



REVIEW



Advanced Feature Selection Techniques in Medical Imaging—A Systematic Literature Review

Sunawar Khan¹, Tehseen Mazhar^{1,2,*}, Naila Sammar Naz¹, Fahad Ahmed¹, Tariq Shahzad³, Atif Ali⁴, Muhammad Adnan Khan^{5,*} and Habib Hamam^{6,7,8,9}

¹School of Computer Science, National College of Business Administration and Economics, Lahore, 54000, Pakistan

²Department of Computer Science and Information Technology, School Education Department, Government of Punjab, Layyah, 31200, Pakistan

³Department of Computer Engineering, COMSATS University Islamabad, Sahiwal Campus, Sahiwal, 57000, Pakistan

⁴Research Management Centre (RMC), Multimedia University, Cyberjaya Campus, 63100, Malaysia

⁵Department of Software, Faculty of Artificial Intelligence and Software, Gachon University, Seongnam-si, 13120, Republic of Korea

⁶Faculty of Engineering, Uni de Moncton, Moncton, NB E1A3E9, Canada

⁷School of Electrical Engineering, University of Johannesburg, Johannesburg, 2006, South Africa

⁸International Institute of Technology and Management (IITG), Av. Grandes Ecoles, Libreville, BP 1989, Gabon

⁹College of Computer Science and Engineering, University of Ha'il, Ha'il, 55476, Saudi Arabia

*Corresponding Authors: Tehseen Mazhar. Email: tehseenmazhar719@gmail.com;

Muhammad Adnan Khan. Email: adnan@gachon.ac.kr

Received: 21 April 2025; Accepted: 04 August 2025; Published: 23 September 2025

ABSTRACT: Feature selection (FS) plays a crucial role in medical imaging by reducing dimensionality, improving computational efficiency, and enhancing diagnostic accuracy. Traditional FS techniques, including filter, wrapper, and embedded methods, have been widely used but often struggle with high-dimensional and heterogeneous medical imaging data. Deep learning-based FS methods, particularly Convolutional Neural Networks (CNNs) and autoencoders, have demonstrated superior performance but lack interpretability. Hybrid approaches that combine classical and deep learning techniques have emerged as a promising solution, offering improved accuracy and explainability. Furthermore, integrating multi-modal imaging data (e.g., Magnetic Resonance Imaging (MRI), Computed Tomography (CT), Positron Emission Tomography (PET), and Ultrasound (US)) poses additional challenges in FS, necessitating advanced feature fusion strategies. Multi-modal feature fusion combines information from different imaging modalities to improve diagnostic accuracy. Recently, quantum computing has gained attention as a revolutionary approach for FS, providing the potential to handle high-dimensional medical data more efficiently. This systematic literature review comprehensively examines classical, Deep Learning (DL), hybrid, and quantum-based FS techniques in medical imaging. Key outcomes include a structured taxonomy of FS methods, a critical evaluation of their performance across modalities, and identification of core challenges such as computational burden, interpretability, and ethical considerations. Future research directions—such as explainable AI (XAI), federated learning, and quantum-enhanced FS—are also emphasized to bridge the current gaps. This review provides actionable insights for developing scalable, interpretable, and clinically applicable FS methods in the evolving landscape of medical imaging.

KEYWORDS: Feature selection; medical imaging; deep learning; hybrid approaches; multi-modal imaging; quantum computing; explainable AI; computational efficiency; dimensionality reduction



1 Introduction

In artificial intelligence-driven medical imaging, FS is an essential component since it enables the management of a substantial quantity of high-dimensional imaging data, enhancing diagnostic accuracy and interpretability. In the first part of this section, we will present the idea of FS, discuss its significance in medical imaging, and then examine the many FS approaches currently in use. These methods include those based on classical techniques, deep learning methods, fusions of classical and deep learning methods, and quantum mechanics. In addition, it details the primary research questions, the aims of the investigation, and the study's contribution. At the end of the section, the structure of the paper is presented to the readers to direct them through the topics contained within the remaining sections of the publication.

1.1 Motivation and Objectives

Medical imaging has become an indispensable tool in modern diagnostics, offering rich data for the detection and treatment of various diseases. However, the high dimensionality and heterogeneity of medical image data pose significant challenges for effective analysis and interpretation. This study is motivated by the need to optimize medical image analysis workflows through advanced feature selection (FS) techniques that enhance accuracy, efficiency, and interpretability. The primary objectives of this review are: (1) to classify and evaluate classical and deep learning-based FS methods used in medical imaging; (2) to explore hybrid and quantum-enhanced FS approaches; and (3) to highlight future directions for clinically applicable and computationally scalable FS strategies.

1.2 Context

Medical imaging is essential to contemporary medicine since it assists in disease detection, diagnosis, and treatment planning. The use of modern imaging modalities, including MRI, CT, PET, ultrasound, and X-ray, results in the production of a large quantity of high-resolution anatomical and functional data [1]. Because these datasets are high dimensional, redundant, and noisy, it is necessary to employ appropriate methodologies to process them using the machine learning models that are currently accessible [2]. AI-driven techniques, in particular, ML and DL, have brought about a revolution in the field of medical image analysis [3]. However, these techniques are severely impacted by the curse of dimensionality, which significantly impacts computational inefficiency, lack of generalization, and consequences of interpretability. FS algorithms have recently been proposed as a fundamental preprocessing step in medical image analysis to address these issues [4].

The procedure known as FS is designed to pick the most pertinent and informative characteristics within the dataset while simultaneously removing irrelevant or redundant features [5]. FS improves diagnostic accuracy, reduces computing costs, and offers insight into the operation of AI-based diagnostic systems, ultimately making the systems more clinically viable and efficient [6]. Traditional FS methods, such as filter, wrapper, and embedding methods, have been utilized in a significant amount of feature optimization; however, they are confronted with the challenge of scaling with huge, high-dimensional imaging datasets features [7]. Approaches to FS based on deep learning, such as CNN, Autoencoders, and Transformers, intrinsically learn the distinguishing features from the medical image and perform better than other methods [8]. The drawback of this, however, is that deep learning models are challenging to comprehend, which leads to a lack of confidence in their application in clinical settings.

Despite this, hybrid FS techniques, which combine traditional FS methods with methods based on deep learning, have gained popularity to overcome the constraints discussed earlier [9]. These methods aim to achieve equivalent FS efficiency, good model interpretability, and diagnostic accuracy [10]. This makes them more suitable for clinical applications in the real world. With medical imaging expanding into differences

of multi-modal data integration (for example, integrating MRI, CT, and PET), FS is also confronted with other issues, such as combining and aligning multiple modalities, which calls for more sophisticated selection algorithms [11]. These challenges underscore the need for advanced solutions in the field of medical imaging.

As the field of quantum computing continues to advance, new FS approaches that use quantum principles have evolved, offering the exciting possibility of providing high-dimensional medical data that can be analyzed more effectively [12]. Even though there are still outstanding problems regarding the practical implementations of quantum FS-based approaches, these challenges include scalability and independent clinical validations [13]. Nevertheless, many significant issues still exist, such as the absence of standard evaluation benchmarks and the lack of comprehension of interpretability, ethics, and computational resources. Solving these problems will improve the reliability and uptake of AI-driven FS techniques for use in clinical settings and raise their transparency.

Considering the growing importance of FS in AI-driven medical imaging, this review aims to:

- To analyze existing FS techniques applied in medical imaging, focusing on classical, deep learning-based, hybrid, and quantum FS approaches.
- To compare the effectiveness of different FS methods regarding accuracy, computational efficiency, and clinical interpretability.
- To explore the role of FS in multi-modal medical imaging and the challenges associated with feature fusion.
- To investigate the emerging applications of quantum-based FS and their potential advantages over classical methods, offering hope for significant improvements in medical imaging.
- To highlight the key challenges, ethical considerations, and future research directions in AI-driven FS for medical imaging.

To achieve these objectives, this review addresses the following research questions:

- **RQ1:** How have classical FS techniques evolved, and what are their limitations in handling high-dimensional medical imaging data?
- **RQ2:** To what extent do deep learning-based FS methods improve diagnostic accuracy, and how do they compare in terms of interpretability?
- **RQ3:** How can hybrid FS techniques integrate classical and deep learning methods to enhance clinical decision-making?
- **RQ4:** What role does FS play in multi-modal medical imaging, and how can feature fusion strategies be optimized?
- **RQ5:** How does quantum computing revolutionize FS in medical imaging, and what are its potential applications?
- **RQ6:** What are the key challenges, ethical considerations, and open research problems in AI-driven FS for medical imaging?

This study provides a comprehensive, structured analysis of FS techniques in medical imaging, making the following key contributions:

- Systematic categorization and comparison of FS methods, covering classical, deep learning, hybrid, and quantum-based approaches.
- Critical evaluation of FS techniques regarding accuracy, computational efficiency, and interpretability.
- Discuss multi-modal FS strategies, focusing on feature fusion and dimensionality reduction challenges.
- Explore quantum computing applications in FS and assess their feasibility for real-world implementation.

- Identification of key challenges, ethical considerations, and open research problems in AI-driven FS for medical imaging.

In addition, this study is a helpful resource for researchers, AI practitioners, and healthcare professionals interested in understanding the current status of FS in medical imaging, the obstacles it faces, and the trends expected to emerge.

Here is how the remainder of this paper is structured. A complete discussion of feature selection strategies is presented in [Section 2](#). [Section 3](#) describes the systematic literature review process, including the inclusion criteria, the search technique, and the data extraction. Techniques are broken down into four categories: classical, deep learning, hybrid, and quantum-based methods, and a comparison of these methods is also included. In the final section, [Section 4](#), the proposed study explores the difficulties and potential future research opportunities associated with applying FS to medical imaging. These problems include the incorporation of multi-modal data, the interpretation of medical imaging models, the resolution of ethical concerns, and the spending of computational resources. In conclusion, [Section 5](#) brings the study to a close by providing a list of the most critical points and a summary of the potential paths the research could take to establish frameworks for artificial intelligence-driven medical imaging.

1.3 Research Gap

Despite the significant advancements in classical and deep learning-based feature selection methods, several limitations remain. Classical techniques often struggle with high-dimensional, multi-modal datasets, lacking the flexibility to adapt to complex spatial and semantic features present in modern medical images. Deep learning-based methods, while powerful, are frequently criticized for their black-box nature and require large amounts of labeled data. Furthermore, limited attention has been paid to privacy-preserving and quantum-enhanced FS frameworks, which are critical in emerging clinical applications. Finally, while hybrid FS strategies have shown promise, systematic evaluations comparing their efficacy across imaging modalities are scarce. These gaps underline the need for a comprehensive review that bridges classical, deep learning, and quantum paradigms to guide future research in this domain.

2 Literature Review

Selecting features through FS reduces medical imaging dataset dimensions by maintaining only the most important diagnostic features for classifying diseases. The three categories of classical FS techniques include filter, wrapper, and embedded methods, which provide specific benefits and constraints. Experimental researchers have successfully used this approach for medical imaging investigations that involve labeling tumors, spotting lesions, and making disease predictions. This part investigates original FS techniques, including their mathematical frameworks and medical imaging applications.

Countervalue selection allows feature classification through statistical assessments that bypass machine learning model operations [\[14\]](#). The methods demonstrate computational efficiency and run independently of classifiers, enabling them to effectively handle large medical imaging datasets [\[15\]](#). The three standard filter techniques for data selection are Pearson Correlation, Mutual Information, and Chi-square tests.

The analysis measures linear interconnections between two variables by using Pearson Correlation. The FS system uses Pearson Correlation to locate important target-related features, removing unneeded attributes [\[16\]](#). The formula for Pearson Correlation coefficient r appears as follows in [Eq. \(1\)](#):

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (1)$$

where X_i and Y_i represent the feature values and target labels, \bar{X} and \bar{Y} are their respective means, and n denotes the total number of samples.

The absolute r value shows strong correlation levels when nearing +1 or -1, but low correlations exist when values approach zero. Applying Pearson Correlation in medical imaging helps eliminate redundant feature elements, thus maintaining independent features that matter for analysis [17].

Mutual Information provides an assessment of variable dependence through its evaluation of information reduction when one variable becomes known [18]. It is defined as in (2):

$$I(X;Y) = \sum_{x \in X} \sum_{y \in Y} P(x,y) \log \left(\frac{P(x)P(y)}{P(x,y)} \right) \quad (2)$$

where:

- $P(x,y)$ is the joint probability distribution of X and Y .
- $P(x)$ and $P(y)$ are the marginal probabilities of X and Y .

The Medical Imaging field benefits from MI-based FS because this approach optimizes the selection of variables with complex nonlinear trends better than traditional methods, such as those used in tumor segmentation procedures [19]. Categorical features and the target variable demonstrate statistical dependence according to the Chi-square test [20]. It is computed in Eq. (3):

$$\chi^2 = \sum \frac{(O - E)^2}{E} \quad (3)$$

where:

- O represents observed frequency.
- E represents expected frequency.

The analysis of discrete feature selection tasks mainly employs this technique to identify significant predictors of disease within medical images based on their histogram-based features [21]. Wrapper methods execute a feature subset evaluation loop by running machine learning model training and validation procedures across all combinations to select the subset demonstrating superior model performance [22]. Although such methods deliver optimized feature collections, their computational cost remains high. Recursive Feature Elimination (RFE) and Forward/Backward Selection are among the preferred wrapper approaches in feature selection applications [23]. Algorithm 1 has been included to illustrate the feature selection process used in the study.

Algorithm 1: Recursive feature elimination (RFE)

Input: Feature set $F = \{f_1, f_2, \dots, f_n\}$, base model (e.g., SVM)

Output: Optimal feature subset F^*

1. Train the base model on the full feature set F
2. Compute feature importance scores
3. Repeat:
 - a. Remove the least important feature
 - b. Retrain the model on the reduced set
 - c. Recompute importance scores
 Until the desired number of features is reached
4. Return final selected feature subset F^*

The feature selection methodology, RFE, accurately produces an optimal set of features because it builds upon a series of benefits. For datasets of all sizes, from small to medium, RFE provides an effective solution to choose essential features that support model performance [24]. The main disadvantage of RFE is its high computational burden, affecting its performance with large-scale datasets with numerous features. The medical imaging field mainly utilizes RFE due to its vital function in identifying discriminative features for applications such as tumor detection and Alzheimer's disease classification to build more precise diagnostic models [25].

A set of standard feature selection techniques include Forward Selection as well as Backward Selection. Forward Selection starts by adding no features and continues by including new features individually to assess their impact on performance at each selection step [26]. The feature selection process of Backward Selection begins with all available features, which are then reduced one by one based on performance criteria [27]. The techniques are broadly used in cardiovascular disease prediction systems because they analyze ECG features successfully [28]. These selection methods improve classification accuracy by choosing optimal wavelet coefficients, leading to more accurate and efficient predictive cardiovascular condition diagnosis [29]. Embedded methods execute feature selection operations within their model training process by integrating the selection procedure with model learning procedures. Massive medical imaging applications favor these methods because they merge efficient performance for practical applications [30]. LASSO (Least Absolute Shrinkage and Selection Operator) allows regression-based feature selection through L1 regularization until coefficients reach absolute zero values, which selects the vital features alone [31]. The LASSO optimization function is defined in Eq. (4):

$$\min \left(\sum_{i=1}^n (y_i - X_i \beta)^2 + \lambda \sum_{j=1}^p |\beta_j| \right) \quad (4)$$

where:

- y_i represents the target variable.
- X_i represents the feature matrix.
- β represents feature weights.
- λ is a regularization parameter controlling sparsity.

LASSO is a popular medical imaging biomarker identification tool for early cancer detection and Alzheimer's disease progression assessment [32]. Random Forest as an ensemble learning method automatically generates feature importance scores through its capability to measure how features reduce classification errors [33]. Feature importance is calculated in Eq. (5):

$$I(f_j) = \sum_{t \in T} \frac{|T|}{|T|} \Delta \text{Gini}_t(f_j) \quad (5)$$

where

- $I(f_j)$ is the importance score of feature f_j .
- $\Delta \text{Gini}_t(f_j)$ represents the decrease in Gini impurity when splitting on feature f_j .
- T is the set of all trees in the forest.

Random Forest-based FS offers several advantages, particularly in the context of high-dimensional and multi-modal medical imaging datasets. The method shows effectiveness when dealing with extensive complex datasets in numerous medical applications [34]. In medical imaging, where data is abundant, feature selection is crucial to pattern identification and machine learning. In medical imaging, feature

selection improves classification accuracy and simplifies data. These methods improve multi-sourced data treatment and medical infection diagnosis [35]. Scientific studies confirm that Random Forest-based FS produces successful results for brain tumor identification through feature prioritization of extracted radiomics characteristics from MRI medical images. The method enhances the diagnostic model's reliability and interpretability to produce better decisions supporting medical settings [36].

FS methods provide efficient solutions that are easy to interpret while having medical imaging dimensional reduction tasks. The filter methods of Mutual Information and Chi-square process fast calculations that help identify features holding crucial statistical information. Extensive computation expenses characterize the model-specific performance optimization of RFE and Forward/Backward Selection methods because they function as wrapper methods. The combination occurs in model training when using LASSO and Random Forest Feature Importance, achieving precise predictions and optimal operational speed. Research protocols based on these strategies achieve their goals even though complex multi-modal imaging data creates challenges that inspire scientists to develop next-generation analysis methods using deep learning frameworks and quantum-filtering techniques. Here's a structured comparison table for classical FS methods in medical imaging in Table 1:

Table 1: Comparison of classical FS methods in medical imaging

Category	Method	Advantages	Disadvantages	Applications
Filter	Pearson correlation [37]	Simple, fast; ignores nonlinearities	Feature selection for radiomics-based tumor classification	Feature selection for radiomics-based tumor classification
Filter	Mutual Information (MI) [38]	Captures nonlinear trends; costly in high-dim data	Brain MRI segmentation, lesion detection	Brain MRI segmentation, lesion detection
Filter	Chi-Square Test [39]	Effective for categorical data; not for continuous features	Histopathology image classification	Histopathology image classification
Wrapper	Recursive Feature Elimination (RFE) [40]	Optimizes subsets; computationally intensive	ECG-based heart disease classification, MRI-based tumor segmentation	ECG-based heart disease classification, MRI-based tumor segmentation
Wrapper	Forward selection [41]	Improves model efficiency; slow for large sets	Feature selection in CT-based lung cancer detection	Feature selection in CT-based lung cancer detection
Wrapper	Backward selection [42]	Efficient dimensionality reduction; risk of early feature loss	Cardiovascular disease risk prediction using ultrasound imaging	Cardiovascular disease risk prediction using ultrasound imaging
Embedded	LASSO regression [43]	Sparse selection; sensitive to λ	Biomarker selection in MRI-based Alzheimer's diagnosis	Biomarker selection in MRI-based Alzheimer's diagnosis

Recent imaging technology relies on feature selection to screen relevant data without overfitting or computation. This paper compares approaches based on computational efficacy, feature interaction, and high-dimensional data, giving a model for clinicians and academics to co-develop new diagnostics tools [44]. Random Forest-based FS offers several advantages, particularly in the context of high-dimensional and multi-modal medical imaging datasets. The method effectively processes complex, voluminous data, allowing for its use across various medical applications [45]. The explainable feature selection process is vital to RF-based FS because it enables medical professionals to discern the most crucial features utilized in diagnosis. Brain tumor diagnostic classification benefits from random forest-based FS through MRI scan

radiomics features, providing accurate predictive assessments of tumor types [46]. This method enhances the diagnostic model's reliability and interpretability, supporting informed medical decisions in real-world clinical settings.

Usually, features are extracted from the images followed by dimensionality reduction, applying an FS technique to prune abundant features and superior classification performance results. Deep learning can automatically extract hierarchical representations of medical images from neural networks without resorting to handcrafted features and predefined statistical measures, as against the classical FS methods [47]. Three DL based FS approaches, i.e., CNN-based FS, autoencoder-based FS, and transformer-based FS, along with their methodologies and their applications in medical imaging, are explored in this section.

The ability to capture spatial hierarchies and extract discriminative features from raw image data has particularly fueled recent advances CNNs as a tool for FS in medical imaging. CNN is a data-driven way of constructing features, unlike the FS techniques in the existing literature, which rely on tailored feature engineering [48]. These CNNs comprise a few layers, namely convolutional, pooling, and fully connected. The feature extraction part in a CNN can be written directly in math form [49]. It is defined in Eq. (6) as follows:

$$F_l = \sigma (W_l * F_{l-1} + b_l) \quad (6)$$

where:

- F represents the feature map at layer l ,
- W_l is the convolutional filter,
- $*$ denotes the convolution operation,
- b_l is the bias term, and
- σ is the activation function (e.g., ReLU).

To eliminate irrelevant features, CNN-based FS techniques identify the most significant feature maps at various convolutional stages [50]. Feature importance scores are computed from such methods as Grad-CAM (Gradient-weighted Class Activation Mapping) or Layer-wise Relevance Propagation (LRP), and the final feature set is selected based on those scores [51].

Since medical image analysis problems, such as brain tumor segmentation, diabetic retinopathy detection, and histopathology image analysis, rely on CNN-based feature selection [52,53], it plays a vital role in medical image analysis [54]. For brain tumor segmentation, CNN-based FS increases the tumor identification accuracy, especially in MRI scans, by using the most relevant region of interest, thus increasing the diagnosis and treatment planning precision and same is applied to detect diabetic retinopathy [55]. CNN-based FS is used in fundus images to identify the vascular abnormalities necessary for the early diagnosis of diabetic eye disease and early intervention, which may help prevent vision loss [56]. For instance, ensemble CNN models have recently shown enhanced classification accuracy in brain tumor detection tasks using MRI scans [52]. Transfer learning-based CNN architectures like Mask R-CNN have been applied to prostate segmentation with promising results [53]. Furthermore, CNN models extract high-level tissue patterns, which are helpful for cancer classification in histopathology image analysis. CNN-based FS attends to extreme regions in the images that would contribute the most to improving the accuracy and efficiency of medical image interpretation in several applications [57]. Additionally, hybrid CNN approaches combined with Bayesian methods have improved privacy-preserving feature extraction in smart imaging systems, such as terahertz-based breast cancer detection [58].

