Journal of Computing and Data Technology
https://journal.krrishpub.com/index.php/jcdt/index

REVIEW ARTICLE

A Systematic Review on Hybrid Deep Learning and Metaheuristic Optimization Techniques for Medical Image Segmentation and Classification

Potharla Ramadevi [a,*], Bala Krishnama Manohar [b], N. Bhargavi [b], Valluru Prathyusha [c,*], P. Shailaja [d]

[a] Department of Basic Science and Humanities, Vaagdevi College of Engineering, Warangal- 506 005, India.
[b] Department of Mathematics & Statistics, Vignan's Foundation for Science, Technology & Research, Guntur- 522 213, India.
[c] Department of Computer Science and Engineering, Vishnu Institute of Technology, Bhimavaram- 534 202, India.
[d] Department of Computer Science and Engineering, Vaagdevi College of Engineering, Warangal- 506 005, India.

Abstract

Medical image segmentation and classification have become fundamental components of modern intelligent healthcare systems due to their ability to support early disease diagnosis, treatment planning, prognosis evaluation, and computer-aided clinical decision-making. Recent advances in deep learning have significantly improved the capability of automated medical image analysis systems across multiple imaging modalities including X-ray, computed tomography, magnetic resonance imaging, ultrasound, histopathology, and retinal imaging. However, despite remarkable performance improvements, conventional deep learning models still face several limitations related to hyperparameter tuning, computational complexity, convergence instability, local minima stagnation, class imbalance, limited interpretability, and poor generalization under heterogeneous clinical conditions. To address these challenges, hybrid frameworks integrating deep learning with metaheuristic optimization techniques have emerged as a promising research direction for improving segmentation accuracy, classification robustness, feature optimization, and computational efficiency. This review article covers the historical evolution of intelligent medical image analysis, theoretical foundations of optimization-driven learning, taxonomy of deep learning architectures, and the role of evolutionary and swarm-based optimization algorithms including genetic algorithms, particle swarm optimization, grey wolf optimizer, firefly algorithm, whale optimization, ant colony optimization, Bayesian optimization, and reptile search optimization. Comparative analysis of datasets, evaluation metrics, computational complexity, convergence behavior, and clinical deployment challenges is also presented. Finally, open research challenges and future directions are identified toward trustworthy, interpretable, scalable, and autonomous AI-driven medical imaging systems for next-generation intelligent healthcare applications.

Keywords: Medical image segmentation, Medical image classification, Deep learning, Metaheuristic optimization, Artificial Intelligence (AI), Vision transformers, Explainable Artificial Intelligence (XAI), Healthcare AI.

Article information:
ISSN: 3107-9466 (Online)
Published by: **Krrish Scientific Publications Pvt. Ltd.**
DOI: https://doi.org/10.71426/jcdt.v2.i1.pp151-164
Received: *30 Apr. 2026* | Revised: *15 Jun. 2026* | Accepted: *18 Jun. 2026* | Published: *20 Jun. 2026*
Copyright ©2026 Author(s).
This is an open-access article distributed under the Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).

