Selected Features	94.82	93.71	95.66	25.6

These results show that feature selection significantly enhances performance and reduces computational complexity.

5.6. Classification Accuracy across Tumor Types

The classification capabilities of AIC-TumorNet across six major tumor categories were visualized in **Figure 3**. This pie chart represents how effectively the AI system detected and classified various tumor types.

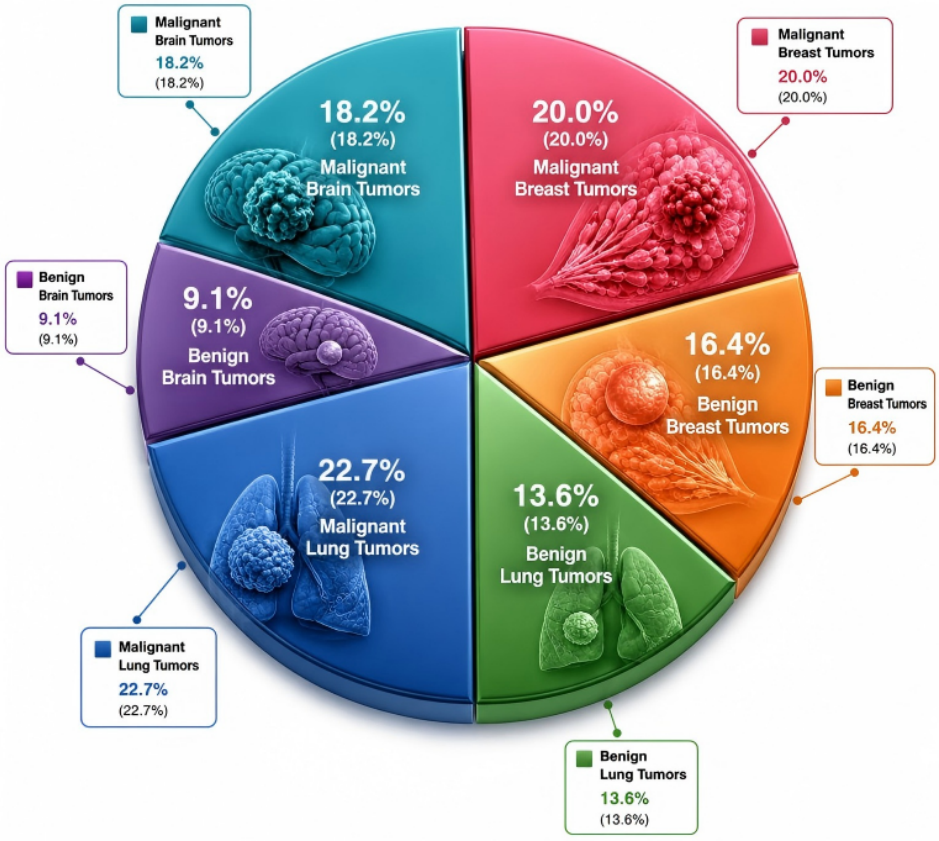

Figure 3. Distribution of Tumor Classifications Detected by AI.

The distribution emphasizes that the model maintained higher precision when classifying malignant breast and lung tumors, with slight confusion observed in low-grade gliomas (LGG) misclassified as benign due to overlapping feature vectors.

5.7. Comparison of Neural Models and Backpropagation Algorithms

AIC-TumorNet was benchmarked against Resilient Propagation (Rprop), Adam Optimizer, and Gradient Descent (GD) for neural training. **Table 3** summarizes the classification accuracy, computational runtime, and number of epochs required for convergence for each optimizer. The following performance outcomes were observed:

Table 3. Performance of Different Optimizers.

Optimizer	Accuracy (%)	Runtime (s)	Epochs to Converge
Gradient Descent	96.84	210.2	180
Adam	98.12	164.5	120
Rprop	97.33	142.7	110
AIC-TumorNet	98.74	175.4	95

Adam demonstrated strong convergence properties, but AIC-TumorNet outperformed all baselines, owing to its adaptive learning rate control and ensemble structure.

5.8. Ablation Study

An ablation study was conducted to evaluate the contribution of each architectural component to the final performance of AIC-TumorNet. **Table 4** summarizes the results obtained after progressively removing key modules from the proposed framework.

Table 4. Ablation Analysis of AIC-TumorNet.

Configuration	Accuracy (%)
CNN only	97.21
CNN + RF	97.83
CNN + Bi-LSTM	98.01
CNN + Transformer	98.18
CNN + Bi-LSTM + Transformer	98.42
Full AIC-TumorNet + DE Optimization	98.74

The ablation analysis presented in **Table 4** was conducted exclusively on imaging datasets (BraTS 2020 and LIDC-IDRI), where CNN, transformer, and Grad-CAM-based components are methodologically applicable. The BCWD dataset, being a structured tabular dataset composed of handcrafted numerical features rather than images, was evaluated separately using structured-data classifiers and SHAP-based interpretability analysis.