Another well-known deep learning approach for FS uses autoencoders (AEs), which are trained to learn compact, low-dimensional feature representations while avoiding losing important information of the high-dimensional medical imaging data [59]. The two main components they consist of are [60]

- **Encoder:** Compresses input data into a lower-dimensional latent space.
- **Decoder:** Reconstructs the original input from the latent space.

FS extends to using a probabilistic latent space, which in Variational Autoencoders (VAEs) helps improve FS by embedding relevant medical information in the extracted features while filtering noise. A VAE's objective function is a reconstruction loss combined with a regularization term [61]. It explains in Eq. (7) as:

$$L = E_{q(z|x)}[\log p(x|z)] - D_{KL}(q(z|x) \| p(z)) \quad (7)$$

where:

- x is the input medical image,
- z is the latent feature representation,
- $q(z|x)$ is the encoder's approximation of the posterior distribution,
- $p(z|x)$ is the decoder's likelihood, and
- D_{KL} is the Kullback-Leibler divergence [62], ensuring the latent space follows a normal distribution.

The sparse autoencoder model effectively reduces high-dimensional data with the least reconstruction error when used unsupervised. It is superior to classical methods in classification problems, particularly when the quantity of labelled data is limited [63]. The author of [64] enforce the sparsity constraint, the sparsity by KL divergence through Eq. (8):

$$\Omega = \sum_{j=1}^h \rho \log \widehat{\rho}_j \rho + (1 - \rho) \log (1 - \widehat{\rho}_j) (1 - \rho) \quad (8)$$

where:

- ρ is the desired sparsity level,
- $\widehat{\rho}_j$ is the average activation of neuron j ,
- h is the number of hidden neurons.

VAE, SAE, and autoencoders are essential for improving disease prediction and detection in medical diagnostics. In particular, autoencoders and CNN architectures have demonstrated utility in classifying early-stage Alzheimer's disease using PET and MRI modalities [65,66]. VAEs are utilized for Alzheimer's disease prediction by extracting relevant biomarkers from MRI scans to distinguish Alzheimer's patients from healthy individuals, enabling early detection and intervention [67]. It is also worth noting that recent research has explored the use of voice biomarkers as prognostic indicators for neurological disorders, such as Parkinson's disease, demonstrating the potential of vocal features combined with machine learning techniques for early diagnosis [68]. For COVID-19 detection, SAEs identify key features of the lungs in chest X-rays or CT scans, allowing for effective and prompt diagnosis [69]. Additionally, in gene expression analysis, autoencoders search for critical genetic features within a high-dimensional medical database to identify genetic markers, which enhances our understanding of complex diseases. Deep learning models significantly impact medical diagnostics, improving accuracy and efficiency [70]. Recently, Transformers, particularly Vision Transformers (ViT), have been recognized as especially promising for medical imaging because they can capture long-range dependencies of an entire image [71]. Unlike CNNs, which utilize local receptive fields, ViTs analyze images in their entirety as sequences of patches and are capable of highlighting

more global features [72]. ViTs segment an image into non-overlapping patches and then map these to a sequence of feature vectors using a linear projection function [73]. Eq. (9) shows that:

$$Z = [Z_1, Z_2, \dots, Z_N] = W_p X + b_p \quad (9)$$

where:

- X is the input image,
- W_p and b_p are learnable parameters,
- Z represents the transformed patch embeddings.

In transformers, the attention mechanism is helpful for FS because the importance scores of features can be computed through the self-attention function [74]. As seen in Eq. (10):

$$\text{Attention}(Q, K, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V \quad (10)$$

where:

- Q, K, V are query, key, and value matrices,
- d_k is the scaling factor.

ViTs have shown promising potential in diverse medical imaging applications, such as representation learning, segmentation, classification, regression, and so on, especially with their capability of modeling long-range dependencies and global contextual features [75]. ViTs offer a significant contribution in brain tumor segmentation because they extract global features from MRI scans to help locate the tumor boundary with high precision, a requirement to accomplish an exact diagnosis and treatment planning [76]. Transformers are excellent for skin lesion classification because they enable the selection of key diagnostic patterns in dermoscopic images and thus better detect and classify various skin conditions like melanoma [77]. For example, ViTs are suitable for biomedical 3D medical imaging since they can efficiently process and represent complex structures, such as 3D CT scans. This capability improves the performance of deep learning models and thus enhances the feature representation and implies better analysis of medical images and, hence, better clinical outcomes. Fig. 1 shows the workflow of Deep Learning based workflow diagram and Table 2 shows the performance of different deep learning feature search techniques.

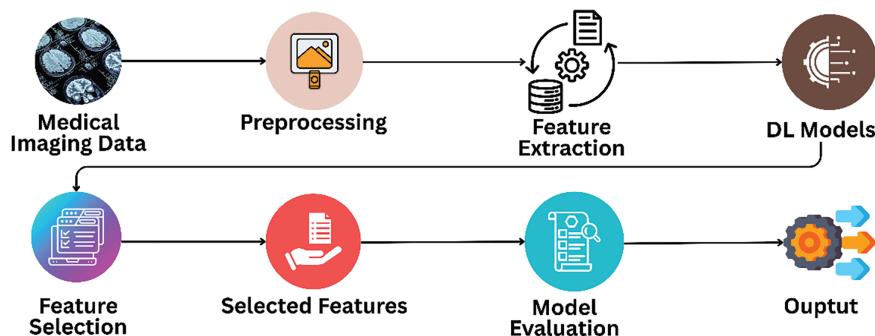


Figure 1: Workflow of Deep Learning-Based Feature Selection (FS) Techniques. This diagram illustrates the typical stages in DL-based FS, including raw medical image input, feature extraction via convolutional or transformer layers, feature importance ranking, and the final selection of diagnostically relevant features used for classification or prediction tasks

Table 2: Performance comparison of deep learning FS techniques

FS technique	Feature extraction mechanism	Computational complexity	Interpretability	Key applications
CNN-Based FS [58]	Hierarchical features from convolution layers	Moderate to High	Low	Tumor detection, Retinopathy, Histopathology
Autoencoder-Based FS [78]	Latent space compression (AE, VAE, SAE)	High	Moderate	Alzheimer's, COVID-19, Gene expression
Transformer-Based FS (ViTs) [79]	Global attention across image patches	Very High	Moderate to High	Brain tumor segmentation, Skin lesion classification, 3D CT/MRI
Transfer Learning FS [80]	Pre-trained CNN or ViT feature encoders	Medium to High	High to Low	MRI-based Alzheimer's classification
Foundation Model FS	Universal representations from large pretrained models	High	High	General-purpose medical imaging, Zero-shot diagnosis

Some recent advances in DL based FS techniques, including CNNs, Autoencoders, and Transformers, are the application of these approaches in medical imaging, which increased feature extraction and selection considerably [81]. CNN-based FS uses convolutional layers to encode spatially meaningful features, Autoencoders encode to compress the data and ensure the retention of vital parts of the information, while Transformers encode to achieve global contextual feature extraction [82]. While there are advantages and disadvantages to each method, combining multiple FS techniques in hybrid methods appears to have great potential to increase accuracy in diagnostic models and reduce the black box characteristic.

Currently, FS in medical imaging is under development through hybrid approaches, which join classical statistical measurement, deep learning architectures, and optimization algorithms [83]. These methods try to overcome the deficiencies of the stand-alone FS techniques and enable their potential to improve accuracy, interpretability, and computational efficiency. Finally, the three major hybrid FS strategies are Classical + Deep Learning Approaches, Metaheuristic optimization-based FS, and Ensemble Learning for FS.

The first one, Classical + Deep Learning FS, merges traditional statistical techniques and deep learning models in feature selection in a way that preserves interpretability in the model [84]. A straightforward solution to this issue is to combine Principal Component Analysis (PCA) with CNNs. PCA reduced dimensionality by selecting the principal components, which were then fed to CNNs for feature extraction and classification. So, this hybrid strategy is aimed to soften the constraints of the 'curse of dimensionality' and to maintain the most usable features for diagnosis [85]. A Genetic Algorithm (GA) with Deep Neural Networks (DNN) represents another popular combination where GA is applied to select the feature subset in simulating natural evolution. This algorithm sequentially selects relevant features according to a fitness function to better generalize and scale down overfitting [86]. The expression of the feature selection process

using GA and deep learning mathematically can be expressed in [Eq. \(11\)](#):

$$F^* = \arg \max \quad \text{Fitness}(F) \quad (11)$$

where F^* is the optimal feature subset, and the fitness function evaluates the classification performance using selected features.

Since dimensional reduction of features has been shown to significantly impact the performance of these models for brain MRI classification, lung nodule detection [87], and histopathology image analysis, these hybrid techniques have proven successful. This review explores the use of DL in medical imaging, particularly for brain tumor analysis, covering segmentation, classification, and prediction techniques. It summarizes key contributions and provides a taxonomy of current research in the field. The article concludes by discussing limitations and future research directions for DL in medical imaging [88].

A second powerful category of algorithms is feature selection based on metaheuristic optimization, where nature-inspired algorithms perform the feature selection. This paper describes these techniques that perform an intelligent search for the optimal feature subsets, which is superior to an exhaustive search [89]. One of the metaheuristic approaches that is mainly used is the GA, which uses selection, crossover, and mutation operations to evolve an optimal feature subset. Particle Swarm Optimization (PSO) is an efficient method by which FS is modeled as a swarm intelligence problem [90]. In PSO, particles (feature subsets) adjust their positions on their own best performance and the best-known feature set. Removing redundant or irrelevant features allows efficient convergence to an optimal solution [91]. Also, Ant Colony Optimization (ACO) follows the foraging behavior of ants to find different feature subsets and refine the selection with pheromone trails [92]. The PSO, which can be used for the FS, is represented mathematically in [Eq. \(12\)](#):

$$V_i(t+1) = wV_i(t) + c_1r_1(P_{i\text{bs}} - X_i(t)) + c_2r_2(G_{\text{bs}} - X_i(t)) \quad (12)$$

where:

- V_i is the velocity of the particle,
- X_i is the particle's position in feature space,
- $P_{i\text{bs}}$ is the best position found by the particle,
- G_{bs} is the best global position among all particles,
- w, c_1, c_2 are weight and learning coefficients.

Due to their ability to enhance classification performance without an intensive computational burden, these metaheuristic techniques have been extensively used in ECG-based heart disease detection, breast cancer classification, and tumor segmentation in MRI.

Ensemble Learning for FS is the third major approach, aggregating multiple FS techniques for better stability and generalization. Instead of using a single selection strategy, ensemble FS methods rely on several models to find the most discriminative features [93]. Bagging-based FS is one frequently used technique that uses bootstrap sampling to create multiple feature subsets and aggregates the selected features through majority voting. Implementing this approach in Random Forest FS reduces variance (and thus overfitting), which is usually what is done when implementing this [94]. Boosting-based FS is the other ensemble feature selection strategy based on improving weak FS models sequentially by focusing on misclassified instances to obtain a better fit and finally refined feature selection [95]. XGBoost FS is a well-known example of such a technique that is efficient and scalable and has been adopted in medical imaging tasks [96]. Finally, Stacking FS, which integrates multiple feature selection models and adopts a meta learner to learn which of those existing feature selection models to use to find the final feature subset with the optimal trade between interpretability and predictive performance. Earlier, I adopted these ensemble FS techniques for multi-modal

medical imaging where combining features from many imaging modalities (e.g., MRI, CT, PET) leads to more accurate diagnostics.

Hybrid FS approaches best solve the challenges with high-dimensional medical imaging data. These comprise classical methods in statistics, deep learning architectures, and metaheuristic optimization tools combined for feature selection optimization, generalization of the model, and clinical interpretability [97]. Future research on some combination of the XAI frameworks and hybrid FSs should be carried out to help ensure clinical utilization of such methods [98]. Moreover, it remains an open research challenge to provide the scalability of quantum computing-based FS in hybrid frameworks and if it could indeed redefine feature selection in medical imaging as shown in [Fig. 2](#).

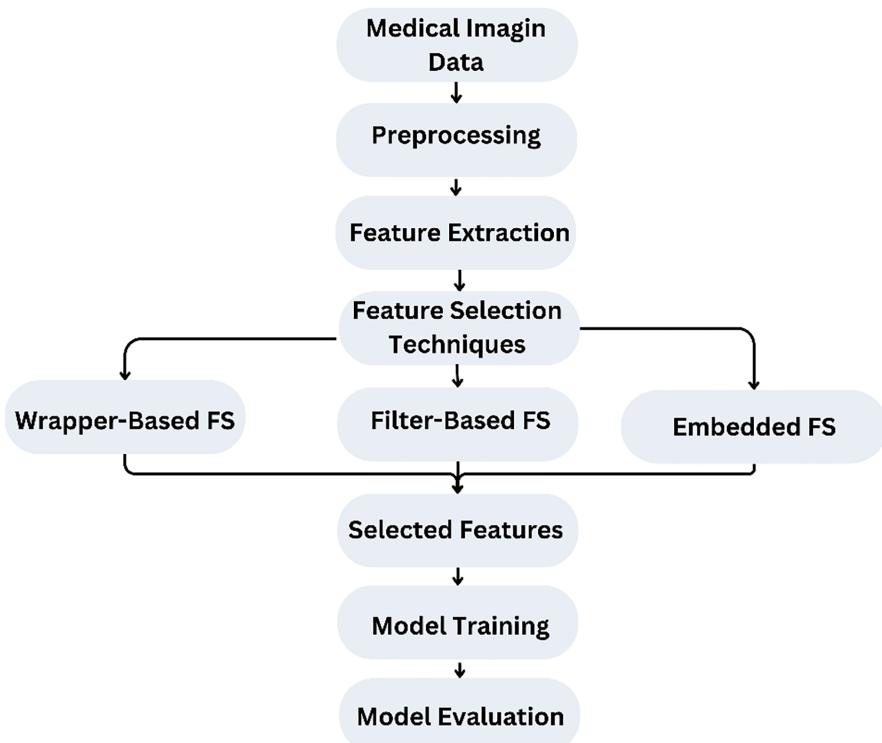


Figure 2: This figure illustrates the process of medical image data analysis, starting from preprocessing and feature extraction to feature selection techniques. It highlights three types of feature selection methods—Wrapper-Based, Filter-Based, and Embedded FS—which lead to selected features for model training and subsequent evaluation

Medical imaging is crucial in disease diagnosis, prognosis, and treatment planning. However, imaging data alone is seldom sufficient to base clinical decisions. Medical imaging, which consists of integrating different imaging modalities like MRI, CT, Positron PET, and US, offers an overall view of pathological conditions [99]. This will capture different anatomical and functional characteristics, e.g., MRI offers high soft tissue contrast, CT provides detailed bone structure, and PET provides metabolic activity [100]. Nevertheless, due to the large dimensionality, redundancy, and heterogeneity, when more than one data modality is integrated, new challenges arise in the FS of multi-modal data. Thus, it is essential to learn the most efficient FS techniques encompassing the most relevant features for each modality while maintaining compatibility [101].

To forcefully maximize the information extracted from multiple imaging sources but maintain collaborative diagnostic acumen, use needs to be made of multimodal feature selection. Feature redundancy is one of the significant challenges in multimodal imaging where the same or highly correlated features lie in different modalities. Concatenating features from different modalities provokes too many dimensions, hindering the performance of the machine learning models [102]. Also, other modalities may have various levels of noise, resolution, or feature scales, which makes direct integration difficult. For FS techniques to be adapted to the specific constraints of single and multimodal imaging, they must detect relevant features for the problem, filter out redundant or conflicting information, and optimize the feature space to minimize run time and preserve clinical interpretability [103].

Different fusion strategies are used to integrate multimodality features effectively. The fusion approaches determine how information should be combined from various modalities and how much these FS methods reduce dimensionality and increase model performance [104]. For instance, fusion techniques that incorporate anisotropic diffusion and cross bilateral filtering have shown promising results in retaining critical anatomical details while suppressing noise across multiple imaging modalities, thereby enhancing diagnostic accuracy [105]. In early fusion, raw features or the preprocessed (also called features) extracted using different modalities are concatenated before applying the FS techniques. This approach enables the deep learning models to learn cross-modal relationships from the beginning [106]. The three standard methods applied in early fusion for FS are PCA, Autoencoders, and Deep Embedding Networks. This method results in such a large dimension of the feature space that dimensionality reduction using Sparse Autoencoders or Variational Autoencoders (VAE) [107] is required to retain only the key features.

Early Fusion, in which features F_1, F_2, \dots, F_n . Different imaging modalities are concatenated into the vector of features. According to Eq. (13):

$$F_{\text{fso}} = f(F_1) \oplus f(F_2) \oplus \dots \oplus f(F_n) \quad (13)$$

where:

- $f(F_i)$ represents the feature transformation function for each modality.
- \oplus denotes the concatenation operation.

Classifier level fusion, or late fusion, refers to a fusion method where each modality is processed separately and decisions are fused later. FS is applied on each modality in isolation, and then the results are fused by an ensemble learning or voting mechanism [108]. Late fusion methods typically take advantage which have several trained classifiers over different feature sets followed by a final decision-making. Late fusion is still interpretable, but it may sacrifice cross-modal dependencies that may be helpful for more effective feature extraction [109]. In Late Fusion, features are initially classified independently, and the last prediction is a weighted sum of individual predictions [110]. As described in Eq. (14):

$$P_{\text{final}} = \sum_{i=1}^n w_i P_i \quad (14)$$

where:

- p_i is the prediction probability from the i -th modality.
- w_i is the weight assigned to each modality's classifier.

Furthermore, in some recent works, hybrid fusion has been proposed, combining early and late fusions. A common approach for fusion methods is to identify hybrid fusion, which tries to retain the multimodal relationship while optimizing the feature selection in terms of efficiency [111]. Advanced cross-modal learning techniques have been introduced to further improve FS in multimodal medical images. These techniques take advantage of the relations between these modalities to better retrieve shared and complementary features [112]. One of the most used approaches for multi-modal FS includes CNNs and Recurrent Neural Networks (RNNs). Both CNNs and RNNs (particularly Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU)) are effective at extracting spatial features from imaging data and sequential dependencies in imaging data, respectively, where it is of interest to learn off of another dimension (e.g., 4D cardiac MRI, fMRI). Cross-modal dependencies across different imaging modalities can be effectively modeled by combining CNN and RNN architectures, improving FS [113].

Another important emerging technique relevant to the domain of feature selection is feature disentanglement. This approach aims to decouple high-level structural information (e.g., anatomical features) from style-based or modality-specific variations (e.g., scanner noise, brightness), allowing only the structurally informative components to be retained for downstream tasks. Feature disentanglement contributes to robust model generalization and interpretability—two of the key goals of feature selection in clinical AI systems. It is commonly implemented through variants of autoencoders or adversarial networks that enforce latent space constraints to isolate distinct feature factors. For instance, disentangled representation learning has been applied in brain MRI analysis to separate pathological features from confounding imaging attributes, improving classification accuracy and reliability. This makes it a compelling complementary technique to FS, especially in multi-institutional or multi-modal imaging studies.

Transformer-based architectures for multi-modal imaging have recently been revolutionized. These self-attention mechanisms help ViTs and Multi Modal Transformers (MMTs) to model complex interaction patterns across different modalities [114]. Due to its capability to learn efficient global feature dependencies, Transformers can be used as an efficient feature selection method for tasks where such relations between different imaging sources must be considered. Vision Transformers have been well applied in the classification of brain tumors, the prediction of Alzheimer's, and the diagnosis of lung cancer; the performance is significantly improved through multiple modalities [115].

The emergence of Foundation Models—large-scale pretrained neural networks trained on massive and diverse datasets—has introduced a paradigm shift in feature extraction and selection. In medical imaging, models such as Vision Transformers (ViTs), CLIP, and Segment Anything Model (SAM) are increasingly adopted as universal feature encoders. These models generate rich, high-level representations that can be further refined through domain-specific feature selection techniques. Foundation Models reduce the need for extensive task-specific training and improve the quality of initial features by embedding semantic, spatial, and contextual priors. When coupled with classical or hybrid FS techniques, they enable better generalization and performance on downstream tasks such as tumor classification, organ segmentation, and disease progression modeling. Thus, Foundation Models serve not only as feature extractors but as enablers of scalable, efficient, and transfer-aware feature selection pipelines in medical imaging.

However, the success of many different FS techniques for multimodal medical imaging presents challenges. A lack of large-scale annotated multi-modal datasets, modality misalignment, and heterogeneous data distributions constrain further progress in this domain. Future research should focus on:

To improve clinical trust in the automated decision-making systems, Explainable AI (XAI) techniques must be developed for the multi-modal FS. When providing interpretable results, XAI can boost the healthcare professionals' confidence in AI-driven decisions by showing how multi-modal data (e.g., medical images and patient records) contribute to the feature selection process [116]. Meanwhile, privacy-preserving

multi-modal feature selection in distributed settings seems promising, and thus, it is investigated with Federated Learning (FL) frameworks for such scenarios in parallel. The approach of FL allows models to train on decentralized data while keeping patient privacy, which is essential to collaborative medical research, while ensuring data security [117]. Blockchain integration in imaging-based AI has emerged as a potential solution to secure data privacy, especially in neurology-related feature extraction pipelines [118]. Moreover, utilizing Quantum Computing with multi-modal FS opens fascinating applications. Processing large-scale imaging data in high dimensions is much more efficient using quantum algorithms, which can amplify this capability to enhance selecting features in complex medical data sets. One such integration of quantum computing can help provide more effective and scalable medical imaging and diagnostics [119]. For example, recent work has demonstrated the effectiveness of hybrid classical–quantum neural networks in enhancing Alzheimer’s disease detection from MRI scans. By combining classical deep learning (ResNet34) for initial feature extraction with quantum variational circuits for dimensionality reduction and classification, these models achieved notably higher accuracy than classical methods alone [120].