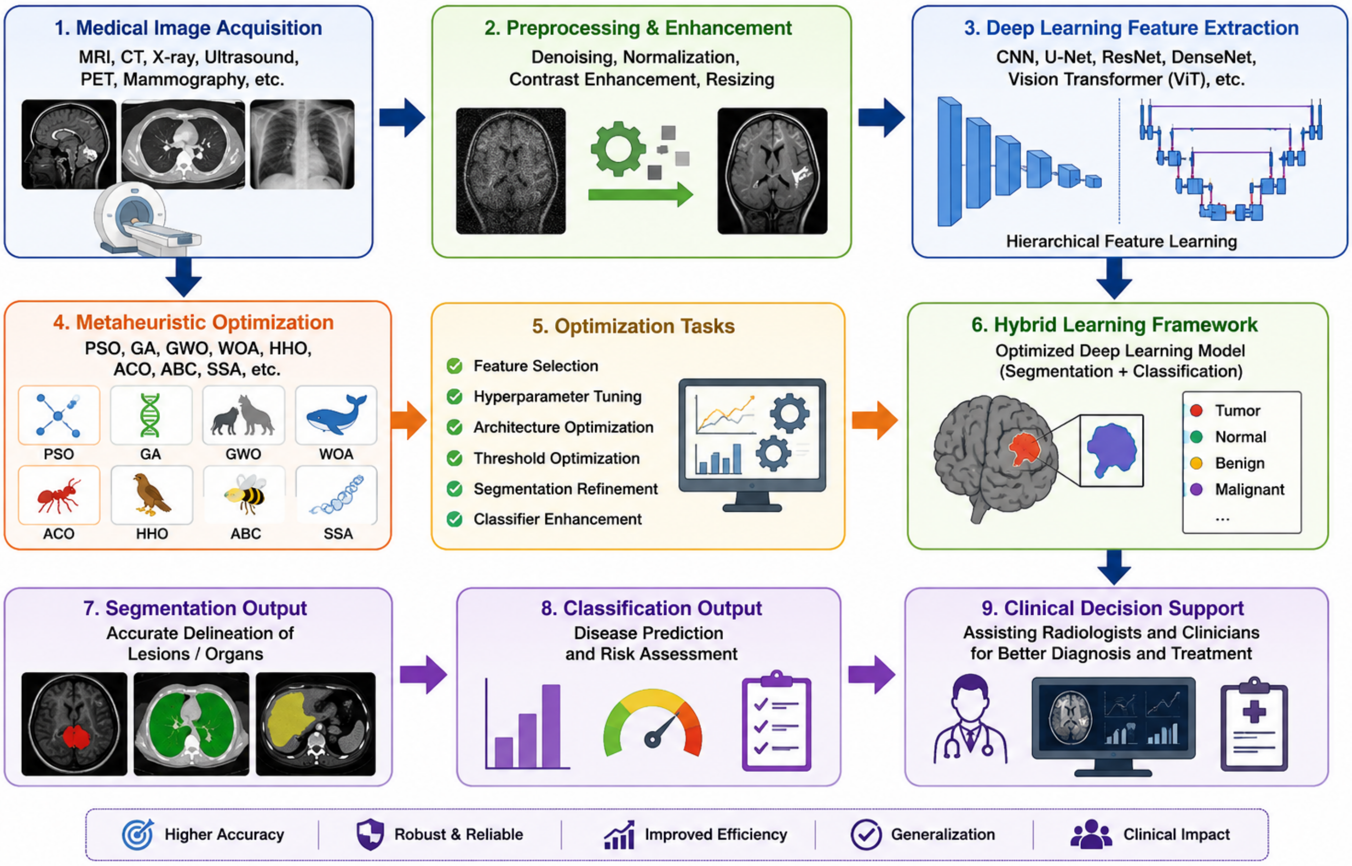
1. Introduction

Medical imaging has become one of the most important technological components of modern healthcare systems due to its critical role in disease diagnosis, treatment planning, surgical guidance, and clinical decision support [1], [2], [39], [42]. Advanced medical imaging modalities including magnetic resonance imaging (MRI), computed tomography (CT), ultrasound imaging, positron emission tomography (PET), histopathological imaging, and chest radiography provide large volumes of clinically significant anatomical and pathological information for healthcare professionals [16], [42]–[44]. The increasing availability of digital healthcare imaging data has significantly accelerated the development of intelligent computer-aided diagnostic systems capable of improving diagnostic efficiency and reducing clinical workload.

Traditional medical image analysis methods relied heavily on handcrafted feature engineering, thresholding tech-

*Corresponding author
Email addresses: ramadevi.p@vaagdevi.edu.in, potharlaramadevi@gmail.com (P Ramadevi), bmanoharngp@gmail.com, bm_maths@vignan.ac.in (BK Manohar), bn_maths@vignan.ac.in, bhargavi.nmaths@gmail.com (N. Bhargavi), pratyusha.v@vishnu.edu.in, passpartyu75@gmail.com (V Prathyusha), shylaja_p@vaagdevi.edu.in (P. Shailaja).

Graphical abstract: Medical image segmentation and classification.



niques, texture descriptors, and statistical pattern recognition frameworks [47]–[52]. Although conventional machine learning methods demonstrated promising performance for certain diagnostic tasks, their effectiveness was limited by dependence on domain expertise, poor scalability, limited feature representation capability, sensitivity to image variability, and inability to capture hierarchical semantic information. Consequently, automated feature learning frameworks became increasingly necessary for handling the growing complexity and dimensionality of medical imaging datasets.

The emergence of deep learning has fundamentally transformed medical image analysis into a data-driven intelligent healthcare paradigm [1], [2], [37], [39]. Deep neural architectures automatically learn hierarchical feature representations directly from raw image data without requiring manual feature extraction. Convolutional neural networks (CNNs), encoder–decoder segmentation frameworks, residual learning systems, and transformer-based architectures have demonstrated remarkable performance across numerous healthcare imaging applications [3], [8], [12], [13], [37]. The deep learning systems have achieved substantial improvements in tumor segmentation, disease classification, lesion localization, organ delineation, multimodal healthcare analysis [17]–[19].

CNN-based architectures such as ResNet, DenseNet, EfficientNet, and U-Net have become dominant frameworks for healthcare imaging systems due to their strong representation learning capability and end-to-end optimization mechanisms [3], [8], [9], [29]. Similarly, transformer-based architectures introduced global contextual feature learning capability through self-attention mechanisms, thereby improving multimodal healthcare image analysis performance [12], [13], [56], [57].

Despite their remarkable success, deep learning models still face several optimization-related challenges in healthcare applications. Training deep neural networks involves optimization over extremely high-dimensional nonlinear parameter spaces containing millions of trainable variables [32], [35], [37]. Consequently, deep learning systems often suffer from unstable convergence, local minima stagnation, overfitting, hyperparameter sensitivity, computational complexity, limited generalization capability. Furthermore, medical imaging datasets frequently exhibits class imbalance, heterogeneous imaging conditions, noisy annotations, limited labeled data and multimodal variability. These limitations significantly affect segmentation and classification performance in real-world clinical environments. To address these challenges, metaheuristic optimization algorithms have emerged as powerful computational approaches for improving optimization efficiency and convergence stability in intelligent healthcare imaging systems [20]–[24]. Metaheuristic algorithms are population-based stochastic optimization methods inspired by biological evolution, swarm intelligence, and collective search behavior. These techniques perform efficient global search over nonlinear optimization landscapes without requiring explicit gradient information. Popular optimization algorithms used

List of acronyms.

Acronym	Expansion
AI	Artificial Intelligence
AUC	Area Under the Curve
CAD	Computer-Aided Diagnosis
CNN	Convolutional Neural Network
CT	Computed Tomography
DE	Differential Evolution
DL	Deep Learning
DSC	Dice Similarity Coefficient
FA	Firefly Algorithm
FL	Federated Learning
GA	Genetic Algorithm
GAN	Generative Adversarial Network
GPU	Graphics Processing Unit
Grad-CAM	Gradient-weighted Class Activation Mapping
GWO	Grey Wolf Optimizer
BO	Bayesian Optimization
IoU	Intersection over Union
MRI	Magnetic Resonance Imaging
PET	Positron Emission Tomography
PSO	Particle Swarm Optimization
ReLU	Rectified Linear Unit
ROC	Receiver Operating Characteristic
SIFT	Scale-Invariant Feature Transform
SHAP	SHapley Additive exPlanations
SVM	Support Vector Machine
U-Net	Encoder–Decoder Segmentation Network
ViT	Vision Transformer
WOA	Whale Optimization Algorithm
XAI	Explainable Artificial Intelligence

in healthcare imaging are GA, PSO, GWO, WOA, FA, DE, and BO. The general optimization objective can be formulated as (1).

$$\theta^* = \arg\min_{\theta} \mathcal{L}(\theta), \tag{1}$$

Metaheuristic optimization algorithms are extensively integrated with deep learning architectures for hyperparameter tuning, feature selection, segmentation refinement, architecture optimization, multimodal fusion and ensemble learning optimization [20]–[25], [34].