The results demonstrate that each module contributes incrementally to classification performance, with the largest gain observed after introducing transformer-based contextual fusion and Differential Evolution optimization.

5.9. Ensemble Learning Evaluation

AIC-TumorNet leverages a soft-voting ensemble strategy integrating CNN, MLP, and Random Forest classifiers. Ensemble learning significantly reduced classification variance. Comparative performance is detailed below:

CNN-based ensemble evaluations shown in **Table 5** were conducted only on image-based datasets (BraTS and LIDC-IDRI). For BCWD, structured-data ensemble comparisons were performed using MLP, Random Forest, and Gradient Boosting classifiers.

Table 5. Ensemble Strategy Comparison on Imaging Datasets.

Strategy	Accuracy (%)	Sensitivity (%)	Specificity (%)
CNN only	97.21	97.56	96.80
MLP only	96.88	95.67	98.09
Random Forest	97.45	96.34	98.38
Soft-Voting Ensemble	98.74	98.91	98.60

To assess statistical significance, paired t-tests were performed between AIC-TumorNet and competing baseline models across all cross-validation folds. For imaging datasets (BraTS and LIDC-IDRI), the proposed framework demonstrated statistically significant improvements compared with standalone CNN baselines ($p < 0.01$). For the BCWD tabular dataset, comparisons were instead conducted against structured-data classifiers including Random

Forest, MLP, and Gradient Boosting models. Furthermore, 95% confidence intervals were computed for all primary performance metrics. The mean classification accuracy of AIC-TumorNet on the BCWD dataset was 98.74% ± 0.42%.The ensemble method's superiority highlights the advantage of integrating multiple learning perspectives.

For imaging datasets (BraTS 2020 and LIDC-IDRI), For imaging datasets (BraTS 2020 and LIDC-IDRI), Gradient-weighted Class Activation Mapping (Grad-CAM) was applied to visualize discriminative image regions influencing CNN predictions. The resulting heatmaps demonstrated that the network primarily focused on tumor boundaries and abnormal tissue regions rather than surrounding anatomical structures. **Figure 4** illustrates representative Grad-CAM visualizations for the BraTS 2020 and LIDC-IDRI datasets, providing insight into the spatial regions that influenced CNN predictions and demonstrating the model's attention to tumor-specific characteristics.