The field of multi-modal FS continues to evolve, and future advancements in hybrid AI models, interpretable learning techniques, and computationally efficient FS strategies are necessary to continue to develop its impact on real-world medical applications. Algorithm 2 has been included to demonstrate the feature selection approach utilized in the study.

Algorithm 2: Multi-modal feature selection using genetic algorithm

Input:

- Multi-modal feature set F = Error! Bookmark not defined.
- Population size P
- Crossover rate Cr
- Mutation rate Mr
- Max generations G

Output:

- Optimized feature subset F_{opt}

Steps:

1. Initialize a population of feature subsets randomly.
2. For each generation g in G :
 - Evaluate the fitness of each feature subset using a classifier (e.g., SVM, CNN).
 - Select the best feature subsets based on fitness scores.
 - Apply the crossover operation with probability.
 - Apply the mutation operation with probability Mr .
 - Update the population with new feature subsets.
3. **Return** the best feature subset F_{opt} .

Fig. 3 shows the multi-modal pipeline for feature selection.

Feature disentanglement has emerged as a critical approach in enhancing generalization and interpretability in medical image analysis. The goal is to separate meaningful anatomical or pathological features from modality-specific variations such as noise, texture inconsistency, or acquisition artifacts. Recent studies demonstrate that channel-level and depth-wise disentanglement strategies can significantly improve downstream tasks like segmentation by isolating domain-invariant representations. For example, Hu et al.

proposed a Contrastive Single Domain Generalization (CSDG) method that disentangles structure and style representations using channel-wise contrastive learning, enabling segmentation models to generalize better without requiring access to target-domain data [121]. Similarly, Wang and Ma introduced a depth disentanglement strategy that decouples local and global feature dependencies in a multi-stage latent space, facilitating fine-grained segmentation of small-scale anatomical structures [122]. These approaches highlight how disentangled representations—focused on structurally relevant details—can serve as effective precursors for robust feature selection. They also underscore the importance of combining spatial-channel attention and detail enhancement modules to isolate noise-free, diagnostic features in high-dimensional imaging data.

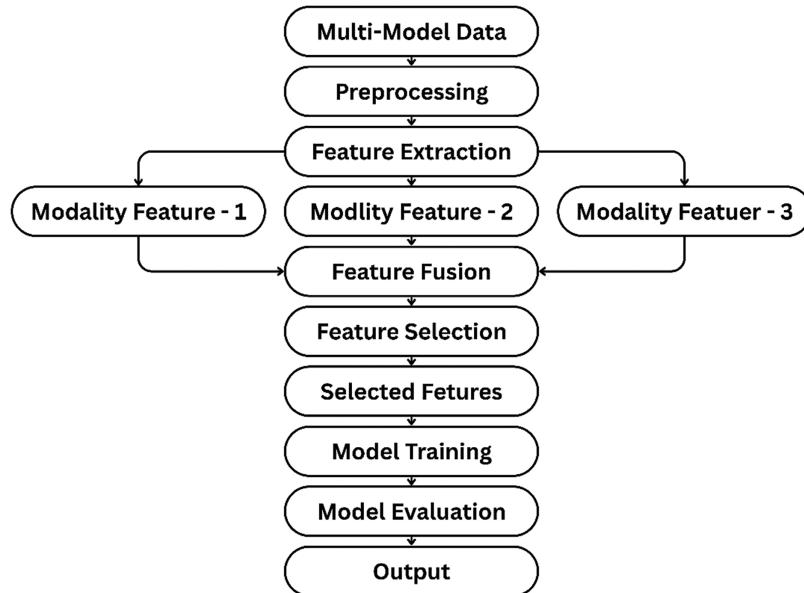


Figure 3: A multi-modal feature selection (FS) pipeline where features from multiple modalities, e.g., images, signals are extracted and fused. The fused features undergo selection, followed by model training, evaluation, and the generation of output predictions

The features of quantum computing generate opportunities to advance classical computing capabilities toward solving feature selection problems in large-scale datasets with numerous dimensions [123]. Feature selection processes using classical methods become unreliable when dealing with big databases since they struggle under high computational demands for handling big feature dimensions and complex relationships between features [124]. Quantum superposition, a fundamental principle of quantum mechanics that allows a quantum system to be in multiple states at the same time, combined with quantum entanglement, a phenomenon where the quantum states of two or more objects are linked together, provides the ability to process large datasets in parallel and improve the optimization of complex problems [125]. Feature selection through the classical method requires users to search various subsets within the feature space while evaluating them. The standard function for optimization appears as follows:

$$\text{Fitness}(F) = \text{Accuracy}(F, \text{Model}) \quad (15)$$

where F represents the selected feature subset, and “*Accuracy*” refers to the classification accuracy achieved by a model (e.g., SVM or CNN) trained on these features.

Various proposals have been made for feature selection using quantum-based algorithms. Quantum feature selection (QFS) offers several theoretical and practical advantages over classical FS methods,

particularly in the context of high-dimensional medical imaging data. First, quantum parallelism allows quantum algorithms to process multiple feature subsets simultaneously, significantly reducing the time required for subset evaluation. Second, entanglement facilitates complex correlations between features to be modeled more naturally, improving the relevance of selected features. Third, algorithms such as Grover's search provide quadratic speedups in exploring large feature spaces, which is especially valuable for real-time diagnostic systems. These capabilities enable QFS to address the scalability and combinatorial complexity challenges that often hinder classical FS approaches in medical imaging applications.

Quantum Genetic Algorithms (QGA) perform global searches more efficiently using quantum computing, giving them the advantage of better exploration of the feature space [126]. The fitness function for a QGA is similar to the fitness function used in classical genetic algorithms. Still, it is expanded with the quantum operations used for better exploration capability [127]. The quantum k-Nearest Neighbors (k-NN) leverages quantum properties to accelerate distance computing and achieves better classification accuracy than the classical k-NN algorithm [128]. The criterion for the quantum k-NN distance function is the Euclidean distance, but quantum parallelism is utilized for faster computation. The relationship in Equation show in (16):

$$d(F_1, F_2) = \sum_{i=1}^n (F_1 - F_2)^2 \quad (16)$$

where F_1 and F_2 and n is several features, and they are feature vectors. As an alternative to the quantum-enhanced version, quantum gates were used to calculate the distance in parallel. As depicted in Eq. (17):

$$\min_{w, b} \frac{1}{2} |w|^2 \quad (17)$$

Subject to

$$y_i (w^T x_i + b) \geq 1, \forall i = 1, \dots, n \quad (18)$$

Furthermore, the Quantum SVM provides the integration of quantum mechanics to speed up the optimization process, which is essential in high-dimensional feature space. One problem in SVM is optimizing the margin when the support vectors in between are maximized [129]. The classical SVM optimization problem is Quantum SVM, a quantum algorithm that uses quantum algorithms to obtain the minimum of this objective function more efficiently, especially for high-dimensional data [130]. One significant benefit of quantum FS over classical approaches is that it achieves substantial speedup when the dimensionality of datasets is large. Parallelism in quantum algorithms means that they can process several features in parallel, while the dimensionality of the data usually constrains classical methods [131]. It results in better searching over the feature space and, hence, better time complexity. Usually, the classical feature selection involves search over a feature space of size. N and requires $O(N)$ steps.

However, with Grover's search algorithm, we can search for such selected features in quantum computers with a quadratic speedup, $\sqrt{O(N)}$:

Classical search complexity: $O(N)$, Quantum search complexity: $O(\sqrt{N})$

In addition, feature interactions are generally easy to compute on a quantum computer and, as a result, can be represented and selected more efficiently than a classical computer. Quantum feature selection, therefore, provides the opportunity to reduce the computational time significantly and even to achieve better machine learning models' overall performance when compared to classical models, when the domain of

interest is medical imaging, genomics, or any other use case where data sets are massive. [Fig. 4](#) illustrates the conceptual diagram of quantum feature selection.

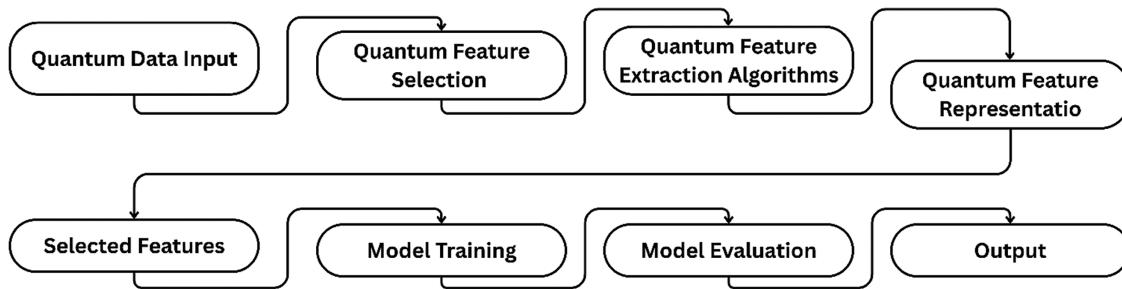


Figure 4: Conceptual diagram of quantum feature selection (FS). It shows the process starting from quantum data input, followed by quantum feature selection, extraction algorithms, and representation. The selected features are then used for model training and evaluation, ultimately leading to the output of the model

3 Methodology (Systematic Review Process)

The proposed study discusses a systematic review methodology we followed to comprehensively analyze the role of FS in medical imaging in carrying out classical, deep learning based, hybrid, and quantum FS techniques. This review uses a methodology that guarantees a structured, transparent, and reproducible approach for recognizing, choosing, and refining the sought research articles.

3.1 Search Strategy

A structured search strategy was designed to ensure the acquisition of comprehensive coverage of FS techniques in medical imaging. Peer-reviewed articles containing relevant information were searched from multiple high-impact databases.

3.1.1 Data Retrieval and Keywords Used

Therefore, an approach to FS techniques in medical imaging is required to achieve a well-defined and systematic search strategy. Various keyword combinations were chosen to cover multiple FS methodologies and their applications over various medical imaging modalities. All the keywords, including “Feature selection in medical imaging,” “Deep learning feature selection for medical images,” “Hybrid feature selection in healthcare,” and “Quantum feature selection for medical imaging,” were searched to widen possible areas of research.

Boolean operators (AND, OR) were further utilized to help refine the search and have greater precision in filtering out relevant literature. For example, the query examples that were used (“Feature Selection” AND “Medical Imaging” AND “Deep Learning”) guaranteed that such studies encompassed the utilization of AI-based FS methods for diagnostic imaging. Moreover, research that combines classical, metaheuristic, and deep learning based FS methods over different imaging modalities can be captured by queries such as (“Hybrid Feature Selection” OR “Metaheuristic FS” AND “MRI” OR “CT” OR “PET”). Additionally, grants were made to explore at the forefront of technologies by searching for studies that applied quantum computing to solve high-dimensional medical imaging data with the search string (“Quantum Feature Selection” AND “High-dimensional Imaging Data”).

By utilizing this structured search approach, high-quality and relevant studies that detail the evolution of FS techniques, their effectiveness in medical imaging, and their future developability could also be extracted as illustrated in [Fig. 5](#).

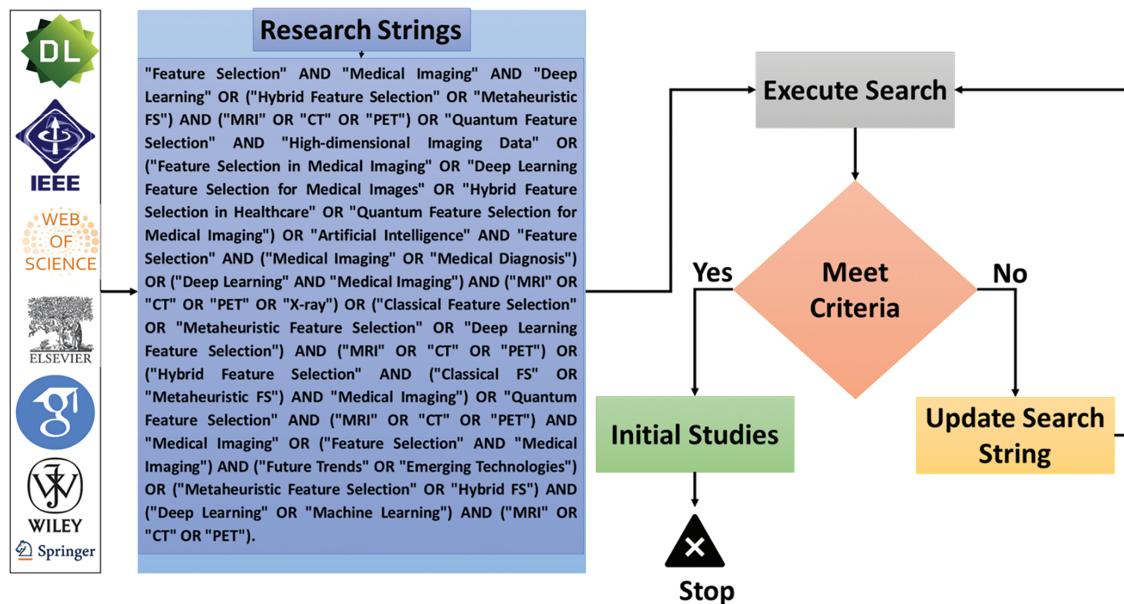


Figure 5: Process of executing a search string in a research database. The flowchart starts with predefined research strings, which include keywords related to feature selection, medical imaging, deep learning, and various technologies. The search is executed, and if the criteria are met, the initial studies are retrieved. If not, the search string is updated accordingly to refine the search, continuing the process until the relevant results are found

3.1.2 Databases Searched

Therefore, a systematic search was performed across well-established electronic databases to ensure a complete and well-structured review of FS techniques in medical imaging. A list of databases was developed and selected with a significant amount of care, including work with biomedical applications, computational support, and AI-based innovations in FS methodologies. As PubMed offers comprehensive coverage of biomedical and clinical imaging research, all the FS techniques applied in the real healthcare setting will be included. A search of IEEE Xplore was used to locate studies related to FS computational and engineering aspects related to AI and machine learning based approaches. At the same time, ScienceDirect offered a vast knowledge base on machine learning, deep learning, and AI-based FS methods used in diagnostics and prediction modeling tasks.

Springer was also used to study research on advanced AI methods in FS, specifically to investigate hybrid and deep learning-based approaches. MDPI was chosen as one of the best focuses on recent advances in medical AI, including FS applications in various imaging modalities like MRI, CT, PET, and ultrasound. Lastly, arXiv was used to investigate new FS methodologies that utilize quantum computing to deal with high-dimensional medical imaging data.

The search was limited to studies published between 2015 and 2025 to keep the information current and reliable and to include the latest FS techniques. This time frame was selected for examining the most modern trends and technological novelties and measuring the latest breakthroughs in the FS methodologies, including the classical statistical methods, the deep learning approaches, hybrid models, and such techniques as quantum-based strategies.

3.2 Inclusion and Exclusion Criteria

Predefined inclusion and exclusion criteria were applied to maintain the quality and relevance of selected studies.

3.2.1 Inclusion Criteria

- Studies focusing on FS techniques applied to major medical imaging modalities (MRI, CT, PET, Ultrasound, ECG).
- Publications between 2015 and 2025 to ensure the inclusion of the latest advancements.
- Peer-reviewed journal and conference papers from reputed sources.

3.2.2 Exclusion Criteria

- Non-English studies (to maintain consistency and accessibility).
- Papers focusing only on general AI-based medical imaging without FS techniques.

This rigorous selection process helped filter out studies that lacked experimental validation or did not focus on FS techniques in medical imaging.

3.3 Data Extraction & Quality Assessment

To ensure high quality and completeness of the review, the evaluation of key aspects regarding how to define the effectiveness and applicability of FS techniques in medical imaging was performed systematically. For each selected study, crucial details were extracted, including the FS technique used, whether classical, deep learning, hybrid, or quantum-based. Information about the datasets used in the studies was also collected in case some studies covered public datasets like BraTS and CheXpert. In contrast, others might have used custom dataset tools specific to their medical imaging task. The reported performance metrics also constitute one of the essential aspects considered; these include standard evaluation measures such as accuracy, AUC ROC, F1 score, and computational efficiency, enabling the comparison between several FS techniques' effectiveness.

Further, rigorous quality assessment was performed for the studies included to ascertain their reliability and validity. This assessment was completed with several key criteria. Given priority to reproducibility: the studies had to have publicly available code, datasets, or detailed methodological descriptions. Moreover, the robustness of the FS techniques was another critical factor to include in the selection of studies, which provided comparative performance evaluations instead of isolated case studies. Finally, the test was conducted to assess the real-world applicability of FS methods by considering the studies to prove the practical implementation of FS in clinical and diagnostic imaging.

[Table 3](#) summarizes and tabulates a comparative overview of the selected studies to provide a structured summary of the reviewed literature. It categorizes them according to applied FS techniques, analyzed datasets, and reported performance metrics. The representation is structured, concise, and informative of advancements and trends in FS for medical imaging and [Fig. 6](#) illustrates the complete process from search to selected studies.

Table 3: Summary of selected studies based on FS techniques, datasets, and performance metrics

Feature selection technique	Imaging modality	Dataset used	Performance metrics
Filter-based FS (Chi-square, Mutual Info)	MRI	BraTS	Accuracy, AUC-ROC

(Continued)

Table 3 (continued)

Feature selection technique	Imaging modality	Dataset used	Performance metrics
CNN-based FS	X-ray	CheXpert	Sensitivity, Specificity
Hybrid FS (PCA + CNN)	PET	Custom dataset	F1-Score, Computational time
Quantum FS (QAOA)	Multi-modal (MRI + CT)	Private dataset	Feature reduction ratio, Model efficiency

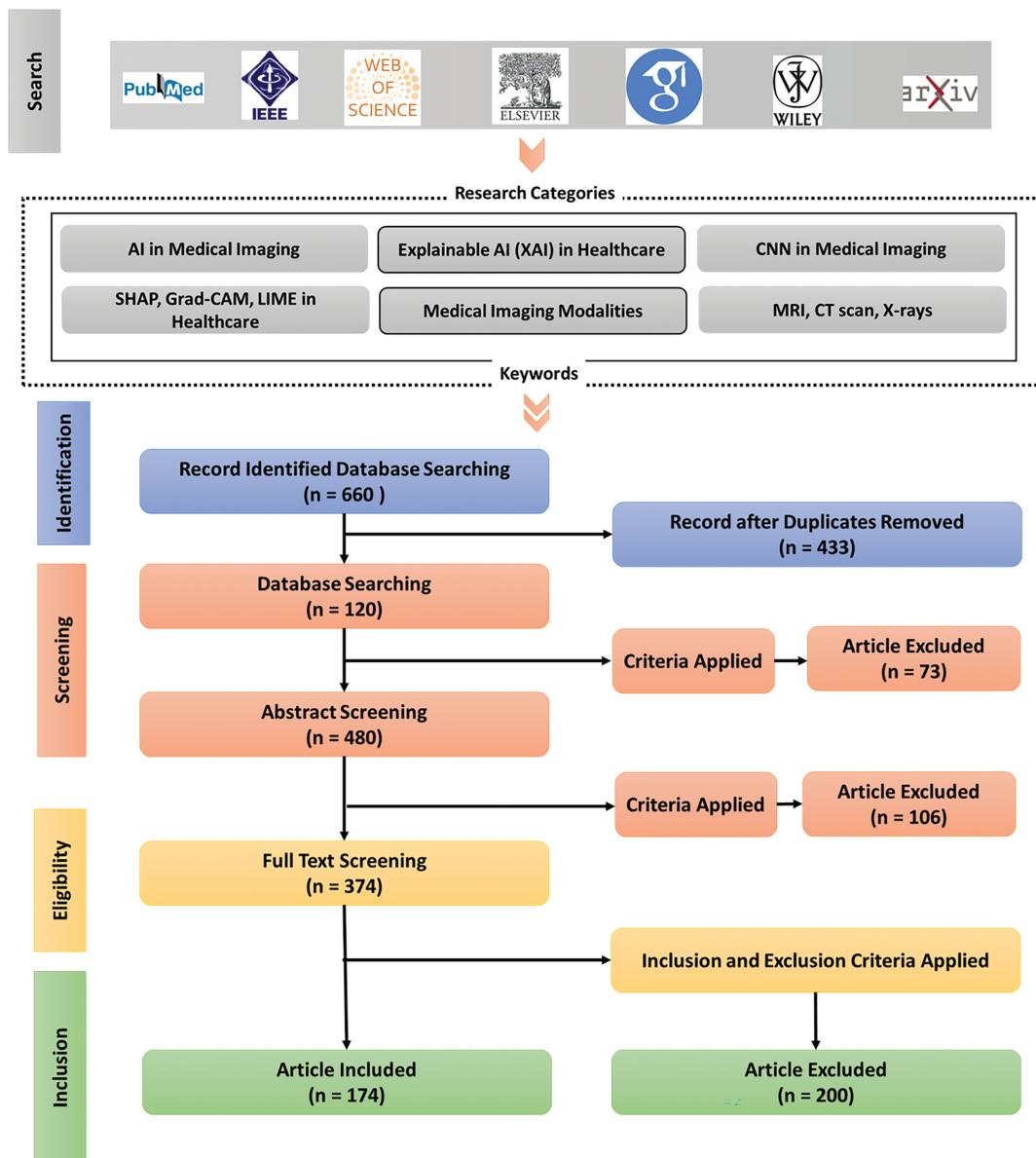


Figure 6: PRISMA flowchart, illustrating the systematic review process. It starts with database searching, where 660 records are identified using specific keywords. After removing duplicates, 433 records remain, followed by abstract screening and full-text screening. Criteria are applied at each stage, resulting in 174 articles included for the final review, while others are excluded based on inclusion and exclusion criteria.