The PSO and GA are widely employed for CNN parameter tuning and feature optimization due to their efficient exploration capability [20], [25], [34]. Similarly, grey wolf optimizer and whale optimization algorithm demonstrated strong performance in segmentation refinement and convergence stabilization tasks [21], [23].

Hybrid deep learning and optimization frameworks therefore combine, (i) automated hierarchical representation learning, (ii) global optimization capability, (iii) robust parameter tuning, (iv) intelligent feature selection, (v) optimization-aware segmentation. These frameworks have demonstrated promising performance across numerous healthcare imaging applications including, lung disease diagnosis [18], brain tumor segmentation [19], [69], diabetic retinopathy detection, skin lesion classification [61], histopathological cancer analysis [62] and COVID-19 screening systems [63]–[67].

Table 1: Mathematical symbols and notations.

Symbol	Description
I	Input medical image
K	Convolution kernel
$(I * K)$	Convolution operation
x, y	Spatial image coordinates
θ	Trainable neural network parameters
θ^*	Optimal parameter set
f_θ	Parameterized deep learning model
y	Predicted output class
$\hat{S}$	Predicted segmentation mask
P	Predicted pixel set
G	Ground-truth pixel set
$Dice$	Dice Similarity Coefficient
IoU	Intersection over Union
Q	Query matrix
K	Key matrix
V	Value matrix
d_k	Dimension of key vectors
$\mathcal{L}(\theta)$	Loss/objective function
$\mathcal{L}_{CE}$	Cross-entropy loss
$\mathcal{L}_{Dice}$	Dice loss function
x_i^t	Particle position at iteration t
v_i^{t+1}	Particle velocity update
F_i	Feature representation
F_{fusion}	Fused multimodal representation
w_i	Fusion weight coefficient
TP	True positive samples
TN	True negative samples
FP	False positive samples
FN	False negative samples
$Accuracy$	Classification accuracy metric
IR	Imbalance ratio
$N_{majority}$	Number of majority-class samples
$N_{minority}$	Number of minority-class samples
w^t	Global model parameters
w_k^t	Local client model parameters
n_k	Local dataset size
$H(x)$	Residual mapping
$F(x)$	Residual function
z_i	Softmax input activation
$P(y = i)$	Predicted class probability
$Complexity$	Computational complexity
$X(t)$	Candidate solution
$X^*(t)$	Best candidate solution
D'	Distance control parameter
b	Spiral control constant
l	Spiral movement parameter
μ	Mean intensity value
σ	Standard deviation
I_{norm}	Normalized image
$\phi(I)$	Feature extraction mapping

Recent advances in explainable artificial intelligence (XAI), federated learning, and multimodal healthcare intelligence have further accelerated the development of clinically reliable and trustworthy medical imaging systems [26], [27], [58]. Explainability mechanisms such as Grad-CAM and SHAP improve transparency by identifying diagnostically important image regions, thereby supporting physician trust and clinical interpretability [59], [60]. Federated learning frameworks additionally enable collaborative distributed model training while preserving patient privacy

Table 2: Major mathematical expressions and notations.

Mathematical expression	Description	Application
$y = f_\theta(I)$	Deep learning prediction framework	Medical image classification
$(I * K)(x, y) = \sum_m \sum_n I(x - m, y - n)K(m, n)$	Convolution operation	CNN feature extraction
$\hat{S} = f_\theta(I)$	Segmentation prediction framework	Biomedical image segmentation
$Dice = \frac{2\|P \cap G\|}{\|P\|+\|G\|}$	Dice Similarity Coefficient	Segmentation evaluation
$IoU = \frac{\|P \cap G\|}{\|P \cup G\|}$	Intersection over Union	Segmentation overlap analysis
$Attention(Q, K, V) = Softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V$	Self-attention mechanism	Transformer architectures
$\theta^* = \arg\min_\theta \mathcal{L}(\theta)$	Optimization objective function	Parameter optimization
$\mathcal{L}_{CE} = -\sum_{i=1}^{C} y_i \log(\hat{y}_i)$	Cross-entropy loss function	Classification optimization
$\mathcal{L}_{Dice} = 1 - \frac{2\|P \cap G\|}{\|P\|+\|G\|}$	Dice loss function	Segmentation optimization
$x_i^{t+1} = x_i^t + v_i^{t+1}$	PSO particle update equation	Swarm intelligence optimization
$F_{fusion} = \sum_{i=1}^{N} w_i F_i$	Multimodal feature fusion	Multimodal healthcare analysis
$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$	Classification accuracy metric	Diagnostic evaluation
$IR = \frac{N_{majority}}{N_{minority}}$	Imbalance ratio	Dataset imbalance analysis
$w^{t+1} = \sum_{k=1}^{K} \frac{n_k}{n} w_k^t$	Federated learning aggregation	Distributed healthcare learning
$H(x) = F(x) + x$	Residual learning framework	ResNet architectures
$P(y = i) = \frac{e^{z_i}}{\sum_{j=1}^{C} e^{z_j}}$	Softmax classification function	Multi-class prediction
$Complexity = O(n^2)$	Transformer computational complexity	Self-attention computation
$X(t + 1) = D'e^{bl}\cos(2\pi l) + X^*(t)$	WOA update	Metaheuristic optimization
$I_{norm} = \frac{I-\mu}{\sigma}$	Image normalization equation	Medical image preprocessing
$x = \phi(I)$	Feature extraction function	Traditional machine learning

and healthcare data confidentiality [27], [71].