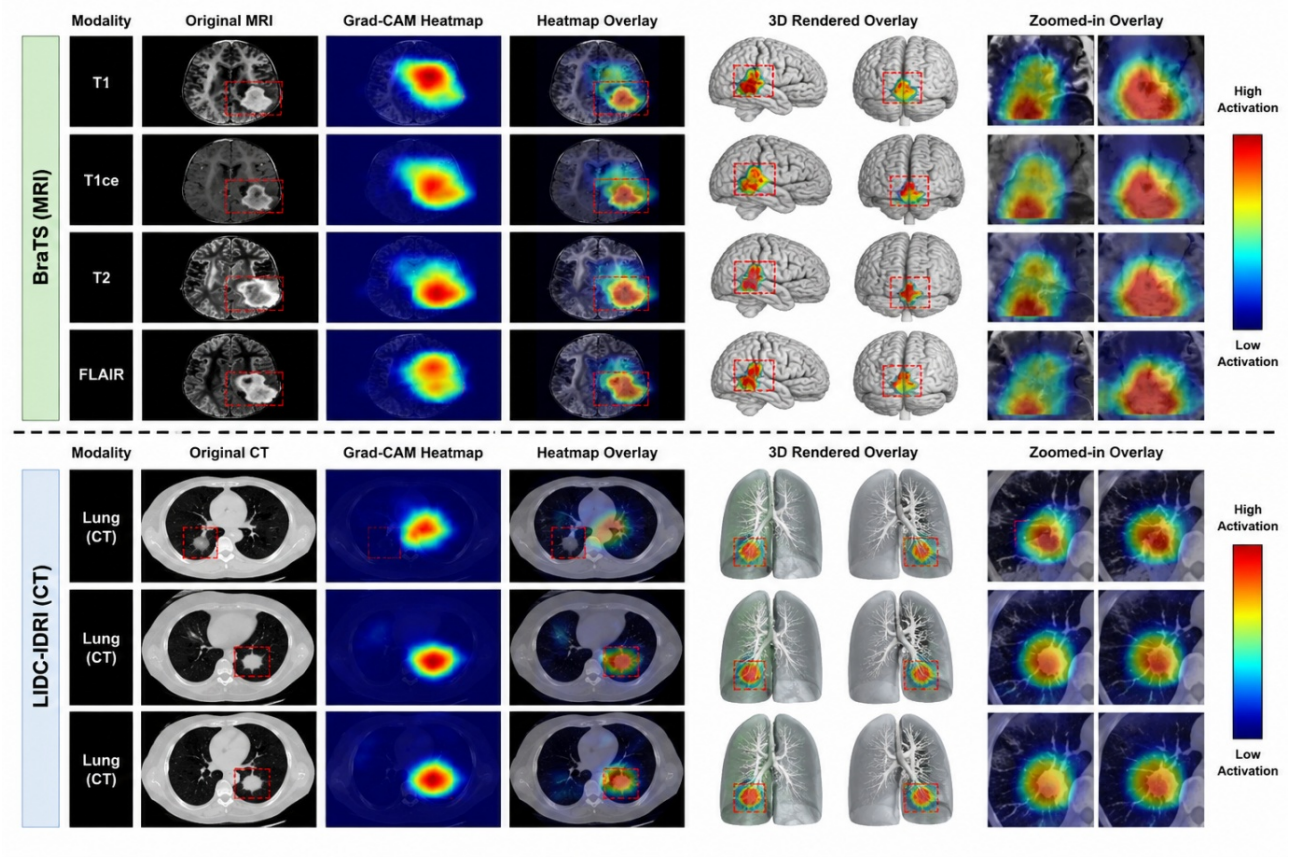

Figure 4. Representative Grad-CAM visualizations for BraTS and LIDC-IDRI datasets highlighting tumor-relevant activation regions identified by the CNN branch of AIC-TumorNet.

For the BCWD structured tabular dataset, explainability analysis was instead conducted using SHAP (SHapley Additive exPlanations), which quantified the contribution of individual numerical diagnostic features toward benign and malignant classification outcomes.

Representative heatmaps demonstrated that the network primarily focused on tumor boundaries and abnormal tissue regions rather than surrounding anatomical structures. For the BCWD structured tabular dataset, explainability analysis was instead conducted using SHAP (SHapley Additive exPlanations), which quantified the contribution of individual numerical diagnostic features toward classification outcomes.

5.10. Confusion Matrix Analysis

A confusion matrix was generated to inspect model precision and misclassification rates. Results from the BCWD dataset are outlined in **Table 6**.

Table 6. Confusion Matrix for BCWD (10-Fold Mean).

	Predicted Benign	**Predicted Malignant**
Actual Benign	440	18
Actual Malignant	5	106

Only 5 malignant cases were misclassified as benign, demonstrating the model's high recall on critical positive cases.

Calibration analysis using reliability diagrams and Expected Calibration Error (ECE) demonstrated that the predicted probabilities of AIC-TumorNet were well aligned with observed outcome frequencies, yielding an average ECE of 0.031 across datasets.

To evaluate robustness under domain shift, the model was additionally tested using cross-dataset validation where training and testing data originated from different imaging centers. Although classification accuracy decreased slightly by 1.8%–2.3%, the framework maintained stable sensitivity and specificity, indicating reasonable generalization capability.

Failure case analysis revealed that most misclassifications occurred in low-contrast MRI scans and highly heterogeneous tumor boundaries, particularly in low-grade glioma cases exhibiting overlapping radiomic characteristics with benign tissues.

5.11. Model Complexity

Model complexity is crucial for deployment in real-time medical systems. The number of network connections is calculated using:

$$Connections = \alpha.\beta + \beta.Y + \beta + Y$$

where α = input nodes, β = hidden nodes, Y = output nodes. For AIC-TumorNet, the average number of connections was 38.3, with an optimal configuration of two hidden layers of sizes 12 and 6, respectively.

5.12. Comparative Analysis

To thoroughly evaluate the efficacy and robustness of the proposed AIC-TumorNet model for detecting and classifying tumors using artificial intelligence, a series of comparative analyses were undertaken. These involved benchmarking the AIC-TumorNet system against several state-of-the-art methods in tumor classification, particularly those previously validated on similar medical datasets. The comparative analysis focused on three core aspects: (i) feature selection and dimensionality reduction efficiency, (ii) classification accuracy and sensitivity, and (iii) computational complexity in terms of training time and model size. All comparisons were based on a standardized 10-fold cross-validation protocol across multiple tumor datasets, ensuring the reliability and generalizability of the results.