Database searching identified a total of 660 records. These records were 433 after being deduplicated. The screening that was done to remove the abstracts ($n = 480$, with extra cases of 47 unique records) removed 73 articles that fail the inclusion criteria. Of 374 articles identified through full-text screening there were 106 excluded, which meant that 174 of the articles were to be included in the eventual synthesis. [Fig. 6](#) and [Table 4](#) give a breakdown of the 200 full-text articles deemed ineligible during eligibility screening.

Table 4: Summary of selected studies

Stage	Action	n
Records identified	Initial database search	660
Duplicates removed	–	227
Records after deduplication	–	433
Additional records searched	Manual/database	47
Total for abstract screening	–	480
Abstracts excluded	Failed initial screening	73
Full-text assessed	Eligibility review	374
Full-text excluded	Did not meet inclusion criteria	200
Final articles included	–	174

4 Challenges, Open Problems, and Future Directions

A lot of progress has been going in developing the existing FS techniques for the application in medical image, but there are some critical challenges and open research problems still to be solved. Nevertheless, deep learning and hybrid FS methods face their own limitation of high computational complexity and scalability to wide ranges of problems making them hard to practice in the real-world applications. Moreover, combining multi modal imaging data is challenging in terms of the redundant features of multiple modalities, the alignment of different modalities, and cross modal learning. Theoretically, FS approaches using quantum computing are quick, but have been held to be impractical on account of hardware and algorithm shortcomings. To develop such FS+DB models, one needs to consider these FS challenges with a multidisciplinary confluence, to have FS models that are efficient, interpretable, and scalable, and seamlessly combined with the clinical workflow. This section explores key challenges, open problems, and potential solutions that will define next generation AI driven medical imaging systems.

4.1 Computational Complexity & Scalability Issues

Although the FS is computationally demanding and a scalability challenge to incorporate in medical imaging, the FS is hugely complex; Challenges exist between FS field and deep learning and hybrid approaches, while great opportunity space exists for innovation and improving. This allows progress towards the direction of the optimization of this integration process and enhancement of all our solutions. Due to their nature, medical imaging datasets are naturally high-dimensional and expensive to process, train, and even make inferences from. Deep FS methods based on CNNs, Autoencoders, and ViTs enjoy better feature extraction performance. However, their feature extraction process tends to be heavy in memory usage and consumes enormous energy and training time. These models come at a high computational burden, making it difficult to use them in real-time medical applications that demand happening fast and efficiently [\[132\]](#). Additionally, because the computational cost is based on feature fusion strategies, the problem of using multimodal medical imaging data. Deep learning Mechanisms of attention Deep learning can improve multimodal picture fusion by considering fusion algorithms, modalities, and metrics. Researchers can use a graphical

taxonomy and difficulties and areas of future research to better understand multimodal fusion problems and choose appropriate approaches [133].

However, one of the significant challenges for deep learning-based FS techniques is that the computational cost of those methods is very high because many parameters are used in deep architectures [134]. For example, the mathematical representation of the complexity of the based FS model can be represented as the following in Eq. (19):

$$C_{FS} = \sum_{l=1}^L O(f_l \cdot k_l^2 \cdot c_l \cdot h_l \cdot w_l) \quad (19)$$

where L denotes the number of layers, f_l represents the number of filters in the l -th layer, k_l^2 corresponds to the kernel size, c_l denotes the number of input channels, and h_l, w_l signify the height and width of the feature map, respectively. The development speed of computational needs escalates exponentially due to longer network depths, thus making standard computing inadequate unless accompanied by high-performance GPUs or TPUs. Excessive memory needs and massive storage demands intensify computational obstacles. Intermediate feature maps generated from the deep learning models are large, thus resulting in enormous memory consumption. This is exacerbated in real-time medical applications where inference speed is the most critical parameter [135].

FS techniques struggle with scale-up potential in situations involving multi-modal imaging data analysis. Combining machine learning with multimodal fusion strategies in human activity recognition using a state-of-the-art convolutional network architecture and a large dataset. The results show a clear performance improvement from multi-modal fusion and a substantial advantage from an early fusion strategy [136]. The processing requirements of multi-modal feature selection rise significantly because of differences between modalities and domain alignment requirements, which causes existing deep learning-based FS methods to be impractical during large-scale clinical deployment [137].

Researchers have explored solutions, such as lightweight deep learning models, federated learning, and quantum computing, to overcome these computational limitations. Implementing the MobileNets and EfficientNets lightweight CNN architectures resulted in substantial parameter reduction while maintaining the performance level of FS. Depth-wise separable convolutions strengthen the computational efficiency of a lightweight CNN-based FS approach, which is shown mathematically in Eq. (20):

$$C_{DW} = O\left(\sum_{l=1}^L k_l^2 \cdot c_l \cdot h_l \cdot w_l + f_l \cdot c_l \cdot h_l \cdot w_l\right) \quad (20)$$

The optimized formulation eliminates calculation redundancy and focuses on the computational path's most crucial feature extraction layers. As such, lightweight models represent a good tradeoff of computation demand vs FS accuracy and are better suited to real-time medical imaging applications.

Moreover, federated learning is another emerging approach to alleviating computational bottlenecks in FS by providing a decentralized learning paradigm for distributed model training across various institutions. Unlike transferring the entire medical imaging dataset to a central server, FS is done locally on distributed nodes in federated learning. This reduces the computational and communication costs while keeping the patient private [138]. In particular, this approach is most useful in healthcare applications where data sharing restrictions must be considered and HIPAA and GDPR compliance is of the essence. Besides achieving scalability, Federated FS models further enable learning collectively over distinct medical centers in a way that reduces the generalizability burden at the cost of computing [139].

Besides classical and deep learning-based FS techniques, quantum computing has recently become a new hope for high-dimensional feature selection. Quantum FS techniques take advantage of quantum parallelism to process large-scale medical imaging datasets quickly [140]. Quantum Variational Circuits (QVCs) allow quantum-based FS to be mathematically formulated in Eq. (21).

$$|\Psi\rangle = U_\theta |0\rangle^{\otimes n} \quad (21)$$

where $|\Psi\rangle$ represents the quantum state encoding feature selection, U_θ denotes the unitary transformation optimized for feature selection, and $|0\rangle^{\otimes n}$ signifies the initial quantum state for n Qubits. On the other hand, classical FS methods suffer from exponential growth in complexity due to a possible exponential increase in the data. Quantum annealers for feature selection in light-weight medical imaging collections. Linear Ising penalties with subsampling and thresholding increase stimulation in simplified use situations, with decent results but uncertain future applicability due to hardware limits [141]. However, quantum FS is still in its infancy and is hindered by practical hardware issues, noise sensitivity, and the fact that the current quantum processors do not scale to acceptable sizes.

Fig. 7 depicts the computation efficiency of classical, deep learning-based, and quantum FS candidates, showing the impact of different FS strategies on computational complexity.

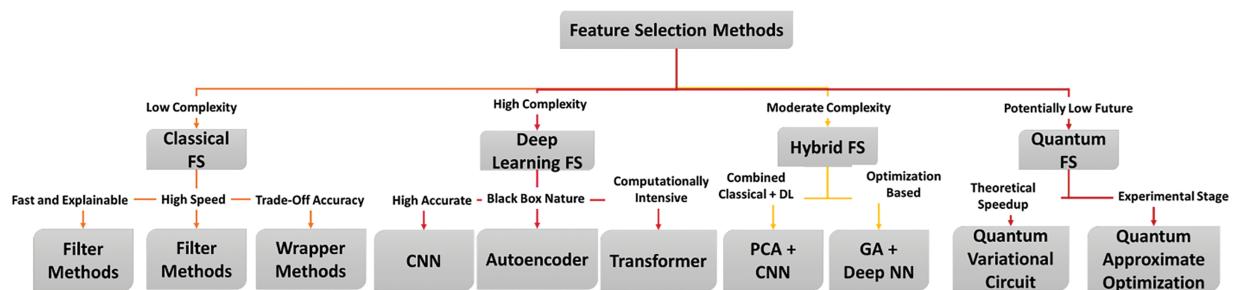


Figure 7: Feature selection methods based on computational efficiency. Classical FS methods are fast and simple but may sacrifice accuracy, while deep learning FS methods offer higher accuracy at the cost of increased complexity. Hybrid FS methods combine both approaches, and quantum FS, still in an experimental stage, promises potential future speedups

This shows that although classical FS techniques are efficient in computation, they are not as capable as the DF extraction of advanced methods. However, for DL-based FS methods, though effective, the computational scalability becomes an issue for their deployment. The hybrid FS models achieve the best tradeoff between good features of classical and deep learning. Federated learning increases scalability, but in terms of privacy, it inherits the complexity of the distributed networks. However, quantum FS is theoretically promising but limited in its practical implementation and current hardware limitations.

Significant challenges to overcome deal with the computational complexity and scalability for FS to reach its full potential in medical imaging. These are very powerful when it comes to neural network-based FS methods, which are also very expensive to compute and, in such cases, require specially optimized techniques such as lightweight architectures, federated learning, or even hybrid methods for reducing computational costs. Furthermore, quantum computing is anticipated to be highly useful in high-dimensional feature selection, but more studies are needed to determine its practical adoption. We will then see interdisciplinary synergies between AI, distributed computing, and quantum technologies endeavouring to improve the computational efficiency and scalability of FS methods to be helpful for the clinic.

4.2 Interpretability & Explainability in FS

Medicine has begun to revolutionize diagnostic accuracy and efficiency using the recent development of AI-based FS approaches. Nevertheless, in the current situation, all these AI-driven FS techniques usually perform well in indicating diagnosis potential. Yet, they are not transparent in clinical practice, restricting their usage in real medical applications [142]. Deep learning models' inherent complexity and opaqueness make them unusable for most healthcare professionals as they are not interpretable, causing them to lack trust and acceptance. In particular, this is a critical issue in healthcare, as incorrect patient care and outcomes depend directly on decisions made using AI models. However, the bridge of this gap has been constituted as a key research area to integrate interpretability and explainability into the FS models [143].

To address the interpretability challenge in AI-based FS, Explainable AI (XAI), SHAP (Shapley Additive Explanations), Grad-CAM (Gradient-weighted Class Activation Mapping), Hybrid FS, etc., have been proposed in the literature. XAI frameworks are aimed to help make the decision process of machine learning models understandable to humans. SHAP is one of the popular methods within XAI that assigns a value for each feature based on its contribution to the model's prediction [144]. The equation to calculate the SHAP value in Eq. (22):

$$\varphi_i(f) = \sum_{S \subseteq N} \frac{|N|!}{|S|! (|N| - |S| - 1)!} [f(S \cup \{i\}) - f(S)] \quad (22)$$

where $\varphi_i(f)$ represents the SHAP value for feature i , and $f(S)$ is the model prediction with a subset S of features. This allows quantification of each feature that affects the final output of the model, addressing prioritization for clinicians to understand the most influential variables and making the decision process transparent.

Besides being a key tool for improving model interpretability, another essential tool is Grad-CAM, especially in medical imaging tasks. Grad-CAM gives an intuitive means to visualize the areas of the image that most influence a CNN's prediction (i.e., those with the strongest and highest activation) [145]. The equation of Grad-CAM in Eq. (23):

$$\text{Grad-CAM}(x) = \text{ReLU} \left(\sum_k \alpha_k A_k \right) \quad (23)$$

In this equation, A_k denotes the feature map from the convolutional layer and is a_k . Each feature map is assigned a gradient-based weight. The weight produces a heatmap showing which regions in the image are key to the model's decision, so clinicians can visually see what parts of an image caused the model to make its decision in Eq. (24):

$$\alpha_k = \sum_{i=1}^Z \sum_j \frac{\partial A_{kij}}{\partial y_c} \quad (24)$$

Here, y_c is the class score for category c , and Z is the total number of pixels in the feature map. This method is beneficial in CNN-based FS models for medical imaging as it allows clinicians to interpret why certain features were chosen in a deep learning framework.

Improving the interpretability of AI-based FS Hybrid FS techniques, which are combinations of classical feature selection methods and deep learning models, also has a specific role in improving the interpretability of AI-based FS. For instance, the dimensionality of high-dimensional medical images can be reduced with

Principal Component Analysis (PCA), and features can be found using CNN. The PCA in Eq. (25) shows as:

$$X_{\text{PCA}} = XW \quad (25)$$

The transformed dataset X is generated from the original dataset X by multiplication with the second matrix W of eigenvectors (principal components). By this hybrid approach, we bridge between regular statistical features and recent deep learning features to select such features efficiently and interpretably. Additionally, given that there is a Hybrid FS, genetic algorithms (GA) can be utilized to achieve such optimization by modeling the natural evolutionary processes. Therefore, the GA fitness function is defined in Eq. (26) as:

$$f(S) = \sum_{i=1}^N \text{accuracy}(S_i) \quad (26)$$

where $f(S)$ is the fitness score for the selected feature subset S , and $\text{accuracy}(S_i)$ represents the performance of a model trained on the feature subset S_i . This method can determine the most critical features, improve the model's interpretability, and retain its high accuracy.

In the end, for AI-based FS methods to be clinically applied, they must also be capable of being interpreted and explained. The SHAP, Grad-CAM, and Hybrid FS approaches, along with XAI, provide promising ways of improving transparency. This will help clinicians know when to trust and use the AI model appropriately in decision-making. By combining these two techniques, an AI-driven FS system can be made both high-performance and clinically viable, leading the way for their wider adoption in healthcare.

Table 5 presents the advantages and limitations of each method.

Table 5: Comparison of interpretability methods in FS

Method	Description	Advantages	Limitations
XAI [146]	AI techniques for making decisions more transparent.	Improves interpretability and provides clear explanations.	It can be computationally expensive and may not fully explain models.
SHAP [147]	Assign importance scores to features.	Precise feature importance works across models.	Computationally intensive, lacks visual representation.
Grad-CAM [148]	Highlights important image regions in CNN decisions.	Provides visual explanations.	Limited to CNNs, only works on images.
Hybrid FS [149]	Combines classical FS with deep learning.	Balances accuracy and interpretability.	It may not continually optimize performance.

Integrating preprocessing and spatially aware decomposition techniques for feature selection has become increasingly important in handling high-dimensional and complex datasets. Preprocessing methods such as data decomposition, spatial interpolation, and dimensionality reduction have proven effective in managing data nonlinearity and spatial dependencies, particularly in domains such as air quality prediction and medical imaging [150]. Advanced feature selection strategies now incorporate multi-criteria evaluation—e.g., minimizing redundancy while maximizing discriminative power—to enhance both interpretability and accuracy [151]. Spatially aware methods, such as relation-aware wrapper techniques, utilize

graph- or tree-based structures to model relationships among features and samples, yielding more relevant and stable feature subsets [152]. These approaches have demonstrated success in medical applications, including breast cancer segmentation, where context-aware spatial decomposition has been used to improve the relevance of extracted features. Furthermore, synergistic frameworks that combine feature selection with distributed classification mechanisms address the computational and heterogeneity challenges posed by large-scale medical data, enabling scalable and efficient model training [153,154]. Hybrid and metaheuristic algorithms—particularly those employing particle swarm optimization or evolutionary multitasking—also contribute by enabling robust search across large feature spaces while facilitating knowledge transfer between related tasks [150,151,155]. These integrated strategies collectively improve classification performance, reduce model complexity, and support scalable analysis pipelines for both centralized and distributed medical imaging scenarios. Despite these advances, ongoing research is needed to balance computational efficiency, interpretability, and robustness to heterogeneous or streaming data, which remain open challenges in modern FS applications.

To complement the detailed discussion of feature selection techniques, we provide an additional comparative summary of representative state-of-the-art works in [Table 6](#). This table highlights the diversity of FS approaches across modalities and their associated evaluation metrics, offering a concise overview of recent advances in the field.

Table 6: Comparative summary of additional state-of-the-art FS works in medical imaging

Ref.	FS methodology	Imaging modality	Dataset used	Performance metrics	Key contribution
[53]	CNN + Grad-CAM	MRI (Prostate)	Private dataset	Accuracy, Visual relevance	Visual interpretability via region localization
[66]	Autoencoder (AE/VAE)	PET, MRI	ADNI	Accuracy, Sensitivity	Biomarker extraction for Alzheimer's detection
[43]	LASSO Regression	MRI	ADNI	AUC, Sensitivity	Sparse FS for Alzheimer's progression
[120]	Hybrid (ResNet34 + QVC)	MRI	Alzheimer's dataset	Accuracy, Feature reduction	Hybrid classical-quantum FS for brain imaging
[86]	GA + DNN	MRI, CT (Multimodal)	Custom dataset	Accuracy, F1-score	Metaheuristic optimization in hybrid FS

4.3 Multi-Modal Feature Selection Challenges

There is great potential in integrating multi-modal data in medical imaging with EEG, MRI, Ultrasound, and other standard medical imaging modalities such as MRI, CT, PET, etc., to make more accurate diagnoses and better inform clinical decision making. However, even with the multi-modal FS amenable to many challenging problems, challenges related to the use of such issues in real-world scenarios need to be addressed [156]. Adaptive feature extraction Deep learning may learn adaptive feature extraction by learning detailed patterns in varied datasets. Multimodal deep learning incorporates other sensory modalities because it is hard to extract useful information from unstructured data [157]. As a result, the two modalities represented by these features often have different resolutions, sampling rates, and structural representations, making them difficult to combine because they may not align with each other across modalities [158].

This challenge becomes more pronounced in multi-modal biomedical data, where visual (e.g., MRI) and non-visual (e.g., lab tests, genomics) data require distinct preprocessing pipelines and semantic alignment. To address this, recent work has focused on deep learning-based multimodal fusion strategies, offering principled frameworks for reconciling heterogeneous biomedical data [159].

Another critical challenge arises from incomplete or missing modalities, which can occur due to technical limitations, motion artifacts, or modality-specific acquisition constraints. When feature selection (FS) algorithms operate under such conditions, they risk introducing biases or disproportionate weighting of partial data, ultimately degrading model performance.

As a result, some solutions to solve these problems are suggested by enhancing integration and selection of multimodal dataset features. Graph Neural Networks (GNNs), Bayesian Fusion, and Self-Suppressed Learning are promising techniques to handle different resolutions of the modalities and missing modality in multimodal FS scenarios.

This work demonstrates that with a powerful tool, GNNs, such relationships in data can be effectively modeled, and projects such as modeling relationships in multimodal medical images have shown GNNs to be powerful. GNNs are suitable for multimodal data because each modality can be considered a node in the graph, and the relation between different features across the modalities can be represented as the edge within these nodes. In general, GNNs are applying for learning the representation of nodes for capturing the local and global dependencies by aggregating the information from the neighbor nodes. This is the equation corresponding to graph convolution operation in GNN shows in Eq. (27):

$$h_v^{(k+1)} = \sigma \left(\sum_{u \in N(v)} W^{(k)} h_u^{(k)} + b^{(k)} \right) \quad (27)$$

where:

- $h_v^{(k+1)}$ is the updated feature representation for node v at the $k + 1$ -th iteration,
- $N(v)$ represents the neighbors of node v ,
- $W^{(k)}$ is the weight matrix for the k -th layer,
- $b^{(k)}$ is the bias term, and
- σ is an activation function (e.g., ReLU).

Considering the multi-modal FS, GNNs can map to combine multi-modal features over which modalities vary. Dependencies are captured between the features, even when resolutions are different or modalities are missing. GNNs lend themselves better to fusing the information more effectively, which is very important when dealing with complex medical image analysis tasks.

Bayesian Fusion is another approach to dealing with the resolution and missing modality challenge in multi-modal FS. This approach considers feature selection as a probabilistic model and computes the probability of various feature subsets based on the available data on all the modalities. Bayesian fusion helps combine data uncertainly and variably between modalities. The general formulation of such Bayesian fusion model is shown in Eq. (28):

$$p(F|M) = \frac{p(M)p(M|F)}{p(F)} \quad (28)$$

where:

- $p(F|M)$ is the posterior probability of the feature set F given the modality set M ,
- $p(M)p(M|F)$ is the likelihood of observing the modalities M given the features F ,

- $p(F)$ is the prior probability of the feature set,
- $p(M)$ is the marginal likelihood of the modalities.

This method uses Bayesian inference to consider the uncertainty in the data caused by missing modalities and different resolutions, and hence, it provides more precise feature selection in multi-modal data sets.

This problem can also be addressed using Self-Supervised Learning (SSL). In contrast, SSL techniques present little burden on the availability of labeled data and utilize the structure of the data itself to produce meaningful representations of input features. SSL can learn features from the available modality even if some modality is missing or incomplete by exploiting the complex and inherently existing relationships among the data. In SSL, the general objective of the function is to minimize the reconstruction error between the original and predicted data, which can be stated in Eq. (29) as:

$$L_{SSL} = \sum_i |x_i - \hat{x}_i|^2 \quad (29)$$

where:

- x_i is the original data point (e.g., a modality),
- \hat{x}_i is the predicted data point after applying the SSL model,
- $|.|^2$ represents the squared Euclidean distance (or some other appropriate loss function).

In the sense of multi-modal FS, SSL could be employed to extract representation from incomplete data, whereas it infers missing information from one modality using the other available media. The main strength of this approach is that when training data is scarce and imaging modalities are unavailable.

In short, solving the problem of resolving different resolutions and missing modalities in multi-modal feature selection can be facilitated by techniques like GNNs, Bayesian Fusion, and Self-Supervised Learning. Bayesian Fusion modulates data fusion and Uncertainty; GNNs can capture the relationship between different modalities; Self-Supervised Learning can generate robust feature representation even without entirely being given information. These solutions improve the effectiveness of multi-modal FS methods so that they can deliver more accurate and reliable diagnostic tools in the presence of real medical imaging data complexities. Fig. 8 illustrates the multimodal features selection with techniques.