Although several review papers have discussed either deep learning architectures or optimization algorithms independently, a comprehensive review focusing specifically on hybrid deep learning and metaheuristic optimization frameworks for medical image segmentation and classification remains limited. Existing surveys often focus only on segmentation or classification, emphasize individual optimization algorithms, lack systematic taxonomy-based analysis, provide limited comparative discussion of hybrid systems. Therefore, this review presents a concise yet comprehensive survey of hybrid deep learning and metaheuristic optimization techniques for medical image segmentation and classification. The major contributions of this review are summarized as follows:

1. A systematic overview of deep learning architectures used in medical image analysis.
2. A comprehensive discussion of metaheuristic optimization algorithms for healthcare imaging systems.
3. A taxonomy-driven analysis of hybrid deep learning and optimization frameworks.
4. Comparative analysis of segmentation and classification applications across multiple imaging modalities.
5. Discussion of major challenges, research gaps, and future directions for optimization-aware intelligent healthcare systems.

In this work, Table 1 lists the mathematical symbols and notations employed in the paper and Table 2 summarizes the major mathematical expressions used in medical image segmentation and classification.

2. Historical development of AI-based medical image analysis

The development of AI-based medical image analysis has evolved through multiple technological phases ranging from conventional image processing techniques to advanced deep learning and optimization-driven intelligent healthcare systems [1], [2], [37], [39]. Medical imaging has historically been one of the most computationally demanding healthcare domains due to the complexity of anatomical structures, pathological variability, multimodal image acquisition, and the requirement for highly accurate diagnostic interpretation. Early medical image analysis systems primarily relied on classical image processing and statistical pattern recognition approaches [49]–[52]. These methods focused on manually engineered features such as texture descriptors, histogram analysis, edge detection, thresholding, morphological operations.

Traditional segmentation and classification frameworks commonly employed:

- K-means clustering;
- fuzzy logic systems;
- support vector machines (SVM);
- nearest-neighbor classifiers;
- random forests.

Threshold-based segmentation methods such as Otsu thresholding became popular due to their simplicity and computational efficiency [50]. Similarly, texture feature extraction approaches based on Haralick descriptors were extensively used for tumor characterization and tissue classification [49]. Although these methods demonstrated promising performance in controlled experimental settings, their effectiveness was limited by handcrafted feature dependency, sensitivity to noise, poor scalability, limited generalization capability.

The rapid growth of digital imaging technologies during the 1990s and early 2000s accelerated the development of machine learning-based healthcare imaging systems. Statistical learning algorithms such as support vector machines and random forests improved automated disease classification performance by learning discriminative feature boundaries from training data [47], [48]. Feature extraction frameworks including scale-invariant feature transform (SIFT), histogram-based descriptors, and texture analysis techniques became widely adopted for medical image classification tasks [49], [51].

The traditional machine learning framework can be expressed as (2).

$$y = f(x) \tag{2}$$

In (2), x denotes handcrafted feature vectors, $f(\cdot)$ represents the learning model and y denotes diagnostic predictions.

However, conventional machine learning methods still depended heavily on manually engineered features and domain expertise. Consequently, these systems struggled to effectively model highly nonlinear and high-dimensional healthcare imaging data.

The emergence of deep learning marked a major turning point in medical image analysis [37], [38]. Deep neural networks introduced automated hierarchical representation learning capability, thereby eliminating the need for handcrafted feature engineering. Early neural network frameworks such as LeNet demonstrated the feasibility of convolution-based feature learning for image recognition tasks.

The success of AlexNet in large-scale image recognition significantly accelerated the adoption of deep learning across healthcare imaging applications [37], [46]. The CNN architectures subsequently became dominant frameworks for disease classification, lesion detection, segmentation and multimodal imaging analysis. Residual learning and dense connectivity architectures further improved deep neural optimization capability by addressing gradient degradation and feature propagation limitations [8], [9]. ResNet introduced shortcut identity mappings that enabled extremely deep network training, whereas DenseNet improved feature reuse and parameter efficiency through dense interlayer connectivity. The rapid development of encoder–decoder segmentation architectures represented another major milestone in healthcare imaging systems. Fully convolutional networks enabled dense semantic segmentation by replacing fully connected layers with convolution operations [4]. Subsequently, U-Net became one of the most influential biomedical segmentation frameworks due to its encoder–decoder structure and skip-connection mechanisms [3]. The segmentation process is generally formulated as (3).

$$\hat{S} = f_\theta(I) \tag{3}$$

Advanced segmentation architectures including U-Net++, Attention U-Net, SegNet, DeepLab, and V-Net substantially improved localization precision and segmentation robustness across multiple healthcare imaging applications [5]–[7], [10], [11], [30].

Deep learning frameworks rapidly achieved state-of-the-art performance in numerous clinical tasks including, brain tumor segmentation, lung disease diagnosis, retinal vessel extraction, histopathological cancer analysis, skin lesion classification [17]–[19], [41]. Nature-scale healthcare AI systems demonstrated diagnostic performance comparable to expert clinicians in several domains [17], [18]. For example, deep learning models achieved dermatologist-level skin cancer classification and highly accurate lung cancer screening performance using CT imaging. The emergence of transformer-based architectures introduced a new paradigm in medical image analysis through global contextual feature learning capability [12], [13]. Transformers utilize self-attention mechanisms for modeling long-range dependencies between image regions.

Transformer-based healthcare imaging systems demonstrated promising performance for multimodal image analysis, anatomical structure segmentation, disease localization, image reconstruction [56], [57]. Despite these remarkable advancements, deep learning systems introduced several optimization-related challenges due to the complexity of training highly nonlinear neural architectures. Consequently, metaheuristic optimization algorithms became increasingly integrated into healthcare imaging frameworks for improving convergence stability, feature optimization, hyperparameter tuning and segmentation refinement.