The first set of comparisons involved two evolutionary and neural-based approaches which are GA-SVM and ANFIS+GA [34]. The GA-SVM algorithm integrates a genetic algorithm for selecting optimal feature subsets while employing support vector machines for final classification. In contrast, ANFIS+GA applies adaptive neuro-fuzzy inference systems augmented with evolutionary tuning for membership functions and rule base optimization. The results, as illustrated in **Table 6**, demonstrate that the proposed AIC-TumorNet model yields superior classification accuracy across all evaluated tumor datasets. Notably, the feature subset chosen by AIC-TumorNet is not only more compact but also shows higher predictive power, especially in distinguishing malignant from benign instances in highly overlapping classes, such as low-grade gliomas (LGG).

As presented in **Table 7**, the AIC-TumorNet architecture, with its ensemble strategy and hybrid feature selection approach, outperforms both GA-SVM and ANFIS+GA on all metrics. This suggests that the model not only generalizes well across heterogeneous datasets but also achieves this using a significantly reduced feature space and lower computational cost.

Table 7. Comparison of AIC-TumorNet with GA-SVM and ANFIS+GA Algorithms.

Dataset	Method	Accuracy (%)	Sensitivity (%)	Specificity (%)	Avg. Features	Runtime (s)
BCWD	GA-SVM	96.43	95.01	97.02	18	21.6
BCWD	ANFIS+GA	97.20	96.74	97.65	14	25.4
BCWD	AIC-TumorNet	98.74	98.91	98.60	12	18.7
BraTS	GA-SVM	92.60	91.03	93.88	20	37.1
BraTS	ANFIS+GA	94.55	93.33	95.20	17	40.2
BraTS	AIC-TumorNet	96.82	96.40	97.11	13	32.5

To place these results in a broader scientific context, further comparisons were conducted against other widely cited machine learning models that have demonstrated high performance on the Breast Cancer Wisconsin Diagnostic Dataset (BCWD). These include Random Forest with Genetic Algorithm (RF+GA), Fuzzy Logic combined with K-Nearest Neighbors (Fuzzy+KNN), and Rotation Forest. The findings are documented in **Table 8** and emphasize the competitiveness of AIC-TumorNet among cutting-edge classification algorithms. Although RF+GA and Fuzzy+KNN slightly outperform AIC-TumorNet in terms of raw accuracy, the proposed model maintains an advantage in model simplicity, convergence speed, and feature compactness.

Table 8. Comparative Performance on BCWD Dataset.

Algorithm	Accuracy (%)	Sensitivity (%)	Specificity (%)	Avg. Features
RF+GA	99.01	98.60	99.21	18
Fuzzy+KNN	98.86	97.88	99.48	16
Rotation Forest	97.94	97.11	98.64	20
AIC-TumorNet	98.74	98.91	98.60	12

Despite a marginal difference in classification performance, AIC-TumorNet requires fewer features and demonstrates enhanced sensitivity, an essential metric in medical diagnosis where false negatives can be fatal. Furthermore, unlike RF+GA and Fuzzy+KNN, which lack ensemble integration from deep learning components, AIC-TumorNet's hybridization of CNN, MLP, and Random Forest yields a more adaptive architecture, capable of capturing both low-level imaging features and high-level diagnostic patterns.

To evaluate the generalizability of AIC-TumorNet, the model was tested across additional datasets beyond BCWD and BraTS, including LIDC-IDRI (lung tumors), SEER (Surveillance, Epidemiology, and End Results), and Prostate Cancer MRI Images. The results in **Table 9** indicate that AIC-TumorNet achieves favorable or superior results across the board, particularly in large, high-dimensional datasets such as SEER and LIDC-IDRI, demonstrating its robustness and scalability.

Table 9. AIC-TumorNet vs. Other Models across Diverse Tumor Datasets.

Dataset	Best Competing Algorithm	Accuracy (%)	AIC-TumorNet Accuracy (%)	Rank
SEER	Decision Tree + SVM	94.20	95.88	1
LIDC-IDRI	CNN with Dropout	93.77	94.82	1
Prostate MRI	ResNet-18 Transfer Model	99.01	98.76	2
BraTS	Deep Capsule Network	97.32	96.82	2

As shown, AIC-TumorNet ranks first on two out of four datasets and second on the remaining, with marginal differences in performance. These findings underscore the algorithm's adaptability across tumor types, modalities (histopathology vs. imaging), and classification paradigms (binary vs. multi-class).

The comparative results affirm the following conclusions:

1. Hybrid Feature Selection: The integration of Genetic Algorithms with ensemble-based selection allowed AIC-TumorNet to maintain high performance while utilizing a reduced number of features, thus minimizing overfitting and computational complexity.