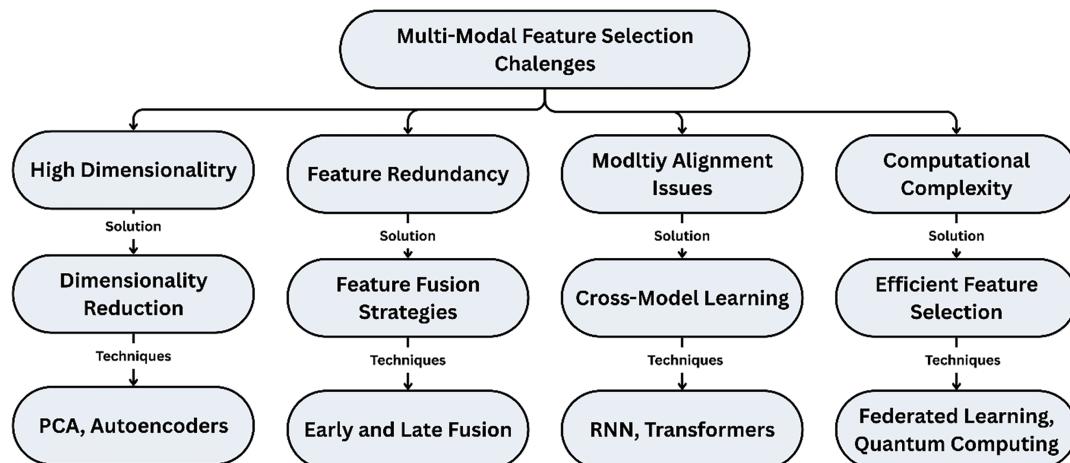


Figure 8: Key challenges in multi-modal feature selection, including high dimensionality, feature redundancy, modality alignment issues, and computational complexity. It also presents corresponding solutions and techniques, such as dimensionality reduction, feature fusion strategies, cross-model learning, and efficient feature selection methods

To facilitate a comprehensive understanding, [Fig. 9](#) presents a taxonomy diagram summarizing the main categories of feature selection techniques covered in this review.

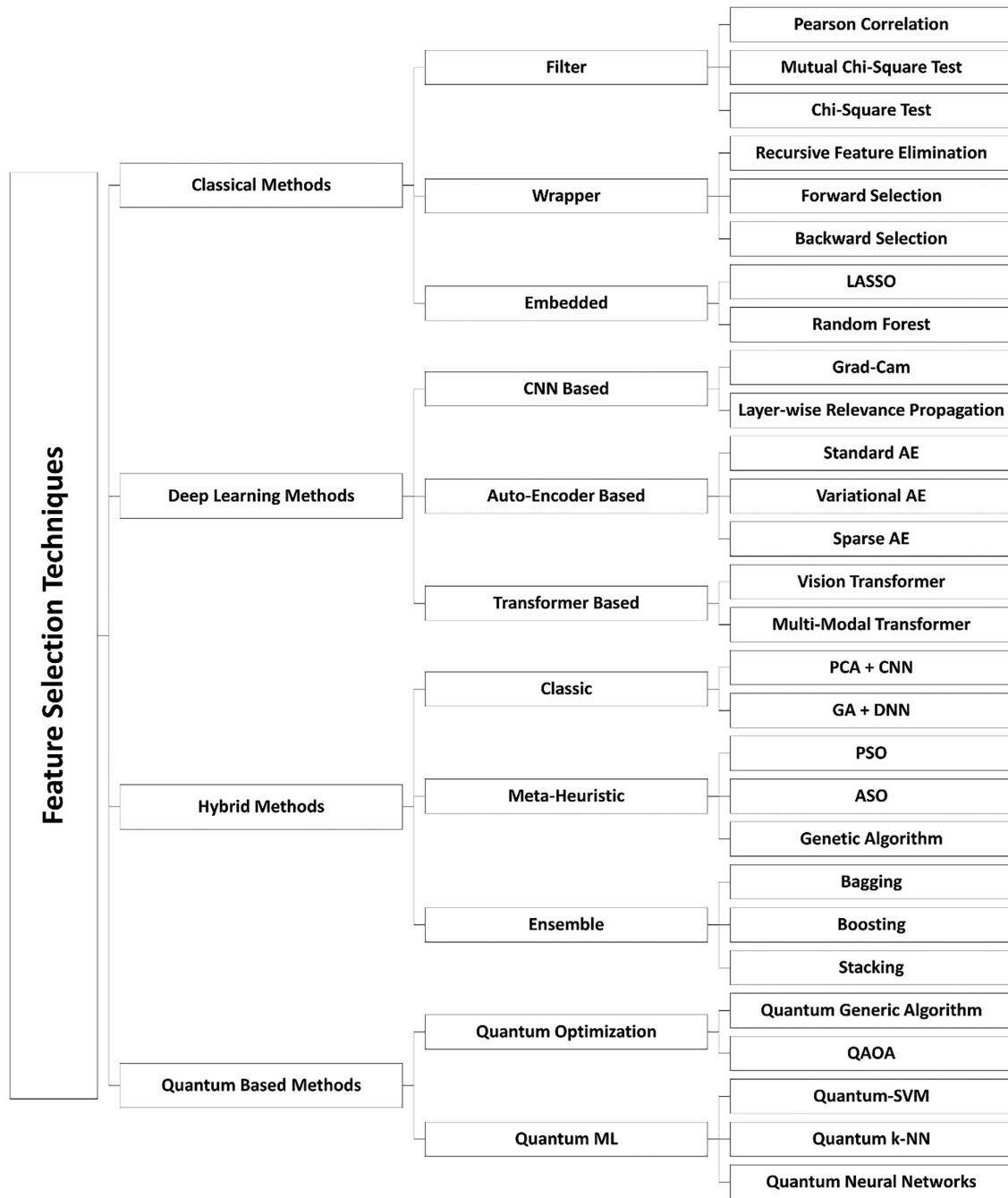


Figure 9: Taxonomy diagram summarizing the main categories of feature selection techniques

4.4 Ethical & Bias Concerns

By adopting AI-based FS models is occurring amongst ethics and bias in healthcare. These have shown tremendous promise in boosting diagnostic accuracy with significant reduction in computational Overhead and self execution of difficult acquired manual tasks in medicine [\[160\]](#). However, AI FS models

can themselves accidentally perpetuate exactly these racial dimensions of healthcare systems (and of our current world) if we fail to consider them. This is likely because these models were trained on historical bias, but disparities in healthcare make historical data a poor representation of the patient base in the future. Such biases can introduce suboptimal (and in some cases harmful) decision-making into subsequent deployment of these AI models into the clinic to assist in disease diagnosis or prediction of patient outcomes [161].

For demonstration, let's say an AI FS model has been trained mainly on data for demographics, say white or male patients, and when our trained model doesn't work correctly or if not, even wrongly do it while practicing on other races or similar gender groups. The issue is that being trained on a limited set of a few subsets of the population, such as medical images, essential features that may be critical for the underrepresented group in the medical image may be missed. However, there might be misdiagnoses, delayed treatment, and illness-related healthcare delivery disparities.

With the proposed bias-aware FS techniques, the risk of forming prejudice on algorithms is reduced through AI model training. Thus, the purpose of imposing fairness constraints is to discover and fix the bias introduced into the feature selection process. Still, we want to pick the features to be used in our model that bring accurate results and do not worsen the existing data bias. The most common method of addressing demographic parity is explicitly penalizing feature subsets to get biased prediction distributions to guarantee that the holdem decisions by the model are fair and comparable between demographic groups. The bias aware FS can be mathematically defined as follows in Eq. (30).

$$L_{\text{total}} = L_{\text{accuracy}} + L_{\text{fairness}} \quad (30)$$

where:

- L_{total} is the total loss function, combining both accuracy and fairness objectives,
- L_{accuracy} is the traditional loss function that measures model performance (e.g., classification accuracy),
- L_{fairness} represents the fairness penalty, which quantifies how much bias is introduced in the feature selection process, and
- L is a regularization parameter that controls the trade-off between accuracy and fairness.

Fairness constraints incorporate fairness considerations by bias-aware FS methods, which aim to guarantee that the features selected for the model do not perpetuate harmful biases that cause unfair or discriminatory outcomes.

Diverse datasets can also be an essential solution in dealing with the bias in AI FS models. When training data sets cannot be guaranteed to include the population, one must ensure they are representative to avoid reinforcing biases in the final model. Especially in the case of medical imaging datasets, the dataset must contain demographic groups that cover a wide range of races, genders, age groups, and socioeconomic statuses. Having various datasets helps the model learn to distinguish features that are not biased toward any specific group, and these learnt features help the generalization and fairness of the model when applied in clinics.

One challenge is creating diverse datasets: creating datasets from underrepresented populations, such as demographics typically underrepresented in healthcare datasets, is logically difficult and resource-intensive. However, collecting datasets for building AI FS models that are accurate and fair for all patients must be more inclusive, including ensuring access to healthcare for all demographic groups.

Besides bias-aware FS and diverse datasets, AI regulation policies regulate bias in deploying AI systems in health care. Governments and regulatory bodies can create rules for the ethical use of AI in the healthcare sector in the most apt manner. The regulations may include auditing AI models for biases, instructing

healthcare professionals to be trained to trust the results provided by the AI models properly, and ensuring the AI models are updated with new data and evolving social norms [162].

For example, data collection, processing, and usage for training AI should be transparent via policies. Further, frameworks for auditing AI systems for fairness, accountability, and transparency (FAT, or Fairness, Accountability, and Transparency) would be introduced to detect and tackle biases in the model. The regulatory framework could also require it to enforce the continued monitoring of deployed models to prevent the introduction or perpetuation of new bias over time. The regulatory equation that aims to ensure AI fairness and accountability could go as follows in Eq. (31).

$$\text{AI Fairness} = \frac{\text{Bias Reduction}}{\text{Bias Detection} + \text{Ethical Oversight}} \quad (31)$$

where:

- Bias Reduction refers to the efforts made to minimize bias during the training and deployment of the AI model,
- Bias Detection involves monitoring AI models post-deployment for emerging biases, and
- Ethical Oversight represents the ongoing review and regulation of AI models to ensure compliance with fairness standards.

Finally, a multi-faceted approach is advocated to mitigate the ethical and biases issues in AI-driven FS for healthcare, ranging from regulatory to organizational and technical matters. FS methods in the presence of bias assist in removing discrimination from feature selection so that the resulting model is not more favorable towards one demography than the other. Diverse datasets are key to training AI models on datasets that represent all the groups proportionally, and AI regulation policies constitute provoking the ethical use of AI to prevent unintentional reinforcement of biases. These strategies are necessary to guarantee that AI healthcare systems will be effective and equitable, thus producing equitable treatment for all patients as illustrated in Table 7.

Table 7: Ethical considerations in AI-based FS

Ethical concern	Description	Proposed solutions	Impact on FS	Challenges
Bias in data [163]	Skewed training data can exacerbate racial, gender, and other demographic biases in AI FS models.	Bias-aware FS techniques Diverse datasets	Ensures model fairness and prevents discrimination.	Challenges include gathering varied data and balancing bias during training.
Fairness and equity [164]	Biased feature selection or training data causes demographic disparities in performance.	Fairness-aware loss functions Fairness audits	Ensures model projections are fair across demographics.	Multiple-criteria fairness definition and measurement.

(Continued)

Table 7 (continued)

Ethical concern	Description	Proposed solutions	Impact on FS	Challenges
Transparency and explainability [165]	AI models, intense learning, are often viewed as “black boxes” with unclear decision-making processes.	XAI frameworks SHAP, Grad-CAM, Hybrid FS	Improves the trust and interpretability of AI systems in clinical settings.	Balancing model performance with explainability, especially in complex models.
Data privacy [166]	Patient data used for AI FS might be sensitive and subject to privacy concerns.	Data anonymization Federated learning Secure data sharing protocols	Ensures that patient data is protected while still allowing for functional AI training.	Ensuring compliance with privacy regulations, such as HIPAA.
Regulatory compliance [167]	AI models may not meet fairness, safety, and efficacy norms in clinical settings.	AI regulation policies Regular audits for fairness and compliance	This framework ensures that AI models fulfill healthcare and ethical requirements.	Clear regulatory frameworks that adapt to new technologies.
Accountability [168]	Assigning blame for healthcare AI model errors.	Clear accountability protocols Human-in-the-loop models	Holds clinicians accountable for ultimate judgments, even with AI.	Determining who is responsible when AI-driven decisions cause harm.

4.5 Future Directions: Quantum-Driven Feature Selection

Quantum Computing has shown great promise in delivering transformational technology among the domains of AI, in one of the domains such as FS in medical imaging. However, classical FS techniques find difficulty in dealing with the ultra-high dimensionality and with data sets that are very large in size, making quantum-enhanced methods promising work in the future. The implication is that with quantum computing, we can process a vast amount of data in parallel and solve complex optimization and intricate feature interactions much more effectively than with classical approaches. But as the progress continues with quantum computing, we anticipate quantum-driven FS techniques for similar purposes and gaining even more capabilities for these AI models in medical imaging and genomics and other data-rich domains.

Secondly, because SVM machines can support the separation of data points, they are very much used in feature selection, classification, and regression tasks. However, these can be spoiled when dealing with high-dimensional spaces due to the computational cost of solving the optimization problems caused by the explosion of the feature space. Quantum-escalated SVMs use quantum parallelism and quantum kernel strategies from quantum registering to perform the optimization and arrange more quickly.

The quantum-enhanced SVMs involve a quantum algorithm to get quantum computation of the inner product (or kernel) of data points and perform the calculation in a much higher dimension feature space. This kernel trick is written mathematically and is called the quantum kernel method in the quantum world illustrates in Eq. (32).

$$K(x, y) = |\langle \psi(x) | \psi(y) \rangle|^2 \quad (32)$$

where:

- $K(x, y)$ is the kernel function between data points x and y ,
- $|\psi(x)\rangle$ and $|\psi(y)\rangle$ are quantum states corresponding to the data points.

The power of this quantum kernel is to implicitly map data from the input space into a higher dimensional kernel feature space, which allows SVMs to perform nonlinear separation more efficiently. Quantum computing can make SVMs faster, less resource-consuming, and suitable for FS tasks in medical imaging and other complex domains like fraud detection.

Feature selection becomes a powerful application of neural networks, in particular in the depth of deep learning. However, their high computational cost in the application to large datasets and complex networks is a problem. Quantum-enhanced Neural Networks (QNNs) try to benefit from quantum circuits to accelerate the training and improve the feature extraction capability of neural networks. Quantum circuits can process data in parallel and perform multivariable operations on a massive dataset with an efficiency that defies classical neural networks.

In QNNs, quantum gates replace classical operations, enabling the network to deal with an exponentially greater superposition of data at a computational power. A quantum neural network can be represented in general structure in Eq. (33) as:

$$\text{output} = \sum_{i=1}^N a_i |\psi_i\rangle \quad (33)$$

where:

- a_i represents the weights of the quantum circuit,
- $|\psi_i\rangle$ is the quantum state corresponding to the input feature.

Since quantum entanglement and superposition are utilized, QNNs can better extract the right features from high-dimensional data than classical networks. Thus, because of the simpler QNNs, FS processes are accelerated with high accuracy, particularly in large-scale medical image tasks.

Cluster or clustering is another feature selection technique, mainly for unsupervised learning problems by not having labels. Quantum clustering uses the fact that certain operations can be performed in superposition on a quantum computer, and thus, several cluster configurations can be explored simultaneously. Quantum K-means is a quantum algorithm that utilizes entanglement and superposition to perform clustering much faster than a classical clustering could, particularly in high-dimensional spaces. It gives a representation of the quantum version of the K-means algorithm in Eq. (34) as:

$$\text{Distance} = |\langle \psi(x) | \psi(\mu) \rangle|^2 \quad (34)$$

where:

- $\psi(x)$ and $\psi(\mu)$ are the quantum states of the data point and the cluster centroid, respectively,
- The distance metric is computed as the overlap between these quantum states.

With the progress of quantum computing technology, it would likely be a vital medical imaging tool with applications such as novel approaches to FS and ultimately revolutionizing how AI can be harnessed in healthcare. [Fig. 10](#) illustrates the quantum feature selection workflow for medical imaging.

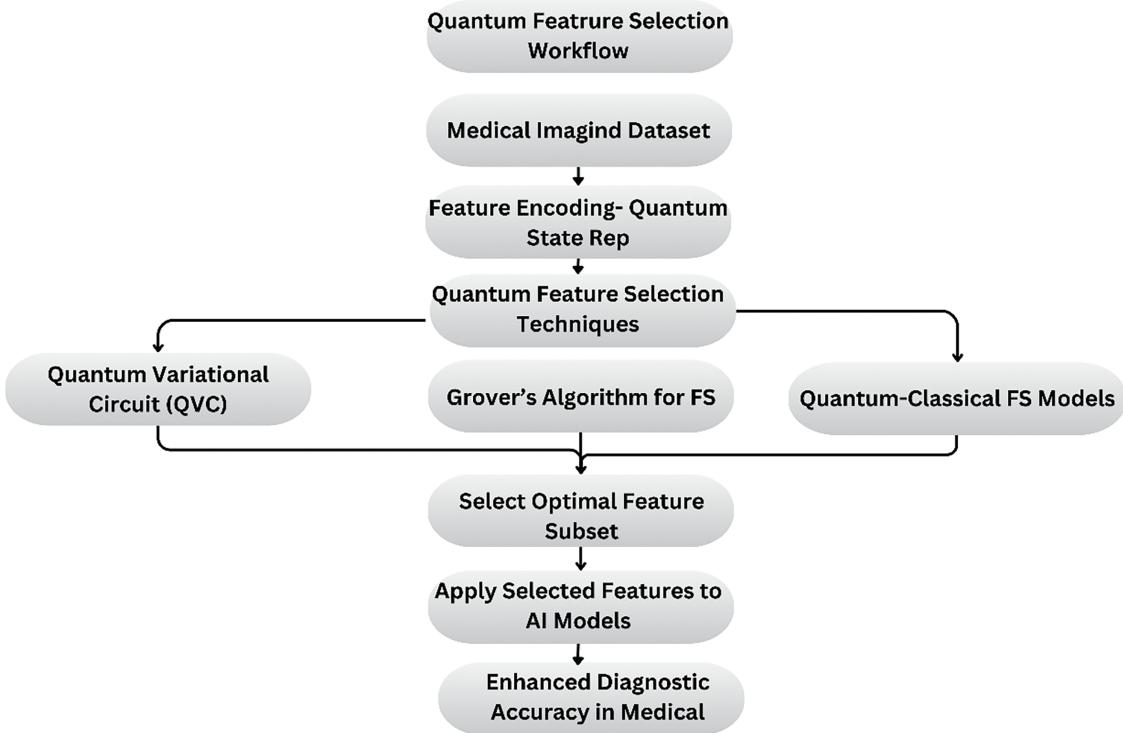


Figure 10: Quantum feature selection workflow applied to medical imaging. It highlights steps such as using a medical imaging dataset, encoding features with quantum state representation, applying quantum feature selection techniques like Grover's Algorithm, and integrating quantum-classical models. The final goal is to select an optimal feature subset and apply it to AI models, enhancing diagnostic accuracy in medical imaging

By processing large numbers of potential cluster configurations in parallel, thanks to quantum clustering, it is faster converging, and thus, more accurate and efficient features can be selected. In particular, this is very advantageous for medical imaging when the data is noisy, high dimensional, and often complex, and finding the correct clusters or groups of features is crucial to improving diagnostic accuracy.

Finally, in summary, quantum-driven feature selection has much room for mitigation of the limitations of classical FS methods, particularly in the field of medical imaging. Future FS models can analyze more extensive and complex datasets with greater computational efficiency using the quantum-enhanced algorithms such as Quantum Enhanced SVMs, Quantum Neural Networks, Quantum Clustering, etc. The advancements expected in these areas will help usher in AI models capable of helping process all these medical imaging data more efficiently, improving both diagnosis speed and accuracy, as well as patient outcome.

5 Discussion and Recommendation

From classical statistic techniques to mainly hybrid and quantum-assisted techniques, along with deep learning, FS problems in medical imaging have been addressed. A review of this set of FS methodologies is provided in this study, and their strengths, limitations, and future research directions are given. Although deep learning and hybrid FS methods are shown to have better accuracy and feature extraction efficiency than

the classical techniques, the former methods still face the challenge of issues related to their interpretability, computational complexity, and scalability. However, the practical implementation of FS in the sensible world of medical imaging to date is only possible because of hardware constraints and slow algorithmic maturity. This paper concludes by synthesizing the significant findings of the review, points to challenges of recent FS techniques, and indicates future research directions to bridge gaps.

It turns out that the classical FS methods (filter, for example, mutual information, Pearson correlation; wrapper: recursive feature elimination; embedded: LASSO, decision tree feature importance) are still in favor of implementation because they are easier to interpret, much more efficient from the computational point of view, and easy to implement. However, since these could not handle high-dimensional and multimodal medical imaging data, AI-based FS techniques came into existence.

FS with deep learning, especially with CNN, Autoencoders, or Transformers, has dramatically improved the extraction and selection of the features. We propose hierarchies of such a model that automatically learn memo hashes and hierarchies of memo hashes without manual feature engineering. However, since they are black boxes, the interpretability to clinicians to trust and validate AI-based FS decisions is an issue. Moreover, these models are computationally heavy, making them unwieldy to use in a clinical setting in real time due to the need for high-performance hardware (GPUs or TPUs) and a performant hyperparameter configuration.