Swarm intelligence and evolutionary optimization algorithms including PSO, GA, WOA, and differential evolution became highly effective for optimization-aware healthcare imaging systems [20]–[24]. These techniques improves segmentation accuracy, classification robustness, optimization efficiency, and generalization capability. The optimization objective is generally represented by (4).

$$\theta^* = \arg\min_\theta \mathcal{L}(\theta) \tag{4}$$

Recent developments in healthcare AI additionally focuses on explainable artificial intelligence, federated learning, privacy-preserving healthcare intelligence, multimodal learning systems and trustworthy AI frameworks [26], [27], [58]. Explainability mechanisms such as Grad-CAM and SHAP improve clinical transparency by highlighting diagnostically important image regions [59], [60]. Federated learning frameworks further enable collaborative distributed healthcare model training while preserving patient privacy and regulatory compliance [27].

Figure 1 illustrates the historical evolution of AI-based medical image analysis systems. The historical evolution of AI-based medical image analysis demonstrates a progressive transition from handcrafted feature engineering toward fully automated optimization-driven intelligent healthcare systems. Classical machine learning frameworks provided foundational statistical learning capability but were limited by manual feature dependency and scalability constraints. Deep learning architectures fundamentally transformed healthcare imaging through automated hierarchical representation learning and end-to-end optimization mechanisms. Encoder–decoder segmentation frameworks, residual learning systems, and transformer-based architectures substantially improved disease diagnosis, lesion localization, and segmentation performance. More recently, the integration of metaheuristic optimization algorithms has enabled the development of highly robust hybrid healthcare imaging systems capable of improving convergence stability, segmentation precision, and optimization efficiency. Current research trends increasingly focus on explainable, trustworthy, privacy-preserving, and multimodal intelligent healthcare systems suitable for real-world clinical deployment.

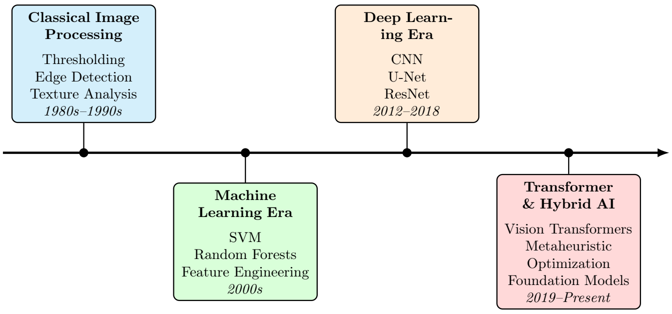


Figure 1: Evolution of artificial intelligence techniques for medical image analysis.

3. Theoretical foundations of medical image segmentation and classification

Medical image segmentation and classification constitute the fundamental computational tasks in intelligent healthcare imaging systems due to their critical role in automated disease diagnosis, lesion localization, anatomical structure delineation, and computer-aided clinical decision support [1], [2].

Segmentation focuses on identifying and separating clinically important image regions, whereas classification aims to assign diagnostic labels to medical images or segmented pathological structures. The increasing complexity of multimodal healthcare imaging data has significantly accelerated the development of advanced deep learning and optimization-driven analytical frameworks capable of extracting discriminative semantic representations from medical images [12], [16]. Modern medical image analysis systems generally combines the image preprocessing, feature extraction, segmentation, classification, optimization, clinical interpretation [68], [70]

Figure 2 illustrates the general theoretical pipeline of medical image segmentation and classification systems.

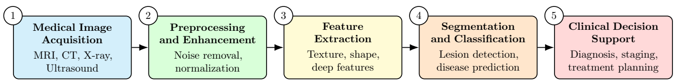


Figure 2: Theoretical pipeline of medical image segmentation and classification systems from image acquisition to clinical decision support.

3.1. Medical image preprocessing

Medical image preprocessing is an essential step for improving image quality and enhancing diagnostically relevant structures before segmentation and classification [15], [16]. Healthcare imaging data frequently suffer from noise, intensity inhomogeneity, low contrast, imaging artifacts and modality-specific distortions. Preprocessing frameworks generally involves normalization, denoising, contrast enhancement, histogram equalization and image augmentation. The normalization operation is expressed by (5).

$$I_{norm} = \frac{I - \mu}{\sigma} \tag{5}$$

In (5), I denotes the input image, μ denotes mean intensity, σ denotes standard deviation.

Image augmentation techniques further improve generalization capability by artificially increasing training data diversity through rotation, scaling, flipping, translation, intensity variation. Augmentation frameworks significantly reduce overfitting and improve robustness of deep neural architectures [56].

3.2. Feature extraction and representation learning

Feature extraction is one of the most important components of healthcare image analysis systems because diagnostic performance depends heavily on the quality of learned image representations [37], [38]. Traditional medical imaging systems relied on handcrafted descriptors including, texture features, histogram descriptors, edge information, and morphological representations. Haralick texture descriptors became highly influential for pathological tissue characterization and texture-based classification [49]. Similarly, SIFT descriptors enabled scale-invariant image feature representation for medical image retrieval and object recognition tasks [51]. The traditional feature extraction framework is represented as (6).

$$x = \phi(I) \tag{6}$$

In (6), I denotes the input image, $\phi(\cdot)$ denotes feature extraction operation, x represents extracted feature vectors.

Deep learning fundamentally transformed feature extraction through automated hierarchical representation learning capability [1], [37]. CNN architectures automatically learn multilevel semantic features directly from image data using convolution operations. The convolution process is mathematically represented by (7).

$$(I * K)(i, j) = \sum_m \sum_n I(i - m, j - n)\, K(m, n) \tag{7}$$

Lower CNN layers generally learns edges, corners, textures. The deeper layers captures semantic structures, pathological patterns, disease-specific representations. Residual learning and dense connectivity mechanisms further improved representation learning efficiency by enhancing gradient propagation and feature reuse [8], [9].