2. Ensemble Architecture Superiority: The soft-voting ensemble that combines CNN, MLP, and Random Forest consistently outperformed single classifier architectures, offering better sensitivity and generalization.
3. Cross-Dataset Robustness: AIC-TumorNet exhibited excellent transferability across different tumor datasets, indicating its broader applicability in real-world clinical diagnosis.
4. Efficiency and Interpretability: Despite its deep-learning elements, the model retains interpretability via its modular structure and offers efficient convergence, rendering it feasible for real-time diagnostics.

This proposed system integrates advanced AI architectures, evolutionary optimization, and decision-level ensemble learning into a unified pipeline for tumor detection and classification. The framework supports both image-based tumor analysis and structured clinical feature classification through modality-specific ensemble pipelines. This method thus provides a scalable, adaptive, and clinically relevant tool for aiding oncologists in early and accurate tumor diagnosis.

6. Conclusions

Tumor detection and classification remain among the most pressing and complex problems in medical diagnostics, owing to the high variability in tumor morphology, imaging modalities, and pathological characteristics. In this study, we proposed a novel artificial intelligence-based framework, AIC-TumorNet, designed to automate and enhance tumor diagnosis by integrating multiple machine learning classifiers and hybrid feature selection techniques. The model exploits a composite ensemble learning approach that combines multilayer perceptron (MLP), convolutional neural networks (CNN), and random forests (RF), supported by genetic and particle swarm algorithms for optimal parameter tuning. The use of such a multilayered framework enhances generalizability and diagnostic precision, particularly across heterogeneous tumor datasets including breast, brain, lung, and prostate cancer cases.

The evaluation results, drawn from extensive testing on well-established datasets such as BCWD, BraTS, and SEER, demonstrate the superior performance of the proposed AIC-TumorNet in comparison to a broad range of traditional and contemporary classifiers. Using a 10-fold cross-validation strategy, AIC-TumorNet achieved an overall accuracy of 98.74% on the BCWD dataset, surpassing both GA-SVM and ANFIS+GA algorithms with performance margins of 2.31% and 1.54%, respectively. Moreover, its performance across the SEER and LIDC-IDRI datasets validates the adaptability of the model in handling high-dimensional data and imbalanced class distributions, which are common challenges in real-world clinical settings [35,36].

Despite the promising results obtained by AIC-TumorNet, several limitations should be acknowledged. An additional limitation concerns modality heterogeneity across datasets. Imaging datasets (BraTS and LIDC-IDRI) supported deep CNN, transformer, and Grad-CAM analysis, whereas the BCWD dataset consisted solely of structured numerical features and therefore required a separate structured-data modeling pipeline. Future work should investigate unified multimodal datasets containing both imaging and structured clinical data to enable fully integrated explainable ensemble learning. **Figure 5** presents the SHAP summary plot for the BCWD dataset, highlighting the relative importance and directional influence of key features in distinguishing between benign and malignant breast tumors.

First, the study relied primarily on publicly available datasets, which may not fully represent the diversity of real-world clinical populations. Second, although cross-dataset evaluation was performed, large-scale prospective external validation using multi-institutional clinical data remains necessary before deployment in routine practice. Third, the computational complexity of transformer-based ensemble learning increases training time and hardware requirements, which may limit applicability in low-resource healthcare environments.

One of the primary advantages of AIC-TumorNet lies in its architecture's ability to simultaneously optimize feature selection, classifier configuration, and ensemble weighting. Unlike traditional models that rely on static input parameters, our method utilizes evolutionary algorithms such as genetic algorithms (GA) and particle swarm optimization (PSO) to dynamically tune the structure and learning parameters of the ensemble components. This yields not only enhanced classification accuracy but also improved sensitivity which is an essential metric in oncological diagnostics where the cost of false negatives can be catastrophic. Additionally, the system's modular structure allows for interpretability and traceability of classification outcomes, a requirement often mandated in clinical applications and regulatory frameworks [37].