Hybrid FS approaches, which mainly contain classical and deep learning methods, have been suggested to counter these things. These methods extend the best of FS techniques in two directions: (i) such models generalize better than FS techniques or AI-driven models, and (ii) the models are interpretable compared to AI-driven models. However, when used to apply to a large-scale, multimodal imaging dataset, the hybrid methods are constrained by scalability and computational optimization issues.

In particular, quantum computing is also demonstrated as a new approach to FS, deep learning, and hybrid methods. Usually, Quantum Feature Selection (QFS) utilizes quantum parallelism and quantum entanglement to use a high-dimensional dataset. Although quantum FS has the theoretical advantage of exponentially improving computational efficiency, its realization in practice faces hardware limitations, quantum noise, and the absence of clinically validated quantum algorithms.

In addition, the studies show that although deep learning and hybrid FS methods need to be further studied, they could significantly improve the efficiency of feature extraction and predictive accuracy over the classical FS methods. On the other hand, classical FS methods are based on preassigned statistical measures. In contrast, AI-driven approaches can learn complex feature representations in an automatic fashion from medical imaging datasets and understand the underlying complex patterns in medical imaging databases. However, they incur this advantage on computational efficiency and interpretability.

Deep learning based FS is faced with the challenge of a blackbox, and medical professionals can hardly explain to what extent the selected features are valuable and why we should exclude other features. To increase transparency, various explainable AI (XAI) techniques, such as SHAP (Shapley Additive Explanations), Grad-CAM (Gradient-weighted Class Activation Mapping), and Layer-wise Relevance Propagation (LRP), have been determined. These methods, though, are not used widely in clinical decision-making because feature selection is not justified.

Additionally, unlike deep learning based FS, using classical FS approaches is a costly operation, computationally speaking. As an example, the computational complexity of CNN-based FS models is illustrated in Eq. (35):

$$C_{FS} = \sum_{l=1}^L O(f_l \cdot k_l^2 \cdot c_l \cdot h_l \cdot w_l) \quad (35)$$

where L is the total number of layers, f_l represents the number of filters, k_l^2 is the kernel size, and h_l, w_l denote the feature map dimensions. This equation summarizes how deep architectures exponentially boost the computation demands, making real-time applicability for medical purposes intractable unless special hardware is considered.

Deep learning-based FS is also pursued to increase scalability. This entails utilizing lightweight CNN architectures, e.g., MobileNets and EfficientNets, to reduce the number of trainable parameters while maintaining FS accuracy. The depth-wise separable convolution is utilized to mathematically optimize the computational efficiency of lightweight FS models. Recent studies have also demonstrated that lightweight CNNs, including MobileNet, can accurately classify multiple stages of Alzheimer's with high accuracy, improving clinical interpretability and efficiency [65]. It illustrates in Eq. (36) as:

$$C_{DW} = O \left(\sum_{l=1}^L k_l^2 \cdot c_l \cdot h_l \cdot w_l + f_l \cdot c_l \cdot h_l \cdot w_l \right) \quad (36)$$

This alleviates computational load by performing no redundant calculations and enables lightweight models to be used for real-time FS in resource-constrained environments.

Federated Learning-Based FS is another promising solution. In this approach, feature selection is performed in a decentralized manner, so sensitive patient data does not have to be transferred to centralized servers. This approach enables several healthcare institutions to contribute to FS tasks securely without significant computational burdens.

Namely, quantum computing provides an entirely disruptive approach to FS with an exponential speedup over classical FS techniques. Quantum Feature Selection (QFS) considers multiple feature sets to evaluate in the superposition and entanglement by the computer to diminish computational time. Quantum FS can be mathematically represented in Eq. (37) as:

$$|\Psi\rangle = U_\theta |0\rangle^{\otimes n} \quad (37)$$

where $|\Psi\rangle$ represents the quantum state encoding selected features, and U_θ is the optimized unitary transformation applied for FS. Quantum classical FS could be more efficient with big multimodal imaging data, unlike classical portable FS, which has scalability issues with high-dimensional data sets.

However, the implementation of quantum FS in practice is still hindered by practical limitations of the current hardware. However, the existing quantum processors are insufficient in qubits, have high noise sensitivity, and are unsuitable for large-scale medical imaging adoption. Hybrid quantum classical FS approaches involving the use of quantum computing along with the deep learning frameworks could be intermediate steps to apply FS in the clinical setting.

5.1 Related Surveys and Positioning

Although several reviews have addressed feature selection or related concepts in medical imaging, none to date provide a comprehensive, systematic review covering classical, deep learning-based, hybrid, and quantum-based feature selection across multi-modal imaging. This subsection highlights selected works and clarifies the unique contribution of this review.

Several recent works provide valuable insights (Table 8):

- [13] Bolón-Canedo and Remeseiro (2019): General FS methods applied in medicine including signals and microarrays, but not focused on imaging FS paradigms.

- [35] Naheed et al. (2020): Broad FS coverage for medical imaging, lacks modern methods like federated or quantum FS.
- [44] Zang et al. (2023): Radiomics-focused FS taxonomy but no integration of hybrid, quantum, or federated FS.
- [168] Kaur and Bohmrah (2025): Review of hybrid deep neural networks and metaheuristic optimization for disease detection. Focuses on hyperparameter tuning and hybrid DNNs, lacks taxonomy or modality-wide FS review.
- [169] Perniciano et al. (2024): Focuses on radiomics, especially SVM and LASSO usage. Imaging-centric but not taxonomy-rich.
- [170] Zouache et al. (2023): Focuses on multi-objective feature selection using Firefly Algorithm (FA) and PSO for COVID-19 detection with CNNs on X-ray images.
- [171] Focus on FS and classification in RNA-Seq using XGBoost and DT for pipeline analysis with limited FS taxonomy.
- [172] Tumor subtype diagnosis using multi-omics and imaging with Filter, Wrapper, and Embedded methods in ML, limited by a narrow biological scope.
- [173] Focus on disease-specific FS applications without broader methodological scope.

Table 8: Comparison of our review with prior surveys on feature selection in medical imaging

Ref.	Year	Focus area	Imaging modalities	FS categories covered	Novel aspects	Limitation
[13]	2019	General FS in medicine	Imaging + non-imaging	Filter, Wrapper, Embedded	Intro to FS for medical apps	No DL or modality focus
[35]	2023	Feature selection in machine learning	X-ray images	Multi-objective feature selection	FA, PSO, and quantum computing	
[44]	2023	Radiomics	CT, MRI	Filter, Wrapper, Embedded	Radiomics FS taxonomy	No hybrid or emerging FS
[168]	2025	DNN + metaheuristics	General	DL + Hybrid	Hyperparameter optimization	No classical or quantum FS
[169]	2024	Feature selection and classification in Radiomics	Radiological images (e.g., CT, MRI, Ultrasound)	Dimensionality reduction, Classification accuracy	SVM classifier and LASSO feature selection; focus on ensemble and hybrid methods	High-dimensionality; algorithm stability study needed
[170]	2024	Multi-objective feature selection for COVID-19 detection	X-ray images	Dimensionality reduction, Classification accuracy	Combination of FA and PSO with quantum computing,	Limited to X-ray images, potential computational complexity
[171]	2024	FS + classification	RNA-Seq	XGBoost, DT	Pipeline analysis	Limited FS taxonomy
[172]	2025	Tumor subtype diagnosis	Multi-omics + imaging	Filter, Wrapper, Embedded	ML + multi-omics	Narrow biological scope
[173]	2024	COVID-19 diagnosis	Chest CT scans	Cuckoo search optimization algorithm, XGBoost	High classification rate	Dependent on feature type
Our Review	2025	FS in medical imaging	MRI, CT, US, X-ray, PET	Classical, DL, Hybrid, Quantum	First to include quantum, federated, ethical analysis	—

The unique contributions of this review are:

- Taxonomy-based comparison across classical, deep learning, hybrid, and quantum FS paradigms.
- Inclusion of multi-modal imaging and challenges such as multi-omics fusion and feature disentanglement.
- Discussion of ethical, privacy-preserving, and federated learning FS aspects.
- Presentation of a PRISMA-based systematic review methodology.
- Comparative focus on interpretability, efficiency, and clinical applicability.

Table 9 presents a comparative summary of the key strengths, limitations, and clinical considerations across classical, deep learning, hybrid, and emerging feature selection (FS) techniques discussed in this review.

Table 9: Comparative summary of feature selection techniques

FS method	Strengths	Limitations	Clinical applicability	Interpretability
Filter Methods	Fast, computationally efficient, independent of ML models	Ignores feature interaction; less effective for complex imaging tasks	High (simple diagnostics)	High
Wrapper Methods	Higher accuracy; model-specific optimization	Computationally expensive; prone to overfitting	Moderate	Medium
Embedded Methods	Balance of performance and efficiency; good model integration	Model-dependent; may miss global optima	Moderate to High	Medium
Deep Learning FS	Automatic feature learning; high predictive accuracy	Black-box nature; high hardware requirements	Limited (real-time use)	Low
Hybrid FS	Combines interpretability and accuracy; balances trade-offs	Scalability issues on large/multimodal datasets	Moderate	Medium to High
Lightweight DL (e.g., MobileNet)	Efficient, scalable; optimized for edge devices	Slight drop in accuracy; still needs tuning	High (point-of-care)	Low
Federated FS	Privacy-preserving; suitable for multi-institutional data	Communication overhead; synchronization issues	Growing	Medium
Quantum FS	Theoretical exponential speed-up; handles high-dimensional data	Experimental phase; hardware and algorithmic immaturity	Future potential	Currently low

(Continued)

Table 9 (continued)

FS method	Strengths	Limitations	Clinical applicability	Interpretability
Explainable AI (XAI)	Enhances transparency of DL-based FS	Not yet widely validated or standardized in clinical use	Promising	High (if integrated well)

5.2 Feature Selection Techniques by Imaging Modality

Table 10 summarizes the most commonly used feature selection (FS) techniques applied across various medical imaging modalities, along with representative applications. This structured overview aids in understanding the domain-specific relevance and adoption trends of FS methods in medical image analysis.

Table 10: Feature selection techniques by imaging modality

Imaging modality	Common FS methods	Applications
MRI	PCA, LASSO, Autoencoders, CNN-based FS, Quantum FS	Brain tumor classification, Alzheimer's detection
CT	Recursive Feature Elimination, Random Forest Importance, CNNs	Lung nodule detection, liver tumor segmentation
X-ray	Chi-square, ReliefF, Transfer Learning with FS	COVID-19 detection, Pneumonia classification
Ultrasound (US)	Mutual Information, Gabor Filters, Wrapper Methods	Breast lesion classification, fetal abnormality detection
PET	LASSO, SHAP-based FS, Federated FS using Tree Models	Cancer staging, brain metabolic analysis

5.3 Recommendations for Future Research

The findings have been reviewed to show that although there has been considerable progress in applying FS methods in medical imaging, there are still several significant challenges in this area. Future works focused on interpretability, computational efficiency, ethical issues, and the coupling of new technologies, such as quantum computing, will help make FS a more practical choice in clinical settings. This leads to the following recommendations across these medical imaging FS methodologies that, in particular, provide a structured roadmap for how one might advance onto these FS methodologies, specifically, more interpretable, more computationally efficient, and more ethical and future technology sound.

Nevertheless, one of the main barriers toward clinical deployment of deep learning is the lack of 'interpretability' in deep learning based FS. Most FS models, such as those derived from Convolutional neural networks (CNNs), Autoencoders, and Transformers, are black box systems, and it is challenging to grasp the reason for including or discarding which feature. Future research should be dedicated to increasing the FS model transparency and trust of FS models based on XAI techniques. SHAP (Shapley Additive Explanations), Grad CAM (Gradient-weighted Class Activation Mapping), or hybrid FS models are not interpretable by a few clinical validation studies. Therefore, for these techniques to be effective and readable to medical professionals, they should be standardized in future work.

Additionally, future FS models that use such a structure must contain domain-specific knowledge to increase their interpretability. Data-driven, existing deep learning-based FS techniques select the features based on their correlations with the prediction tasks, which could be caused by some anomalies and have no clinical significance. By using the medical knowledge of radiomic features and human-in-the-loop AI systems, the FS model transparency can be strengthened while ensuring the selected features comply with clinical frameworks for making a decision.

While the diagnostic ability and feature extraction capability of deep learning FS techniques are substantially improved, the major drawbacks are high computational cost and scalability issues preventing their practical deployment in real-time medical applications. The construction of lightweight FS architectures (MobileNets and EfficientNets) that reduce the trainable parameters but do not significantly affect performance is adopted to improve the computational efficiency. Depthwise separable convolutions are adopted in these architectures to reduce computational complexity and maintain the capacity for discriminative feature selection. To this end, these architectures should be optimized for mobile and edge computing with real-time point of care ultrasound (POCUS) analysis and mobile ECG-based heart disease classification.

In addition, federated learning based FS is an important area for research in the future because it is helpful for distributed learning with confidential preservation, where one learns across several datasets without pooling them in centralized storage. It is essential in healthcare because patient privacy regulations (HIPAA, GDPR) apply when you share data. Fortunately, with FS techniques, researchers can federate learning with FS and utilize multi-institutional medical imaging databases with data privacy and security in mind. Nevertheless, the federated FS models should be optimized towards efficient communication to avoid bandwidth costs without compromising model performance.

The medical imaging datasets, however, suffer from a common problem: they are imbalanced and often contain demographic bias, which leads to real-world FS models that are neither generalized nor fair. More accurately, the FS techniques may choose the significant lures in the overrepresented patient groups but fail to detect the diagnostic patterns about the patient groups that constitute the minority, and such bias can influence clinical outcomes. In the future, research centers should be trained on diverse and representative datasets, i.e., have a variation in age, gender, ethnicity, and disease subtypes.

There is, however, an acute need to accelerate the process of regulatory guidelines for the ethical use of AI in medical imaging on aspects of FS model fairness, transparency, and accountability, among others. Yet, the current rules for utilizing FS techniques in light of AI have not been completely developed. However, the World Health Organization (WHO) and the FDA are discussing regulative AI frameworks in healthcare. Future research includes developing auditable FS models, wherein explaining what decides selected features in a given model aids conformance to ethical AI principles.

Quantum computing offers the potential for computationally efficient FS research compared to classical methods in an almost exponential fashion. Quantum Feature Selection (QFS) can process high-dimensional data in polynomial time using quantum superposition and entanglement, making it an attractive prospect for solving multimodal medical Imaging using quantitative methods. Practical realization, however, still represents a daunting problem due to faults in qubits, restrictions in hardware, and algorithms that are still a work in progress.

Future work in this area involves developing hybrid quantum-classical FS models due to incorporating quantum-assisted FS methods into existing AI frameworks. As an instance, VariationalQC VQCs can, for example, be used to optimize FS processes but have to be validated on actual-world medical imaging data sets. At the same time, Grover's Search Algorithm can be applied to optimize FS processes but must also be validated on real-world medical imaging data sets. The quantum enhanced FS can be mathematically as

follows:

$$|\Psi\rangle = U_\theta |0\rangle^{\otimes n} \quad (38)$$

where $|\Psi\rangle$ represents the quantum state encoding feature selection, U_θ is the unitary transformation optimized for FS, and $|0\rangle^{\otimes n}$ denotes the initial quantum state with n Qubits. Such a formulation implies a significant reduction in the computational complexity of quantum algorithms that implement FS, but practical deployment involves substantially upgrading quantum hardware.

Future work on quantum FS will include using quantum error correction techniques to improve the noise and decoherence issues in the hardware. One proposed way to address the issue of quantum fault tolerance to general programming (i.e., for use in FS tasks in medical imaging) is using quantum error-correcting codes (QECCs) such as Shor's Code and Surface Code. Nevertheless, these methods are still in the early experimental phases, leaving many studies to be done in the direction of scalable quantum architectures that can tackle many feature selection tasks.

For future research in FS for medical imaging, better explainability, improved computational efficiency, reduced bias, and advanced methodologies that leverage quantum computation are necessary. Standardization and clinical validation of explainable AI techniques should be done to enable greater transparency and trust in FS models. Lightweight FS architectures and federated learning are enabled for efficient computation to perform FS in real time for clinical applications. Fairness in AI-driven FS techniques is vital to address bias and ethical concerns, especially when addressing patient groups not well-represented in the data. Lastly, quantum computing features a revolutionary use for FS, which, however, still needs to mature hardware and algorithmically before seeing any use in real-world medical image processing workflows.

In addition to SHAP and Grad-CAM, LIME (Local Interpretable Model-agnostic Explanations) offers another potent approach to explainability-driven feature selection [174]. LIME works by perturbing input features and observing the corresponding changes in model predictions, thereby constructing a locally interpretable approximation of the model's behavior. This technique is particularly beneficial in identifying which features have the greatest influence on individual predictions, making it highly suitable in patient-specific diagnostic scenarios. However, its reliance on surrogate models and sensitivity to sampling methods can sometimes limit its reliability in high-dimensional medical imaging data. Future research should explore how LIME can be integrated with domain-specific constraints to ensure that the explanations it produces are not only technically valid but also clinically meaningful. In particular, combining LIME with structured prior knowledge (e.g., known radiomic markers or anatomical relevance) may improve both interpretability and trustworthiness of the FS process in real-world medical settings.

5.4 Limitations of Proposed Study

Even though the current review adheres to such stringent methodological standards, it is important to point out that it does have a few drawbacks. To begin, studies that are not conducted in English are going to be disregarded; as a result, there is a risk of bias in language, and the research as a whole would be unable to capture literature that is relevant and was published in other languages. Second, the sources that contained grey literature, such as technical reports, preprints, and dissertations, were not included. This may have resulted in the absence of explanations for new work that was not specifically documented. Third, despite the fact that the selection of the study and the extraction of the data were carried out independently by two reviewers, there was no calculation of inter-rater concordance measures (such as the Cohen Kappa statistic). This is a feature that, if calculated, would have increased the reliability of the selection process. The fact that the use of automation tools was purposefully omitted throughout the screening process is another evidence that this choice was made. Despite the fact that this choice has the ability to reduce the amount

of human error, it was decided to employ manual validation in order to obtain full coverage. Fourth, there is no implementation of meta-analysis, which limits the ability to quantitatively evaluate heterogeneity and variation in impact sizes across the studies that have been found.

6 Conclusion

Feature selection is a rapidly evolving field in medical imaging, where deep learning and hybrid methods are increasingly surpassing classical approaches. Although classical techniques are affordable and interpretable, they often lack the capacity to process high-dimensional and heterogeneous imaging data effectively. In contrast, deep learning-based feature selection models—particularly CNNs, autoencoders, and transformers—can learn complex, higher-order imaging features. However, these methods tend to be less interpretable, computationally intensive, and difficult to scale in clinical environments. Hybrid feature selection approaches, which integrate classical and AI-driven methods, offer a middle ground by enhancing feature generation while maintaining a degree of interpretability. Yet, their clinical integration remains constrained by high computational demands and the challenges associated with feature fusion in multimodal imaging applications. A significant insight from this review is the potential of quantum computing to revolutionize feature selection. QFS leverages quantum superposition and entanglement to reduce the computational complexity of classical FS methods. Although promising in theory, QFS still faces practical challenges, including limitations in current quantum hardware, lack of algorithmic maturity, and the need for robust clinical validation. Nonetheless, hybrid quantum-classical frameworks may offer a transitional path toward integrating quantum FS with existing AI-driven systems. Future research should prioritize the development of robust and explainable AI-based FS models that can deliver both performance and transparency for clinical decision-making. There is a growing need for computationally efficient solutions, such as lightweight FS architectures and federated learning frameworks, especially in privacy-sensitive healthcare environments. Moreover, fairness-aware FS algorithms—coupled with demographically representative datasets—must be explored to address bias and equity concerns. Finally, the high scalability potential of quantum FS models makes them attractive candidates for future multimodal imaging applications. By addressing these technological and ethical challenges, feature selection techniques can be more seamlessly integrated into clinical AI systems, ultimately improving diagnostic accuracy, computational efficiency, and trust in real-world medical settings.

Abbreviations and Notations

Abbreviation	Full Form
ACO	Ant Colony Optimization
AUC-ROC	Area Under the Receiver Operating Characteristic Curve
BraTS	Brain Tumor Segmentation (dataset)
CheXpert	Chest X-ray dataset
CNN	Convolutional Neural Network
CT	Computed Tomography
DL	Deep Learning
ECG	Electrocardiogram
EEG	Electroencephalogram
FAT	Fairness, Accountability, and Transparency
FS	Feature Selection
GA	Genetic Algorithm
GDPR	General Data Protection Regulation
Grad-CAM	Gradient-weighted Class Activation Mapping

GRU	Gated Recurrent Unit
GNN	Graph Neural Network
HIPAA	Health Insurance Portability and Accountability Act
KL divergence	Kullback-Leibler Divergence
k-NN	k-Nearest Neighbors
LASSO	Least Absolute Shrinkage and Selection Operator
LRP	Layer-wise Relevance Propagation
LSTM	Long Short-Term Memory
MI	Mutual Information
ML	Machine Learning
MRI	Magnetic Resonance Imaging
PCA	Principal Component Analysis
PET	Positron Emission Tomography
POCUS	Point of Care Ultrasound
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSO	Particle Swarm Optimization
QAOA	Quantum Approximate Optimization Algorithm
QFS	Quantum Feature Selection
QGA	Quantum Genetic Algorithm
QNN	Quantum Neural Network
QSVM	Quantum Support Vector Machine
QVC	Quantum Variational Circuit
ReLU	Rectified Linear Unit
RF	Random Forest
RFE	Recursive Feature Elimination
SAE	Sparse Autoencoder
SHAP	Shapley Additive Explanations
SSL	Self-Supervised Learning
SVM	Support Vector Machine
US	Ultrasound
VAE	Variational Autoencoder
ViT	Vision Transformer
VQC	Variational Quantum Circuit
WHO	World Health Organization
XAI	Explainable Artificial Intelligence
XGBoost	Extreme Gradient Boosting (algorithm)
Symbol	Meaning
X_i, Y_i	Feature values and target labels
\bar{X}, \bar{Y}	Mean of feature and target values
n	Number of samples
r	Pearson correlation coefficient
$I(X; Y)$	Mutual Information between variables X and Y
$P(x, y), P(x), P(y)$	Joint and marginal probabilities
χ^2	Chi-square statistic
O, E	Observed and expected frequencies
β	Feature weight vector
λ	Regularization parameter (LASSO)
F_l	Output feature map at layer l in a CNN
W_l	Convolutional weights at layer l

σ	Activation function (e.g., ReLU)
$q(z x), p(z x)$	Posterior and likelihood in VAEs
DKL	Kullback–Leibler divergence
Z	Transformer patch embeddings
Q, K, V	Query, Key, and Value matrices (transformer attention)
F^*	Optimal feature subset
V_i, X_i	Velocity and position in PSO
G_{bs}, P_{ibs}	Global and individual best solutions (PSO)
$d(F_1, F_2)$	Distance between feature vectors
P_i, w_i	Prediction and weight in late fusion

Acknowledgement: Not applicable.