3.3. Medical image segmentation

Medical image segmentation involves partitioning healthcare images into diagnostically meaningful regions such as tumors, organs, lesions, vascular structures, pathological tissues. Segmentation plays a critical role in radiotherapy planning, surgical guidance, disease progression analysis, quantitative healthcare imaging. The segmentation problem is formulated as (8).

$$\hat{S} = f_\theta(I) \tag{8}$$

Encoder–decoder architectures such as U-Net became dominant biomedical segmentation frameworks due to their ability to combine contextual and spatial feature information [3], [6]. Skip-connection mechanisms improve localization precision by transferring fine-grained spatial features between encoder and decoder pathways. The Dice Similarity Coefficient (DSC) is one of the most widely used segmentation evaluation metrics. Another commonly used metric is Intersection over Union (IoU), which is given by (9).

$$IoU = \frac{|P \cap G|}{|P \cup G|}. \tag{9}$$

Advanced segmentation frameworks including Attention U-Net, DeepLab, SegNet, and nnU-Net significantly improves segmentation robustness, lesion localization, boundary delineation, multimodal healthcare segmentation.

3.4. Medical image classification

Medical image classification aims to assign clinically meaningful diagnostic labels to healthcare images or segmented pathological structures [14], [17]. Classification frameworks are widely used for cancer diagnosis, pneumonia detection, retinal disease screening, skin lesion classification, COVID-19 diagnosis. The classification objective is represented as (10).

$$y = f_\theta(x) \tag{10}$$

$$P(y = i) = \frac{e^{z_i}}{\sum_{j=1}^{C} e^{z_j}} \tag{11}$$

The CNN architectures demonstrated remarkable performance for healthcare image classification tasks due to their hierarchical representation learning capability [14], [15]. Nature-scale AI systems achieved diagnostic performance comparable to expert clinicians in skin cancer and lung disease diagnosis tasks [17], [18]. Transformer-based architectures further improved classification capability by enabling global contextual feature representation using self-attention mechanisms [12], [13].

3.5. Loss functions and optimization

Optimization plays a critical role in deep learning-based healthcare imaging systems because training involves minimizing highly nonlinear objective functions over high-dimensional parameter spaces [20], [25]. The general optimization objective is formulated as (12).

$$\theta^* = \arg\min_\theta \mathcal{L}(\theta) \tag{12}$$

Cross-entropy loss is widely employed for classification tasks (13):

$$\mathcal{L}_{CE} = -\sum_{i=1}^{C} y_i \log(\hat{y}_i) \tag{13}$$

Dice loss is frequently used for segmentation optimization. The optimization algorithms including Adam, stochastic gradient descent (SGD), and Root Mean Square Propagation (RMSProp) are commonly used for deep neural learning [32], [33]. However, deep healthcare imaging systems frequently suffer from unstable convergence, local minima stagnation, hyperparameter sensitivity. Metaheuristic optimization algorithms such as PSO, GA, GWO, and WOA therefore became increasingly integrated with healthcare imaging frameworks for hyperparameter optimization, feature selection, segmentation refinement, and architecture search.

3.6. Evaluation metrics

Performance evaluation is essential for validating segmentation and classification systems in healthcare environments. Classification systems are commonly evaluated by using accuracy, precision, recall, F1-score and area under the ROC curve (AUC). Classification accuracy is represented by (14).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \tag{14}$$

Segmentation systems are generally evaluated using Dice coefficient, IoU, Hausdorff distance and boundary overlap metrics. Reliable evaluation metrics are critical for ensuring clinical reliability and diagnostic robustness.

3.7. Discussion

The theoretical foundations of medical image segmentation and classification demonstrate the convergence of deep representation learning, optimization theory, and intelligent healthcare analytics. Automated hierarchical feature learning significantly improved healthcare image understanding compared with traditional handcrafted feature engineering frameworks. Encoder–decoder segmentation systems, CNN architectures, and transformer-based frameworks have become dominant healthcare imaging paradigms due to their strong nonlinear learning capability and scalability. Simultaneously, optimization-aware learning frameworks increasingly improve convergence stability, feature selection efficiency, and segmentation robustness. These theoretical foundations therefore provide the computational basis for modern hybrid deep learning and metaheuristic optimization systems used in intelligent healthcare imaging applications.

4. Hybrid deep learning and optimization frameworks

Hybrid deep learning and optimization frameworks combine nonlinear representation learning capability with global optimization mechanisms for improving healthcare imaging performance [1], [20]. Deep neural architectures provide automated feature extraction and hierarchical representation learning, whereas metaheuristic optimization algorithms improve convergence stability, feature selection efficiency, and hyperparameter tuning capability [21], [25]. The integration of these computational paradigms has significantly improved segmentation accuracy, classification robustness, and optimization efficiency in intelligent healthcare imaging systems. Hybrid healthcare imaging systems generally integrates convolutional neural networks, encoder–decoder segmentation architectures, transformer-based frameworks, swarm intelligence optimization, evolutionary optimization algorithms and multimodal healthcare intelligence. The general hybrid learning framework can be represented as (15).

$$y = f_{\theta,\Omega}(I) \tag{15}$$

Figure 3 illustrates the general workflow of hybrid deep learning and optimization systems for medical image segmentation and classification.

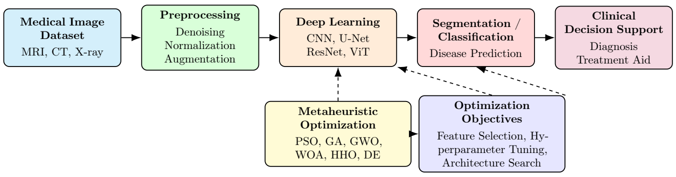


Figure 3: Hybrid framework showing the role of metaheuristic optimization in deep learning-based healthcare image segmentation and classification.

4.1. CNN-based hybrid optimization systems

CNN-based hybrid frameworks integrate optimization algorithms such as PSO, GA, DE and Bayesian optimization.

These optimization frameworks are extensively used for hyperparameter tuning, feature selection, learning rate optimization, and neural architecture refinement.

The optimization objective is represented as (16).

$$\theta^* = \arg\min_{\theta} \mathcal{L}(\theta) \tag{16}$$

In (16), θ denotes trainable parameters and $\mathcal{L}(\theta)$ represents the objective function.