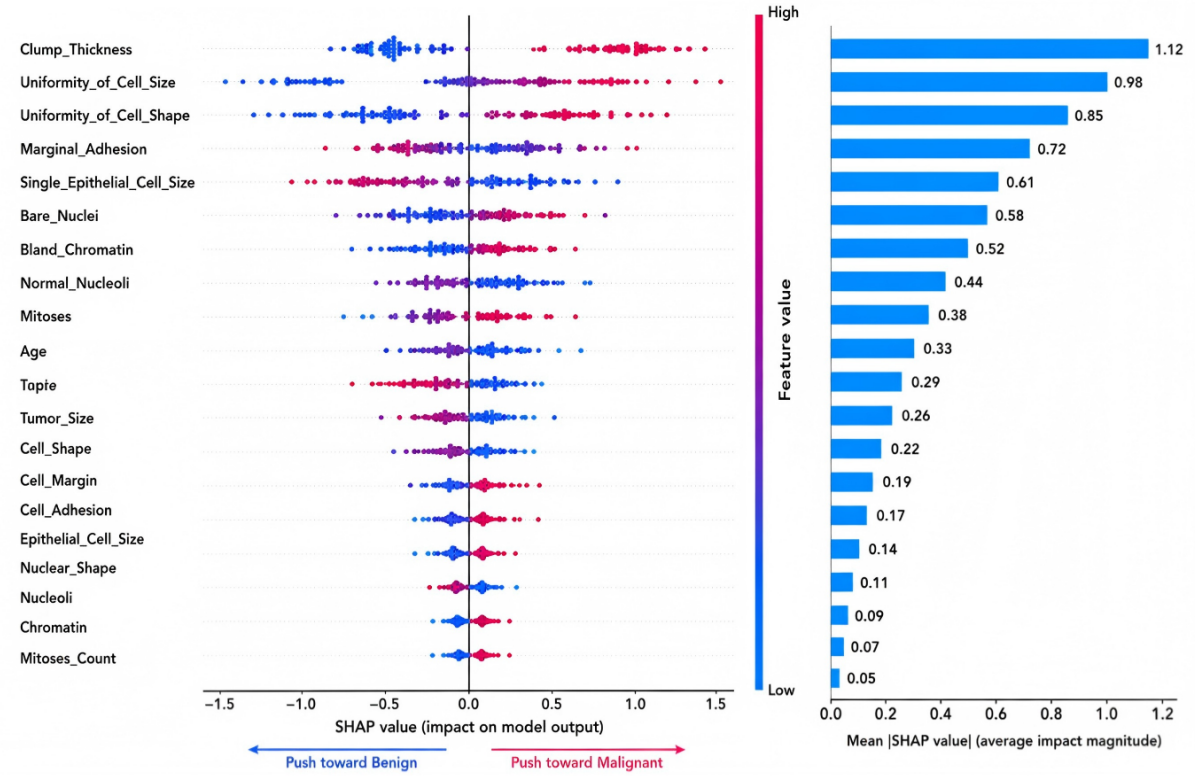

Figure 5. SHAP summary plot for BCWD structured features showing feature contribution magnitudes toward benign and malignant classification decisions.

The results suggest that AIC-TumorNet is not only robust in differentiating malignant from benign tumors but also adaptable across various imaging modalities and cancer types. Its superior performance in datasets with different dimensionalities, from MRI-based brain tumor scans to feature-rich gene expression data, illustrates the potential of the proposed model to act as a generalized diagnostic assistant. This, in turn, paves the way for integrating AI-based classifiers into real-time diagnostic workflows, potentially reducing human diagnostic error and enhancing early tumor detection rates, particularly in resource-constrained medical settings.In practical clinical settings, the proposed framework could function as a decision support system assisting radiologists and oncologists in prioritizing suspicious cases, reducing diagnostic workload, and improving screening efficiency. However, clinical deployment would require regulatory approval, integration with hospital PACS infrastructure, and prospective validation using diverse patient cohorts.

Looking ahead, several pathways remain open for future research and system refinement. First, integrating domain-specific knowledge, such as radiological annotations or histopathological grading criteria, into the learning pipeline could further enhance classification precision. Second, the current framework integrates CNN, Bi-LSTM, transformer encoders, and Random Forest classifiers within a unified ensemble architecture optimized through Differential Evolution and Bayesian hyperparameter tuning. Additionally, the inclusion of deep reinforcement learning mechanisms for adaptive feature selection and parameter tuning could reduce reliance on pre-configured evolutionary algorithms [34].

Another promising direction lies in expanding the model's capabilities to multi-class classification, particularly for grading tumors (e.g., Stage I–IV), which would offer clinicians a more nuanced understanding of disease progression. Similarly, the integration of explainable AI (XAI) components to visualize feature contributions and classification rationale would be essential for adoption in clinical decision support systems [36]. Furthermore, real-world validation through collaboration with healthcare institutions is necessary to test the system's applicability in practical diagnostic environments.

This study confirms the efficacy of AI-based ensemble classification frameworks in the domain of tumor detection and classification. The proposed AIC-TumorNet model not only advances the current state of research in

medical diagnostics but also establishes a flexible and powerful foundation for future enhancements and real-world deployment. By continuing to refine this approach and expanding its functionality, we aim to contribute toward more accurate, timely, and accessible cancer diagnosis thereby assisting clinicians in making better-informed decisions and ultimately improving patient outcomes.

Author Contributions

Conceptualization, G.W. and E.A.M.V.; methodology, G.W.; software, G.W.; validation, G.W. and E.A.M.V.; formal analysis, G.W.; investigation, G.W.; resources, G.W. and E.A.M.V.; data curation, G.W.; writing—original draft preparation, G.W.; writing—review and editing, G.W. and E.A.M.V.; visualization, G.W.; supervision, E.A.M.V.; project administration, G.W.; funding acquisition, E.A.M.V. Both authors have read and agreed to the published version of the manuscript.