Funding Statement: The authors received no specific funding for this study.

Author Contributions: Study conception and design: Sunawar Khan, Tehseen Mazhar, Naila Sammar Naz and Fahad Ahmed; data collection: Sunawar Khan, Naila Sammar Naz, Habib Hamam and Tariq Shahzad; analysis and interpretation of results: Tariq Shahzad, Atif Ali, Muhammad Adnan Khan and Habib Hamam; draft manuscript preparation: Sunawar Khan, Tehseen Mazhar, Habib Hamam and Fahad Ahmed. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: All data used in this article is included within the article itself.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.

Supplementary Materials: The supplementary material is available online at <https://www.techscience.com/doi/10.32604/cmc.2025.066932/s1>.

References

1. Abhisheka B, Biswas SK, Purkayastha B, Das D, Escargueil A. Recent trend in medical imaging modalities and their applications in disease diagnosis: a review. *Multimed Tools Appl.* 2024;83(14):43035–70. doi:10.1007/s11042-023-17326-1.
2. Barbieri MC, Grisci BI, Dorn M. Analysis and comparison of feature selection methods towards performance and stability. *Expert Syst Appl.* 2024;249(5):123667. doi:10.1016/j.eswa.2024.123667.
3. Rahman A, Debnath T, Kundu D, Khan MSI, Aishi AA, Sazzad S, et al. Machine learning and deep learning-based approach in smart healthcare: recent advances, applications, challenges and opportunities. *AIMS Public Health.* 2024;11(1):58–109. doi:10.3934/publichealth.2024004.
4. Crespo Márquez A. The curse of dimensionality. In: *Digital maintenance management*. Cham: Springer International Publishing; 2022. p. 67–86. doi: 10.1007/978-3-030-97660-6_7.
5. Ebrahimi Warkiani M, Moattar MH. A comprehensive survey on recent feature selection methods for mixed data: challenges, solutions and future directions. *Neurocomputing.* 2025;623(3):129372. doi:10.1016/j.neucom.2025.129372.
6. Cai J, Luo J, Wang S, Yang S. Feature selection in machine learning: a new perspective. *Neurocomputing.* 2018;300:70–9. doi:10.1016/j.neucom.2017.11.077.
7. Manikandan G, Abirami S. Feature selection is important: state-of-the-art methods and application domains of feature selection on high-dimensional data. In: *Applications in ubiquitous computing*. Cham: Springer International Publishing; 2020. p. 177–96. doi: 10.1007/978-3-030-35280-6_9.

8. Zhang Y, Liu Y, Chen CH. Review on deep learning in feature selection. In: The 10th International Conference on Computer Engineering and Networks. Singapore: Springer Singapore; 2020. p. 439–47. doi:10.1007/978-981-15-8462-6_49.
9. Azevedo BF, Rocha AMAC, Pereira AI. Hybrid approaches to optimization and machine learning methods: a systematic literature review. *Mach Learn.* 2024;113(7):4055–97. doi:10.1007/s10994-023-06467-x.
10. Piri J, Mohapatra P, Dey R, Acharya B, Gerogiannis VC, Kanavos A. Literature review on hybrid evolutionary approaches for feature selection. *Algorithms.* 2023;16(3):167. doi:10.3390/al6030167.
11. Pinto-Coelho L. How artificial intelligence is shaping medical imaging technology: a survey of innovations and applications. *Bioengineering.* 2023;10(12):1435. doi:10.3390/bioengineering10121435.
12. Ullah U, Garcia-Zapirain B. Quantum machine learning revolution in healthcare: a systematic review of emerging perspectives and applications. *IEEE Access.* 2024;12(5):11423–50. doi:10.1109/access.2024.3353461.
13. Bolón-Canedo V, Remeseiro B. Feature selection in image analysis: a survey. *Artif Intell Rev.* 2020;53(4):2905–31. doi:10.1007/s10462-019-09750-3.
14. Chen RC, Dewi C, Huang SW, Caraka RE. Selecting critical features for data classification based on machine learning methods. *J Big Data.* 2020;7(1):52. doi:10.1186/s40537-020-00327-4.
15. Britto AS, Sabourin R, Oliveira LES. Dynamic selection of classifiers—a comprehensive review. *Pattern Recognit.* 2014;47(11):3665–80. doi:10.1016/j.patcog.2014.05.003.
16. Gong H, Li Y, Zhang J, Zhang B, Wang X. A new filter feature selection algorithm for classification task by ensembling Pearson correlation coefficient and mutual information. *Eng Appl Artif Intell.* 2024;131(22):107865. doi:10.1016/j.engappai.2024.107865.
17. Qu G, Hariri S, Yousif M. A new dependency and correlation analysis for features. *IEEE Trans Knowl Data Eng.* 2005;17(9):1199–207. doi:10.1109/TKDE.2005.136.
18. Beraha M, Metelli AM, Papini M, Tirinzoni A, Restelli M. Feature selection via mutual information: new theoretical insights. In: 2019 International Joint Conference on Neural Networks (IJCNN); 2019 Jul 14–19; Budapest, Hungary. p. 1–9. doi:10.1109/ijcnn.2019.8852410.
19. Sharma N, Sharma M, Singhal A, Vyas R, Malik H, Afthanorhan A, et al. Recent trends in EEG-based motor imagery signal analysis and recognition: a comprehensive review. *IEEE Access.* 2023;11:80518–42. doi:10.1109/access.2023.3299497.
20. Miola AC, Miot HA. Comparing categorical variables in clinical and experimental studies. *J Vasc Bras.* 2022;21(7104):e20210225. doi:10.1590/1677-5449.20210225.
21. Fira M, Goras L, Costin HN. Evaluating sparse feature selection methods: a theoretical and empirical perspective. *Appl Sci.* 2025;15(7):3752. doi:10.3390/app15073752.
22. Njoku UF, Abelló Gamazo A, Bilalli B, Bontempi G. Wrapper methods for multi-objective feature selection. In: 26th International Conference on Extending Database Technology (EDBT 2023); 2023 Mar 28–31; Ioannina, Greece. p. 697–709.
23. Kuzudisli C, Bakir-Gungor B, Bulut N, Qaqish B, Yousef M. Review of feature selection approaches based on grouping of features. *PeerJ.* 2023;11(8):e15666. doi:10.7717/peerj.15666.
24. Selvaraj P, Sivaprakash S. Feature extraction and feature selection in medical images. In: Intelligent computing techniques in biomedical imaging. Amsterdam: Elsevier; 2025. p. 83–97. doi: 10.1016/b978-0-443-15999-2.00008-6.
25. Popov A. Feature engineering methods. In: Advanced methods in biomedical signal processing and analysis. Amsterdam: Elsevier; 2023. p. 1–29. doi: 10.1016/b978-0-323-85955-4.00004-1.
26. Borboudakis G, Tsamardinos I. Forward-backward selection with early dropping. *J Mach Learn Res.* 2019;20(8):1–39.
27. Mittal N, Kumar A. Analysis of supervised feature selection in bioinformatics. In: Blockchain applications for healthcare informatics. Amsterdam: Elsevier; 2022. p. 431–46. doi: 10.1016/b978-0-323-90615-9.00008-6.
28. Tayyeb M, Umer M, Alnowaiser K, Sadiq S, Abdulmajid Eshmawi A, Majeed R, et al. Deep learning approach for automatic cardiovascular disease prediction employing ECG signals. *Comput Model Eng Sci.* 2023;137(2):1677–94. doi:10.32604/cmes.2023.026535.

29. Yuan G, Li X, Qiu P, Zhou X. Feature selection method based on wavelet similarity combined with maximum information coefficient. *Inf Sci.* 2025;699(8):121801. doi:10.1016/j.ins.2024.121801.
30. Ros F, Riad R. Feature and dimensionality reduction for clustering with deep learning. Berlin, Germany: Springer; 2024.
31. Wang S, Celebi ME, Zhang YD, Yu X, Lu S, Yao X, et al. Advances in data preprocessing for biomedical data fusion: an overview of the methods, challenges, and prospects. *Inf Fusion.* 2021;76(2):376–421. doi:10.1016/j.inffus.2021.07.001.
32. Cui X, Xiao R, Liu X, Qiao H, Zheng X, Zhang Y, et al. Adaptive LASSO logistic regression based on particle swarm optimization for Alzheimer's disease early diagnosis. *Chemom Intell Lab Syst.* 2021;215(3):104316. doi:10.1016/j.chemolab.2021.104316.
33. Paul A, Mukherjee DP, Das P, Gangopadhyay A, Chinthia AR, Kundu S. Improved random forest for classification. *IEEE Trans Image Process.* 2018;27(8):4012–24. doi:10.1109/TIP.2018.2834830.
34. Wang Q, Nguyen TT, Huang JZ, Nguyen TT. An efficient random forests algorithm for high dimensional data classification. *Adv Data Anal Classif.* 2018;12(4):953–72. doi:10.1007/s11634-018-0318-1.
35. Naheed N, Shaheen M, Ali Khan S, Alawairdh M, Attique Khan M. Importance of features selection, attributes selection, challenges and future directions for medical imaging data: a review. *Comput Model Eng Sci.* 2020;125(1):315–44. doi:10.32604/cmes.2020.011380.
36. Aria M, Cuccurullo C, Gnasso A. A comparison among interpretative proposals for random forests. *Mach Learn Appl.* 2021;6(2):100094. doi:10.1016/j.mlwa.2021.100094.
37. Biesiada J, Duch W. Feature selection for high-dimensional data—a Pearson redundancy based filter. In: *Computer recognition systems 2.* Berlin/Heidelberg, Germany: Springer; 2007. p. 242–9. doi: 10.1007/978-3-540-75175-5_30.
38. Nguyen HB, Xue B, Andrae P. Mutual information estimation for filter based feature selection using particle swarm optimization. In: *Applications of evolutionary computation.* Cham: Springer International Publishing; 2016. p. 719–36. doi: 10.1007/978-3-319-31204-0_46.
39. Ding GY, Tan WM, Lin YP, Ling Y, Huang W, Zhang S, et al. Mining the interpretable prognostic features from pathological image of intrahepatic cholangiocarcinoma using multi-modal deep learning. *BMC Med.* 2024;22(1):282. doi:10.1186/s12916-024-03482-0.
40. Liu W, Wang J. Recursive elimination–election algorithms for wrapper feature selection. *Appl Soft Comput.* 2021;113(1–4):107956. doi:10.1016/j.asoc.2021.107956.
41. Toğçaçar M, Ergen B, Cömert Z. Detection of lung cancer on chest CT images using minimum redundancy maximum relevance feature selection method with convolutional neural networks. *Biocybern Biomed Eng.* 2020;40(1):23–39. doi:10.1016/j.bbe.2019.11.004.
42. Saim MM, Ammor H. Comparative study of feature selection algorithms for cardiovascular disease prediction with artificial neural networks. In: *Smart applications and data analysis.* Cham: Springer Nature Switzerland; 2024. p. 218–29. doi: 10.1007/978-3-031-77040-1_16.
43. Sharma S, Mandal PK. A comprehensive report on machine learning-based early detection of Alzheimer's disease using multi-modal neuroimaging data. *ACM Comput Surv.* 2023;55(2):1–44. doi:10.1145/3492865.
44. Zhang W, Guo Y, Jin Q. Radiomics and its feature selection: a review. *Symmetry.* 2023;15(10):1834. doi:10.3390/sym15101834.
45. Mirza B, Wang W, Wang J, Choi H, Chung NC, Ping P. Machine learning and integrative analysis of biomedical big data. *Genes.* 2019;10(2):87. doi:10.3390/genes10020087.
46. Wang H. Research on the application of random forest-based feature selection algorithm in data mining experiments. *Int J Adv Comput Sci Appl.* 2023;14(10). doi:10.14556/ijacsa.2023.0141054.
47. Kumar D, Pandey RC, Mishra AK. A review of image features extraction techniques and their applications in image forensic. *Multimed Tools Appl.* 2024;83(40):87801–902. doi:10.1007/s11042-023-17950-x.
48. Razzaghi P, Abbasi K, Bayat P. Learning spatial hierarchies of high-level features in deep neural network. *J Vis Commun Image Represent.* 2020;70:102817. doi:10.1016/j.jvcir.2020.102817.

49. Yu H, Yang LT, Zhang Q, Armstrong D, Deen MJ. Convolutional neural networks for medical image analysis: state-of-the-art, comparisons, improvement and perspectives. *Neurocomputing*. 2021;444(9):92–110. doi:10.1016/j.neucom.2020.04.157.
50. Afshar M, Usefi H. Optimizing feature selection methods by removing irrelevant features using sparse least squares. *Expert Syst Appl*. 2022;200(D1):116928. doi:10.1016/j.eswa.2022.116928.
51. Panati C, Wagner S, Brüggenwirth S. Feature relevance evaluation using grad-CAM, LIME and SHAP for deep learning SAR data classification. In: 2022 23rd International Radar Symposium (IRS); 2022 Sep 12–14; Gdansk, Poland: IEEE; 2022. p. 457–62.
52. Asif RN, Naseem MT, Ahmad M, Mazhar T, Khan MA, Khan MA, et al. Brain tumor detection empowered with ensemble deep learning approaches from MRI scan images. *Sci Rep*. 2025;15(1):15002. doi:10.1038/s41598-025-99576-7.
53. Shabbir M, Suhail Z, Hafeez N, Saqib N, Farooq M, Guizani S, et al. Prostate segmentation in MRI images using transfer learning based mask RCNN. *Curr Med Imaging*. 2024;20:e15734056305021. doi:10.2174/0115734056305021240603114137.
54. Li M, Jiang Y, Zhang Y, Zhu H. Medical image analysis using deep learning algorithms. *Front Public Health*. 2023;11:1273253. doi:10.3389/fpubh.2023.1273253.
55. Huda N, Ku-Mahamud KR. CNN-based image segmentation approach in brain tumor classification: a review. *Eng Proc*. 2025;84(1):66. doi:10.3390/engproc2025084066.
56. Asia AO, Zhu CZ, Althubiti SA, Al-Alimi D, Xiao YL, Ouyang PB, et al. Detection of diabetic retinopathy in retinal fundus images using CNN classification models. *Electronics*. 2022;11(17):2740. doi:10.3390/electronics11172740.
57. Sarvamangala DR, Kulkarni RV. Convolutional neural networks in medical image understanding: a survey. *Evol Intell*. 2022;15(1):1–22. doi:10.1007/s12065-020-00540-3.
58. Satpathy S, Khalaf OI, Shukla DK, Algburi S, Hamam H. Consumer electronics based smart technologies for enhanced terahertz healthcare having an integration of split learning with medical imaging. *Sci Rep*. 2024;14(1):10412. doi:10.1038/s41598-024-58741-0.
59. Berahmand K, Daneshfar F, Salehi ES, Li Y, Xu Y. Autoencoders and their applications in machine learning: a survey. *Artif Intell Rev*. 2024;57(2):28. doi:10.1007/s10462-023-10662-6.
60. Chen S, Guo W. Auto-encoders in deep learning—a review with new perspectives. *Mathematics*. 2023;11(8):1777. doi:10.3390/math11081777.
61. Tang C, Bian M, Liu X, Li M, Zhou H, Wang P, et al. Unsupervised feature selection via latent representation learning and manifold regularization. *Neural Netw*. 2019;117(9):163–78. doi:10.1016/j.neunet.2019.04.015.
62. Joyce JM. Kullback-leibler divergence. In: International encyclopedia of statistical science. Berlin/Heidelberg, Germany: Springer; 2011. p. 720–2. doi: 10.1007/978-3-642-04898-2_327.
63. Liu J, Li C, Yang W. Supervised learning via unsupervised sparse autoencoder. *IEEE Access*. 2018;6:73802–14. doi:10.1109/access.2018.2884697.
64. Lapucci M, Levato T, Rinaldi F, Sciandrone M. A unifying framework for sparsity-constrained optimization. *J Optim Theory Appl*. 2023;199(2):663–92. doi:10.1007/s10957-023-02306-0.
65. Tufail AB, Anwar N, Othman MTB, Ullah I, Khan RA, Ma YK, et al. Early-stage Alzheimer’s disease categorization using PET neuroimaging modality and convolutional neural networks in the 2D and 3D domains. *Sensors*. 2022;22(12):4609. doi:10.3390/s22124609.
66. Mohi ud din dar G, Bhagat A, Ansarullah SI, Ben Othman MT, Hamid Y, Alkahtani HK, et al. A novel framework for classification of different Alzheimer’s disease stages using CNN model. *Electronics*. 2023;12(2):469. doi:10.3390/electronics12020469.
67. Ehrhardt J, Wilms M. Autoencoders and variational autoencoders in medical image analysis. In: Biomedical image synthesis and simulation. Amsterdam: Elsevier; 2022. p. 129–62. doi: 10.1016/b978-0-12-824349-7.00015-3.

68. Naeem I, Ditta A, Mazhar T, Anwar M, Saeed MM, Hamam H. Voice biomarkers as prognostic indicators for Parkinson's disease using machine learning techniques. *Sci Rep.* 2025;15(1):12129. doi:10.1038/s41598-025-96950-3.
69. Safdarian N, Dabanloo NJ. Detection and classification of COVID-19 by lungs computed tomography scan image processing using intelligence algorithm. *J Med Signals Sens.* 2021;11(4):274–84. doi:10.4103/jmss.JMSS_55_20.
70. Khan H, Borah N, Begum SS, Alam A, Soudy M. Transformer networks and autoencoders in genomics and genetic data interpretation: a case study. In: Deep learning in genetics and genomics. Amsterdam: Elsevier; 2025. p. 399–423. doi: 10.1016/b978-0-443-27523-4.00004-4.
71. Khan A, Rauf Z, Khan AR, Rathore S, Khan SH, Shah NS, et al. A recent survey of vision transformers for medical image segmentation. *arXiv:2312.00634.* 2023.
72. Khan A, Rauf Z, Sohail A, Khan AR, Asif H, Asif A, et al. A survey of the vision transformers and their CNN-transformer based variants. *Artif Intell Rev.* 2023;56(3):2917–70. doi:10.1007/s10462-023-10595-0.
73. Mankki JJ, Bochenina K. Vision transformers in brain image segmentation [master thesis]. Helsinki, Finland: University of Helsinki; 2025.
74. Choi SR, Lee M. Transformer architecture and attention mechanisms in genome data analysis: a comprehensive review. *Biology.* 2023;12(7):1033. doi:10.3390/biology12071033.
75. Aburass S, Dorgham O, Al Shaqsi J, Abu Rumman M, Al-Kadi O. Vision transformers in medical imaging: a comprehensive review of advancements and applications across multiple diseases. *J Imag Inform Med.* 2025. doi:10.1007/s10278-025-01481-y.
76. Sundar GN, Narmadha D, Jerry NA, Thangavel SK, Shanmugam SK, Ajibesin AA. Brain tumor detection and classification using vision transformer (ViT). In: 2024 3rd International Conference on Automation, Computing and Renewable Systems (ICACRS); 2024 Dec 4–6; Pudukkottai, India. p. 562–7. doi:10.1109/ICACRS62842.2024.10841703.
77. Adebiyi A, Abdalnabi N, Simoes EJ, Becevic M, Smith EH, Rao P. Transformers in skin lesion classification and diagnosis: a systematic review. *medRxiv.* 2024. doi:10.1101/2024.09.19.24314004.
78. Jiao CN, Gao YL, Ge DH, Shang J, Liu JX. Multi-modal imaging genetics data fusion by deep auto-encoder and self-representation network for Alzheimer's disease diagnosis and biomarkers extraction. *Eng Appl Artif Intell.* 2024;130(5):107782. doi:10.1016/j.engappai.2023.107782.
79. Shamshad F, Khan S, Zamir SW, Khan MH, Hayat M, Khan FS, et al. Transformers in medical imaging: a survey. *Med Image Anal.* 2023;88(1):102802. doi:10.1016/j.media.2023.102802.
80. Ghazal TM, Abbas S, Munir S, Khan MA, Ahmad M, Issa GF, et al. Alzheimer disease detection empowered with transfer learning. *Comput Mater Contin.* 2022;70(3):5005–19. doi:10.32604/cmc.2022.020866.
81. Ersavas T, Smith MA, Mattick JS. Novel applications of convolutional neural networks in the age of transformers. *Sci Rep.* 2024;14(1):10000. doi:10.1038/s41598-024-60709-z.
82. Jogin M, Mohana, Madhulika MS, Divya GD, Meghana RK, Apoorva S. Feature extraction using convolution neural networks (CNN) and deep learning. In: 2018 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT); 2018 May 18–19; Bangalore, India. p. 2319–23. doi:10.1109/RTEICT42901.2018.9012507.
83. Maier A, Köstler H, Heisig M, Krauss P, Yang SH. Known operator learning and hybrid machine learning in medical imaging—a review of the past, the present, and the future. *Prog Biomed Eng.* 2022;4(2):022002. doi:10.1088/2516-1091/ac5b13.
84. Agarwal NB, Yadav DK. A comprehensive analysis of classical machine learning and modern deep learning methodologies. *Int J Eng Res Technol.* 2024;13(5):1–16. doi:10.17577/IJERTV13IS050275.
85. Odhiambo Omuya E, Onyango Okeyo G, Waema Kimwele M. Feature selection for classification using principal component analysis and information gain. *Expert Syst Appl.* 2021;174(11):114765. doi:10.1016/j.eswa.2021.114765.
86. Such FP, Madhavan V, Conti E, Lehman J, Stanley KO, Clune J. Deep neuroevolution: genetic algorithms are a competitive alternative for training deep neural networks for reinforcement learning. *arXiv:1712.06567.* 2017.