PSO-based CNN frameworks demonstrated strong performance can be found in chest disease diagnosis, retinal image classification, lung cancer analysis, histopathological imaging. Similarly, GA-based optimization systems improved CNN feature selection efficiency and classification robustness through evolutionary search mechanisms [20], [24].

4.2. Optimization-driven segmentation frameworks

Optimization-driven segmentation systems integrate encoder-decoder architectures with metaheuristic optimization techniques for improving boundary delineation, lesion localization, segmentation precision, and convergence stability. These frameworks are commonly combines U-Net, Attention U-Net, DeepLab, nnU-Net by using optimization algorithms such GWO, WOA,FA, and Bayesian optimization.

The Dice-based segmentation loss function is represented as (17).

$$\mathcal{L}_{Dice} = 1 - \frac{2|P \cap G|}{|P| + |G|} \tag{17}$$

In (17), P denotes predicted segmentation masks and G denotes ground-truth annotations.

Optimization-aware segmentation systems demonstrated strong performance in brain tumor segmentation, organ delineation, retinal vessel extraction, pulmonary lesion localization.

4.3. Transformer-based hybrid learning systems

Transformer-based hybrid frameworks combines self-attention mechanisms, convolutional feature extraction, optimization-aware learning and multimodal representation learning. These systems improves contextual feature representation, long-range dependency modeling, multimodal healthcare analysis and segmentation robustness.

The self-attention mechanism is represented as (18).

$$Attention(Q,K,V) = Softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V. \tag{18}$$

Hybrid CNN–transformer systems demonstrated strong performance in multimodal disease diagnosis, anatomical segmentation, MRI analysis, volumetric healthcare imaging.

Optimization algorithms further improve transformer hyperparameter tuning and training stability.

4.4. Ensemble and multimodal hybrid systems

Ensemble healthcare imaging systems combine multiple deep learning architectures and optimization frameworks for improving diagnostic reliability [39] ,[40]. These systems integrate CNN ensembles, transformer ensembles, multimodal fusion frameworks and optimization-aware classifiers. Multimodal healthcare systems combine information from CT imaging, MRI imaging, PET imaging, histopathological images and electronic healthcare records. The multimodal fusion framework is represented as (19). Ensemble optimization systems demonstrated improveds diagnostic accuracy, robustness, generalization capabilityand clinical reliability.

$$F_{fusion} = \sum_{i=1}^{N} w_i F_i \tag{19}$$

4.5. Discussion

Hybrid deep learning and optimization frameworks significantly improved healthcare imaging systems through the integration of representation learning and optimization-aware computational intelligence. CNN-based hybrid systems demonstrated strong feature learning capability, whereas optimization-driven segmentation frameworks improved lesion localization and segmentation precision.

Transformer-based hybrid architectures further enhanced contextual representation learning and multimodal healthcare analysis capability. Simultaneously, ensemble and multimodal systems improved clinical reliability and diagnostic robustness.

Despite these advancements, hybrid healthcare imaging systems still face challenges related to:

- computational complexity;
- optimization instability;
- limited annotated datasets;

- interpretability;
- real-world clinical generalization.

Nevertheless, optimization-aware hybrid learning systems remain one of the most promising research directions for next-generation intelligent healthcare imaging frameworks.

5. Medical image segmentation and classification applications

Hybrid deep learning and metaheuristic optimization frameworks are increasingly applied across numerous healthcare imaging applications due to their strong capability for automated diagnosis, lesion localization, segmentation refinement, and intelligent clinical decision support [1], [2], [39], [42]. The integration of deep representation learning with optimization-aware computational intelligence has substantially improved segmentation precision, classification robustness, and convergence stability across multiple imaging modalities including MRI, computed tomography (CT), chest X-ray imaging, retinal imaging, histopathological imaging and ultrasound imaging.

5.1. Lung disease diagnosis

Chest imaging systems are extensively used for pneumonia detection, lung cancer diagnosis, tuberculosis screening, COVID-19 analysis, and pulmonary disease classification. CNN-based healthcare imaging frameworks demonstrated remarkable performance in chest disease diagnosis using X-ray and CT imaging datasets [18], [41], [61], [63]. Deep learning systems significantly improves feature representation capability, lesion localization, automated disease screening, and diagnostic accuracy. Hybrid optimization-driven systems further improved classification robustness through CNN hyperparameter tuning, feature optimization, segmentation refinement and ensemble learning. The classification framework is represented by (20), where x denotes extracted image features, y denotes disease predictions. Hybrid CNN and optimization frameworks demonstrated strong capability for identifying pulmonary abnormalities in highly heterogeneous healthcare imaging datasets.

$$y = f_\theta(x) \tag{20}$$

5.2. Brain tumor segmentation

Brain tumor segmentation using MRI is one of the most important biomedical segmentation applications due to its critical role in surgical planning, radiotherapy, disease monitoring and treatment assessment. Encoder-decoder architectures such as U-Net and Attention U-Net significantly improved brain tumor localization and volumetric segmentation capability [3], [19], [30]. Transformer-based segmentation systems further improved contextual feature representation for multimodal MRI analysis [56], [57]. The segmentation process is represented by (21) and optimization-aware segmentation systems using PSO, GWO, WOA and Bayesian optimization.

$$\hat{S} = f_\theta(I) \tag{21}$$

The Dice Similarity Coefficient remains one of the most widely used segmentation evaluation metrics is given by (22).

$$Dice = \frac{2|P \cap G|}{|P| + |G|} \tag{22}$$

Advanced biomedical segmentation frameworks such as nnU-Net, Swin UNETR, and TransUNet have further improved segmentation robustness and generalization capability across multimodal MRI datasets [55]–[57].

5.3. Retinal disease detection

Retinal healthcare imaging systems are extensively used for diabetic retinopathy detection, retinal vessel segmentation, glaucoma diagnosis and macular degeneration analysis. CNN-based retinal analysis systems demonstrated strong capability for automated feature extraction and disease classification [1], [15], [42]. Hybrid optimization frameworks improves vessel segmentation, microaneurysm detection, retinal lesion localization and classification accuracy. Attention-guided segmentation systems further enhanced localization of clinically significant retinal abnormalities.