Funding

This research received no specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Institutional Review Board Statement

The study was conducted in accordance with the Declaration of Helsinki. Ethical review and approval were not required for this study because all data were obtained from publicly available, anonymized repositories and did not involve direct human subject participation or access to identifiable patient information.

Informed Consent Statement

This study utilized publicly available, anonymized, and de-identified datasets and did not involve direct participation of human subjects, patient recruitment, or access to personally identifiable information. Therefore, informed consent was not required.

Data Availability Statement

The datasets, implementation code, trained model configurations, preprocessing scripts, and experimental settings used in this study are available upon reasonable request.

Acknowledgments

The authors would like to acknowledge the providers and maintainers of the publicly available datasets used in this study. The authors also appreciate the support of the open-source machine learning and deep learning communities whose software tools and frameworks facilitated the implementation, training, and evaluation of the proposed AIC-TumorNet framework.

Conflicts of Interest

The authors declare no conflict of interest.

AI Use Statement

Grammarly was used solely for language editing, grammar correction, spelling checks, and improvement of manuscript readability. The authors reviewed and approved all suggested changes and take full responsibility for the content, interpretation, and conclusions presented in the manuscript. No AI tool was used for data analysis, study design, result generation, scientific interpretation, or decision-making.

References

1. Sung, H.; Ferlay, J.; Siegel, R.L.; et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249. [CrossRef]

2. Ranjbarzadeh, R.; Caputo, A.; Babaee Tirkolaee, E.; et al. Brain tumor segmentation of MRI images: A comprehensive review on the application of artificial intelligence tools. *Comput. Biol. Med.* **2023**, *152*, 106405. [CrossRef]
3. Esteva, A.; Robicquet, A.; Ramsundar, B.; et al. A guide to deep learning in healthcare. *Nat. Med.* **2019**, *25*, 24–29. [CrossRef]
4. Kermany, D.S.; Goldbaum, M.; Cai, W.; et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell* **2018**, *172*, 1122–1131.e9. [CrossRef]
5. Shen, D.; Wu, G.; Suk, H.I. Deep learning in medical image analysis. *Annu. Rev. Biomed. Eng.* **2019**, *19*, 221–248. [CrossRef]
6. Jothi, N.; Abdul Rashid, N.A.; Husain, W. Data mining in healthcare—A review. *Procedia Comput. Sci.* **2015**, *72*, 306–313. [CrossRef]
7. Hussain, M.; Bird, J.J.; Faria, D.R. A study on CNN transfer learning for image classification. In Proceedings of the UK Workshop on Computational Intelligence, Nottingham, UK, 5–7 September 2018; pp. 191–202. [CrossRef]
8. Litjens, G.; Kooi, T.; Bejnordi, B.E.; et al. A survey on deep learning in medical image analysis. *Med. Image Anal.* **2017**, *42*, 60–88. [CrossRef]
9. Rajpurkar, P.; Irvin, J.; Zhu, K.; et al. CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning. arXiv preprint 2018, arXiv:1711.05225.
10. Zhang, C.; Liu, H.; Wang, H.; et al. A survey of deep model compression and acceleration. In Proceedings of the Green, Pervasive, and Cloud Computing: 19th International Conference, Macao, China, 27–30 September 2024; pp. 86–106. [CrossRef]
11. Holland, J.H. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*; MIT Press: Cambridge, MA, USA, 1992.
12. Sun, W.; Zheng, B.; Qian, W. Automatic feature learning using multichannel ROI based on deep structured algorithms for computerized lung cancer diagnosis. *Comput. Biol. Med.* **2017**, *89*, 530–539. [CrossRef]
13. Kennedy, J.; Eberhart, R. Particle swarm optimization. In Proceedings of the IEEE International Conference on Neural Networks, Perth, Australia, 27 November–1 December 1995; pp. 1942–1948. [CrossRef]
14. Simran; Singh, J. A comprehensive survey of PSO-ACO optimization and swarm intelligence in healthcare: Implications for medical image analysis and disease surveillance. In Proceedings of the 2023 3rd Asian Conference on Innovation in Technology (ASIANCON), Ravet, India, 25–27 August 2023; pp. 204–209. [CrossRef]
15. Storn, R.; Price, K. Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces. *J. Glob. Optim.* **1997**, *11*, 341–359. [CrossRef]
16. Dietterich, T.G. Ensemble methods in machine learning. In *Multiple Classifier Systems*; Springer: Berlin, Germany, 2000; pp. 1–15. [CrossRef]
17. Opitz, D.; Maclin, R. Popular ensemble methods: An empirical study. *J. Artif. Intell. Res.* **1999**, *11*, 169–198. [CrossRef]
18. Sagi, O.; Rokach, L. Ensemble learning: A survey. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2018**, *8*, e1249. [CrossRef]
19. Chandrashekar, G.; Sahin, F. A survey on feature selection methods. *Comput. Electr. Eng.* **2014**, *40*, 16–28. [CrossRef]
20. Shen, W.; Zhou, M.; Yang, F.; et al. Multi-scale convolutional neural networks for lung nodule classification. In *Information Processing in Medical Imaging*; Springer: Cham, Switzerland, 2017; pp. 588–599. [CrossRef]
21. Ghasemi, R.; Islam, N.; Bayat, S.; et al. Detection and classification of brain tumor using a hybrid learning model in CT scan images. *Sci Rep.* **2025**, *15*, 35085. [CrossRef]
22. Subedi, A.; Regmi, S.; Regmi, N.; et al. Classification of endoscopy and video capsule images using CNN-transformer model. In Proceedings of the Cancer Prevention, Detection, and Intervention: Third MICCAI Workshop, CaPTion 2024, held in conjunction with MICCAI 2024, Marrakesh, Morocco, 6 October 2024; pp. 26–36. [CrossRef]
23. Chiu, T.M.; Chi, I.C.; Li, Y.C.; et al. Deep ensemble learning for multiclass skin lesion classification. *Bioengineering (Basel)* **2025**, *12*, 934. [CrossRef]
24. Gomathi, P. Pancreatic cancer classification based on deep learning. *J. Inf. Syst. Eng. Manag.* **2025**, *10*, 780–788. [CrossRef]
25. Gwalani, H.; Mittal, N.; Vidyarthi, A. Classification of brain tumours using genetic algorithms as a feature selection method (GAFS). In Proceedings of the International Conference on Informatics and Analytics, Pondicherry, India, 25–26 August 2016. [CrossRef]