87. Haq I, Mazhar T, Malik MA, Kamal MM, Ullah I, Kim T, et al. Lung nodules localization and report analysis from computerized tomography (CT) scan using a novel machine learning approach. *Appl Sci.* 2022;12(24):12614. doi:10.3390/app122412614.
88. Nadeem MW, Ghamsi MAA, Hussain M, Khan MA, Khan KM, Almotiri SH, et al. Brain tumor analysis empowered with deep learning: a review, taxonomy, and future challenges. *Brain Sci.* 2020;10(2):118. doi:10.3390/brainsci10020118.
89. Nsibbi M, Manita G, Korbaa O. Advances in nature-inspired metaheuristic optimization for feature selection problem: a comprehensive survey. *Comput Sci Rev.* 2023;49(2):100559. doi:10.1016/j.cosrev.2023.100559.
90. Abdel-Basset M, Abdel-Fatah L, Sangaiah AK. Metaheuristic algorithms: a comprehensive review. In: Computational intelligence for multimedia big data on the cloud with engineering applications. Amsterdam: Elsevier; 2018. p. 185–231. doi: 10.1016/b978-0-12-813314-9.00010-4.
91. Marini F, Walczak B. Particle swarm optimization (PSO). A tutorial. *Chemom Intell Lab Syst.* 2015;149:153–65. doi:10.1016/j.chemolab.2015.08.020.
92. Yilmaz Eroglu D, Akcan U. An adapted ant colony optimization for feature selection. *Appl Artif Intell.* 2024;38(1):2335098. doi:10.1080/08839514.2024.2335098.
93. Venkatesh B, Anuradha J. A review of feature selection and its methods. *Cybern Inf Technol.* 2019;19(1):3–26. doi:10.2478/cait-2019-0001.
94. Xu Y, Yu Z, Cao W, Philip Chen CL. A novel classifier ensemble method based on subspace enhancement for high-dimensional data classification. *IEEE Trans Knowl Data Eng.* 2023;35(1):16–30. doi:10.1109/TKDE.2021.3087517.
95. Ros F, Guillaume S. From supervised instance and feature selection algorithms to dual selection: a review. In: Sampling techniques for supervised or unsupervised tasks. Cham: Springer International Publishing; 2019. p. 83–128. doi: 10.1007/978-3-030-29349-9_4.
96. Sagi O, Rokach L. Approximating XGBoost with an interpretable decision tree. *Inf Sci.* 2021;572(2–3):522–42. doi:10.1016/j.ins.2021.05.055.
97. Binder M, Moosbauer J, Thomas J, Bischl B. Multi-objective hyperparameter tuning and feature selection using filter ensembles. In: Proceedings of the 2020 Genetic and Evolutionary Computation Conference. Cancún, Mexico: ACM; 2020. p. 471–79. doi:10.1145/3377930.3389815.
98. Petrauskas V, Damuleviciene G, Dobrovolskis A, Dovydaitis J, Janaviciute A, Jasinevicius R, et al. XAI-based medical decision support system model. *Int J Sci Res Publ.* 2020;10(12):598–607. doi:10.29322/ijrsp.10.12.2020. p10869.
99. Hussain S, Mubeen I, Ullah N, Shah SSUD, Khan BA, Zahoor M, et al. Modern diagnostic imaging technique applications and risk factors in the medical field: a review. *Biomed Res Int.* 2022;2022(1):5164970. doi:10.1155/2022/5164970.
100. Wang Z, Yi R, Wen X, Zhu C, Xu K. Cardiovascular medical image and analysis based on 3D vision: a comprehensive survey. *Meta-Radiology.* 2024;2(4):100102. doi:10.1016/j.metrad.2024.100102.
101. Ray P, Reddy SS, Banerjee T. Various dimension reduction techniques for high dimensional data analysis: a review. *Artif Intell Rev.* 2021;54(5):3473–515. doi:10.1007/s10462-020-09928-0.
102. Han Q, Hu L, Gao W. Feature relevance and redundancy coefficients for multi-view multi-label feature selection. *Inf Sci.* 2024;652(1):119747. doi:10.1016/j.ins.2023.119747.
103. Chai W, Wang G. Deep vision multimodal learning: methodology, benchmark, and trend. *Appl Sci.* 2022;12(13):6588. doi:10.3390/app12136588.
104. Luo F, Wu D, Pino LR, Ding W. A novel multimodel medical image fusion framework with edge enhancement and cross-scale transformer. *Sci Rep.* 2025;15(1):11657. doi:10.1038/s41598-025-93616-y.
105. Lepcha DC, Dogra A, Goyal B, Chohan JS, Koundal D, Zagaria A, et al. Multimodal medical image fusion based on pixel significance using anisotropic diffusion and cross bilateral filter. *Hum-Cent Comput Inf Sci.* 2022;12. doi:10.22967/HCIS.2022.12.015.
106. Pawłowski M, Wróblewska A, Sysko-Romańczuk S. Effective techniques for multimodal data fusion: a comparative analysis. *Sensors.* 2023;23(5):2381. doi:10.3390/s23052381.

107. Jiao T, Guo C, Feng X, Chen Y, Song J. A comprehensive survey on deep learning multi-modal fusion: methods, technologies and applications. *Comput Mater Contin.* 2024;80(1):1–35. doi:10.32604/cmc.2024.053204.
108. Huang X, Ma T, Jia L, Zhang Y, Rong H, Alnabhan N. An effective multimodal representation and fusion method for multimodal intent recognition. *Neurocomputing.* 2023;548(2):126373. doi:10.1016/j.neucom.2023.126373.
109. Huang SC, Pareek A, Seyyedi S, Banerjee I, Lungren MP. Fusion of medical imaging and electronic health records using deep learning: a systematic review and implementation guidelines. *npj Digit Med.* 2020;3(1):136. doi:10.1038/s41746-020-00341-z.
110. Pereira LM, Salazar A, Vergara L. On comparing early and late fusion methods. In: *Advances in computational intelligence.* Cham: Springer Nature Switzerland; 2023. p. 365–78. doi: 10.1007/978-3-031-43085-5_29.
111. Wang Y, Xu X, Yu W, Xu R, Cao Z, Heng TS. Combine early and late fusion together: a hybrid fusion framework for image-text matching. In: *2021 IEEE International Conference on Multimedia and Expo (ICME); 2021 Jul 5–9; Shenzhen, China.* p. 1–6. doi:10.1109/ICME51207.2021.9428201.
112. Wang J, Yu L, Tian S. Cross-attention interaction learning network for multi-model image fusion via transformer. *Eng Appl Artif Intell.* 2025;139(5):109583. doi:10.1016/j.engappai.2024.109583.
113. Mienye ID, Swart TG, Obaido G. Recurrent neural networks: a comprehensive review of architectures, variants, and applications. *Information.* 2024;15(9):517. doi:10.3390/info15090517.
114. Dai Y, Gao Y, Liu F. TransMed: transformers advance multi-modal medical image classification. *Diagnostics.* 2021;11(8):1384. doi:10.3390/diagnostics11081384.
115. Islam S, Elmekki H, Elsebai A, Bentahar J, Drawel N, Rjoub G, et al. A comprehensive survey on applications of transformers for deep learning tasks. *Expert Syst Appl.* 2024;241(8):122666. doi:10.1016/j.eswa.2023.122666.
116. Hulsen T. Explainable artificial intelligence (XAI): concepts and challenges in healthcare. *AI.* 2023;4(3):652–66. doi:10.3390/ai4030034.
117. Bai Z, Wang M, Guo F, Guo Y, Cai C, Bie R, et al. SecMdp: towards privacy-preserving multimodal deep learning in end-edge-cloud. In: *2024 IEEE 40th International Conference on Data Engineering (ICDE); 2024 May 13–16; Utrecht, The Netherlands.* p. 1659–70. doi:10.1109/ICDE60146.2024.00135.
118. Mazhar T, Shah SFA, Inam SA, Awotunde JB, Saeed MM, Hamam H. Analysis of integration of IoMT with blockchain: issues, challenges and solutions. *Discov Internet Things.* 2024;4(1):21. doi:10.1007/s43926-024-00078-1.
119. Xia Q, Li Q. QuantumFed: a federated learning framework for collaborative quantum training. In: *2021 IEEE Global Communications Conference (GLOBECOM); 2021 Dec 7–11; Madrid, Spain.* p. 1–6. doi:10.1109/GLOBECOM46510.2021.9685012.
120. Shahwar T, Zafar J, Almogren A, Zafar H, Rehman A, Shafiq M, et al. Automated detection of Alzheimer's via hybrid classical quantum neural networks. *Electronics.* 2022;11(5):721. doi:10.3390/electronics11050721.
121. Hu S, Liao Z, Xia Y. Devil is in channels: contrastive single domain generalization for medical image segmentation. In: *Medical image computing and computer assisted intervention–MICCAI 2023.* Cham: Springer Nature Switzerland; 2023. p. 14–23. doi: 10.1007/978-3-031-43901-8_2.
122. Wang J, Ma H. Depth disentanglement strategy of latent space for medical image segmentation. *Biomed Signal Process Control.* 2024;92(1):106102. doi:10.1016/j.bspc.2024.106102.
123. Mandal AK, Chakraborty B. Quantum computing and quantum-inspired techniques for feature subset selection: a review. *Knowl Inf Syst.* 2025;67(3):2019–61. doi:10.1007/s10115-024-02282-5.
124. Islam MR, Lima AA, Das SC, Mridha MF, Prodeep AR, Watanobe Y. A comprehensive survey on the process, methods, evaluation, and challenges of feature selection. *IEEE Access.* 2022;10(8):99595–632. doi:10.1109/access.2022.3205618.
125. Koprinkov IG. The quantum superposition principle: a reconsideration. *arXiv:2311.02391.* 2023.
126. Lahoz-Beltra R. Quantum genetic algorithms for computer scientists. *Computers.* 2016;5(4):24. doi:10.3390/computers5040024.
127. Ardelean SM, Udrescu M. Hybrid quantum search with genetic algorithm optimization. *PeerJ Comput Sci.* 2024;10(3):e2210. doi:10.7717/peerj.cs.2210.
128. Basheer A, Afham A, Goyal SK. Quantum k -nearest neighbors algorithm. *arXiv:2003.09187.* 2020.

129. Wang H. A novel feature selection method based on quantum support vector machine. *Phys Scr.* 2024;99(5):056006. doi:10.1088/1402-4896/ad36ef.

130. Ding C, Bao TY, Huang HL. Quantum-inspired support vector machine. *IEEE Trans Neural Netw Learning Syst.* 2022;33(12):7210–22. doi:10.1109/tnnls.2021.3084467.

131. Qi H, Xiao S, Liu Z, Gong C, Gani A. Variational quantum algorithms: fundamental concepts, applications and challenges. *Quantum Inf Process.* 2024;23(6):224. doi:10.1007/s11128-024-04438-2.

132. Li P, Pei Y, Li J. A comprehensive survey on design and application of autoencoder in deep learning. *Appl Soft Comput.* 2023;138(7553):110176. doi:10.1016/j.asoc.2023.110176.

133. Kalamkar S, Geetha Mary A. Multimodal image fusion: a systematic review. *Decis Anal J.* 2023;9(3):100327. doi:10.1016/j.dajour.2023.100327.

134. Talaei Khoei T, Ould Slimane H, Kaabouch N. Deep learning: systematic review, models, challenges, and research directions. *Neural Comput Appl.* 2023;35(31):23103–24. doi:10.1007/s00521-023-08957-4.

135. Wolters C, Yang X, Schlichtmann U, Suzumura T. Memory is all you need: an overview of compute-in-memory architectures for accelerating large language model inference. *arXiv:2406.08413.* 2024.

136. Gadzicki K, Khamsehshari R, Zetzsche C. Early vs late fusion in multimodal convolutional neural networks. In: 2020 IEEE 23rd International Conference on Information Fusion (FUSION); 2020 Jul 6–9; Rustenburg, South Africa. p. 1–6.

137. Nekouie N, Romoozi M, Esmaeili M. A new evolutionary ensemble learning of multimodal feature selection from microarray data. *Neural Process Lett.* 2023;55(5):6753–80. doi:10.1007/s11063-023-11159-7.

138. Ji S, Tan Y, Saravirat T, Yang Z, Liu Y, Vasankari L, et al. Emerging trends in federated learning: from model fusion to federated X learning. *Int J Mach Learn Cyber.* 2024;15(9):3769–90. doi:10.1007/s13042-024-02119-1.

139. Scheibner J, Raisaro JL, Troncoso-Pastoriza JR, Ienca M, Fellay J, Vayena E, et al. Revolutionizing medical data sharing using advanced privacy-enhancing technologies: technical, legal, and ethical synthesis. *J Med Internet Res.* 2021;23(2):e25120. doi:10.2196/25120.

140. Valdez F, Melin P. A review on quantum computing and deep learning algorithms and their applications. *Soft Comput.* 2023;2022(18):1–20. doi:10.1007/s00500-022-07037-4.

141. Nau MA, Nutricati LA, Camino B, Warburton PA, Maier AK. Quantum annealing feature selection on light-weight medical image datasets. *arXiv:2502.19201.* 2025.

142. Al-Antari MA. Artificial intelligence for medical diagnostics-existing and future AI technology!. *Diagnostics.* 2023;13(4):688. doi:10.3390/diagnostics13040688.

143. Marey A, Arjmand P, Alerab ADS, Eslami MJ, Saad AM, Sanchez N, et al. Explainability, transparency and black box challenges of AI in radiology: impact on patient care in cardiovascular radiology. *Egypt J Radiol Nucl Med.* 2024;55(1):183. doi:10.1186/s43055-024-01356-2.

144. Sharma NA, Chand RR, Buksh Z, Ali ABMS, Hanif A, Beheshti A. Explainable AI frameworks: navigating the present challenges and unveiling innovative applications. *Algorithms.* 2024;17(6):227. doi:10.3390/al17060227.

145. Linardatos P, Papastefanopoulos V, Kotsiantis S. Explainable AI: a review of machine learning interpretability methods. *Entropy.* 2020;23(1):18. doi:10.3390/e23010018.

146. Ali S, Abuhmed T, El-Sappagh S, Muhammad K, Alonso-Moral JM, Confalonieri R, et al. Explainable artificial intelligence (XAI): what we know and what is left to attain trustworthy artificial intelligence. *Inf Fusion.* 2023;99(3):101805. doi:10.1016/j.inffus.2023.101805.

147. Letoffe O, Huang X, Asher N, Marques-Silva J. From SHAP scores to feature importance scores. *arXiv:2405.11766.* 2024.

148. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: visual explanations from deep networks via gradient-based localization. In: 2017 IEEE International Conference on Computer Vision (ICCV); 2017 Oct 22–29; Venice, Italy. p. 618–26. doi:10.1109/ICCV.2017.74.

149. Singh R, Bengani V, Saini K. Hybrid learning systems: integrating traditional machine learning with deep learning techniques. 2024 [cited 2025 Jun 10]. Available at: <https://www.researchgate.net/publication/380910904>.

150. Yu C, Tan J, Cheng Y, Mi X. Data analysis and preprocessing techniques for air quality prediction: a survey. *Stoch Environ Res Risk Assess.* 2024;38(6):2095–117. doi:10.1007/s00477-024-02693-4.

151. Theng D, Bhoyar KK. Feature selection techniques for machine learning: a survey of more than two decades of research. *Knowl Inf Syst.* 2024;66(3):1575–637. doi:10.1007/s10115-023-02010-5.
152. Liu Z, Yang J, Wang L, Chang Y. A novel relation aware wrapper method for feature selection. *Pattern Recognit.* 2023;140(1):109566. doi:10.1016/j.patcog.2023.109566.
153. Li J, Cheng K, Wang S, Morstatter F, Trevino RP, Tang J, et al. Feature selection. *ACM Comput Surv.* 2018;50(6):1–45. doi:10.1145/3136625.
154. Chen K, Xue B, Zhang M, Zhou F. Evolutionary multitasking for feature selection in high-dimensional classification via particle swarm optimization. *IEEE Trans Evol Comput.* 2022;26(3):446–60. doi:10.1109/TEVC.2021.3100056.
155. Dokeroglu T, Deniz A, Kiziloz HE. A comprehensive survey on recent metaheuristics for feature selection. *Neurocomputing.* 2022;494(13):269–96. doi:10.1016/j.neucom.2022.04.083.
156. Teoh JR, Dong J, Zuo X, Lai KW, Hasikin K, Wu X. Advancing healthcare through multimodal data fusion: a comprehensive review of techniques and applications. *PeerJ Comput Sci.* 2024;10(9):e2298. doi:10.7717/peerj-cs.2298.
157. Bayoudh K. A survey of multimodal hybrid deep learning for computer vision: architectures, applications, trends, and challenges. *Inf Fusion.* 2024;105:102217. doi:10.1016/j.inffus.2023.102217.
158. Santosh KC, Rizk R, Bajracharya SK. Understanding data—modalities and preprocessing. In: Cracking the machine learning code: technicality or innovation? Singapore: Springer Nature Singapore; 2024. p. 13–24. doi: 10.1007/978-981-97-2720-9_2.
159. Mittermaier M, Raza MM, Kvedar JC. Bias in AI-based models for medical applications: challenges and mitigation strategies. *npj Digit Med.* 2023;6(1):113. doi:10.1038/s41746-023-00858-z.
160. Hanna MG, Pantanowitz L, Jackson B, Palmer O, Visweswaran S, Pantanowitz J, et al. Ethical and bias considerations in artificial intelligence/machine learning. *Mod Pathol.* 2025;38(3):100686. doi:10.1016/j.modpat.2024.100686.
161. Alvarez JM, Colmenarejo AB, Elbaid A, Fabbrizzi S, Fahimi M, Ferrara A, et al. Policy advice and best practices on bias and fairness in AI. *Ethics Inf Technol.* 2024;26(2):31. doi:10.1007/s10676-024-09746-w.
162. Min A. Artificial intelligence and bias: challenges, implications, and remedies. *J Soc Res.* 2023;2(11):3808–17. doi:10.55324/josr.v2i11.1477.
163. Fletcher RR, Nakashimana A, Olubeko O. Addressing fairness, bias, and appropriate use of artificial intelligence and machine learning in global health. *Front Artif Intell.* 2021;3:561802. doi:10.3389/frai.2020.561802.
164. Balasubramaniam N, Kauppinen M, Rannisto A, Hiekkonen K, Kujala S. Transparency and explainability of AI systems: from ethical guidelines to requirements. *Inf Softw Technol.* 2023;159(4):107197. doi:10.1016/j.infsof.2023.107197.
165. Murdoch B. Privacy and artificial intelligence: challenges for protecting health information in a new era. *BMC Med Ethics.* 2021;22(1):122. doi:10.1186/s12910-021-00687-3.
166. Mennella C, Maniscalco U, De Pietro G, Esposito M. Ethical and regulatory challenges of AI technologies in healthcare: a narrative review. *Heliyon.* 2024;10(4):e26297. doi:10.1016/j.heliyon.2024.e26297.
167. Habli I, Lawton T, Porter Z. Artificial intelligence in health care: accountability and safety. *Bull World Health Organ.* 2020;98(4):251–6. doi:10.2471/BLT.19.237487.
168. Bohmrah MK, Kaur H. Advanced hybridization and optimization of DNNs for medical imaging: a survey on disease detection techniques. *Artif Intell Rev.* 2025;58(4):122. doi:10.1007/s10462-024-11049-x.
169. Perniciano A, Loddo A, Di Ruberto C, Pes B. Insights into radiomics: impact of feature selection and classification. *Multimed Tools Appl.* 2025;84(26):31695–721. doi:10.1007/s11042-024-20388-4.
170. Zouache D, Got A, Alarabiat D, Abualigah L, Talbi EG. A novel multi-objective wrapper-based feature selection method using quantum-inspired and swarm intelligence techniques. *Multimed Tools Appl.* 2024;83(8):22811–35. doi:10.1007/s11042-023-16411-9.
171. Wang J, Zhang Z, Wang Y. Utilizing feature selection techniques for AI-driven tumor subtype classification: enhancing precision in cancer diagnostics. *Biomolecules.* 2025;15(1):81. doi:10.3390/biom15010081.

172. Kanya Kumari L, Naga Jagadesh B. An adaptive teaching learning based optimization technique for feature selection to classify mammogram medical images in breast cancer detection. *Int J Syst Assur Eng Manag.* 2024;15(1):35–48. doi:10.1007/s13198-021-01598-7.
173. Singh LK, Khanna M, Monga H, Singh R, Pandey G. Nature-inspired algorithms-based optimal features selection strategy for COVID-19 detection using medical images. *New Gener Comput.* 2024;42(4):761–824. doi:10.1007/s00354-024-00255-4.
174. Ghadi YY, Saqib SM, Mazhar T, Almogren A, Waheed W, Altameem A, et al. Explainable AI analysis for smog rating prediction. *Sci Rep.* 2025;15(1):8070. doi:10.1038/s41598-025-92788-x.