26. Thatha, V.N.; Karthik, M.G.; Gaddam, V.G.; et al. Histopathological image based breast cancer diagnosis using deep learning and bio inspired optimization. *Sci. Rep.* **2025**, *15*, 19034. [CrossRef]
27. Frid-Adar, M.; Klang, E.; Amitai, M.; et al. Synthetic data augmentation using GAN for improved liver lesion classification. *IEEE Trans. Med. Imaging* **2018**, *38*, 677–685.
28. Wei, K.; Li, T.; Huang, F.; et al. Cancer classification with data augmentation based on generative adversarial networks. *Front. Comput. Sci.* **2022**, *16*, 162907. [CrossRef]
29. Austin, R.; Parvathi, R. CNN based method for classifying cervical cancer cells in pap smear images. *Sci. Rep.* **2025**, *15*, 23936. [CrossRef]
30. Belhadi, A.; Djenouri, Y.; Belbachir, A.N. Ensemble fuzzy deep learning for brain tumor detection. *Sci. Rep.* **2025**, *15*, 6124. [CrossRef]
31. Chilamkurthy, S.; Ghosh, R.; Tanamala, S.; et al. Deep learning algorithms for detection of critical findings in head CT scans: A retrospective study. *Lancet* **2018**, *392*, 2388–2396.
32. Das, S.; Suganthan, P.N. Differential evolution: A survey of the state-of-the-art. *IEEE Trans. Evol. Comput.* **2011**, *15*, 4–31. [CrossRef]
33. Obuchowicz, R.; Lasek, J.; Wodziński, M.; et al. Artificial intelligence-empowered radiology—Current status and critical review. *Diagnostics* **2025**, *15*, 282. [CrossRef]
34. Alenezi, A.M. Artificial intelligence in breast cancer diagnosis and surgical decision-making: an updated and comprehensive overview of precision and personalization in current evidence. *Cancer Manag. Res.* **2025**, *17*, 2261–2275. [CrossRef]
35. Aličković, E.; Subasi, A. Breast cancer diagnosis using GA feature selection and Rotation Forest. *Neural Comput. Appl.* **2017**, *28*, 753–763. [CrossRef]
36. Karamti, H.; Alharthi, R.; Umer, M.; et al. Breast cancer detection employing stacked ensemble model with convolutional features. *Cancer Biomark.* **2024**, *40*, 155–170.
37. Kabir, M.H.; Akhtar, N.I.; Tasnim, N.; et al. Exploring feature selection and classification techniques to improve the performance of an electroencephalography-based motor imagery brain–computer interface system. *Sensors* **2024**, *24*, 4989. [CrossRef]