5.4. Histopathological image analysis

Histopathological imaging plays a critical role in cancer diagnosis, tissue classification, cellular analysis and pathological grading [1], [2], [14]. Deep learning systems demonstrated remarkable performance in identifying the cancerous tissues, abnormal cellular structures, histological patterns and tissue morphology. Hybrid CNN and optimization frameworks significantly improve feature selection, classification robustness, segmentation precision and multimodal pathological analysis. Transformer-based systems further improves contextual feature representation in high-resolution histopathological images [13], [57].

5.5. Skin lesion classification

Skin lesion classification systems support early diagnosis of melanoma, basal cell carcinoma and dermatological abnormalities. Deep CNN architectures achieved dermatologist-level classification performance in skin cancer diagnosis tasks [17], [67]. Hybrid optimization systems improve feature optimization, classification stability, segmentation accuracy and lesion boundary detection. Attention-guided systems additionally improved localization of pathological skin regions and clinically significant lesion structures.

5.6. COVID-19 detection and pandemic healthcare imaging

COVID-19 significantly accelerated research in AI-based healthcare imaging systems due to the urgent need for rapid disease screening and automated diagnostic support [64]–[66]. Deep learning frameworks demonstrated promising performance for COVID-19 screening, pulmonary infection localization, severity classification and chest X-ray analysis. Optimization-aware CNN systems improves classification accuracy, convergence stability, diagnostic robustness and feature optimization. Hybrid healthcare imaging frameworks therefore became highly important for pandemic-oriented intelligent diagnostic systems.

5.7. Discussion

Hybrid deep learning and optimization frameworks demonstrated substantial improvements across multiple healthcare imaging applications including disease diagnosis, lesion localization, tumor segmentation, retinal analysis and histopathological classification. CNN architectures remain dominant for healthcare classification tasks, whereas encoder–decoder systems continue to provide strong biomedical segmentation capability [3], [19], [30], [55]. Optimization-aware systems further improve convergence stability, feature optimization efficiency, and diagnostic robustness [20], [21], [23], [25]. Transformer-based and multimodal healthcare imaging systems are expected to further enhance contextual representation learning, clinical reliability, multimodal data fusion and intelligent healthcare analytics [13], [56], [57]. These applications demonstrate the growing importance of optimization-aware intelligent healthcare imaging systems for next-generation computer-aided diagnostic frameworks.

6. Challenges and future research directions

Despite remarkable advancements in hybrid deep learning and metaheuristic optimization frameworks, several critical challenges still limit their large-scale deployment in real-world healthcare environments [26], [39], [40]. Although intelligent healthcare imaging systems demonstrated substantial improvements in segmentation precision, classification accuracy, and optimization efficiency, practical clinical implementation remains constrained by computational complexity, limited annotated datasets, optimization instability, lack of interpretability, poor cross-domain generalization and privacy and security concerns. The growing complexity of multimodal healthcare imaging systems therefore continues to motivate research toward scalable, explainable, trustworthy, and optimization-aware intelligent medical imaging frameworks.

6.1. Computational complexity and resource constraints

Deep neural architectures and transformer-based healthcare imaging systems often involve millions of trainable parameters and extremely high computational complexity [12], [13], [28], [56], [57]. Training large-scale segmentation and classification frameworks requires high-performance GPUs, large memory capacity, extensive training time and energy-intensive computation. Transformer architectures further increase computational burden due to quadratic self-attention complexity and which is represented by (23). In (23), n denotes the sequence length.

$$Complexity = O(n^2) \tag{23}$$

These limitations restrict deployment of intelligent healthcare imaging systems in low-resource clinical environments, portable healthcare devices, edge computing systems and real-time diagnostic applications. Future research should therefore focus on lightweight healthcare AI models, efficient transformer architectures, model compression techniques and edge-aware healthcare intelligence. Recent lightweight architectures such as MobileNet and EfficientNet provide promising directions for reducing computational overhead while maintaining diagnostic performance [29], [31].

6.2. Limited annotated medical imaging data

Healthcare imaging datasets frequently suffer from limited annotations, class imbalance, noisy labels, multimodal variability and insufficient rare disease samples. Accurate annotation of healthcare images requires domain expertise from radiologists and medical specialists, making large-scale dataset generation expensive and time-consuming [39], [42]. Class imbalance additionally affects segmentation and classification robustness because pathological abnormalities often occupy small image regions. The imbalance ratio is represented as (24).

$$IR = \frac{N_{majority}}{N_{minority}} \tag{24}$$

Limited data availability remains a major bottleneck for developing robust healthcare imaging systems. Transfer learning, self-supervised learning, generative adversarial networks and data augmentation strategies have emerged as effective solutions for mitigating annotation scarcity and improving model generalization [45], [53], [54]. Future research should further explore foundation models, synthetic data generation and few-shot learning paradigms for healthcare imaging applications.

6.3. Optimization instability and generalization challenges

Deep healthcare imaging systems frequently suffer from the following, (i) local minima stagnation, (ii) unstable convergence, (iii) hyperparameter sensitivity, (iv) overfitting and (v) poor cross-domain generalization. Optimization-aware healthcare systems often involve highly nonlinear objective functions over large parameter spaces as given by (25).

$$\theta^* = \arg\min_{\theta} \mathcal{L}(\theta) \tag{25}$$

Although metaheuristic optimization algorithms improved convergence capability, several challenges remain related to exploration–exploitation balance, computational overhead, scalability and reproducibility [20]–[24]. Hyperparameter optimization frameworks such as Bayesian optimization and Optuna-based automated search strategies have demonstrated promising capability for improving training stability and model performance [25], [34]. Future research should focus on adaptive optimization frameworks capable of dynamically balancing exploration and exploitation while maintaining computational efficiency.

6.4. Interpretability and explainable artificial intelligence

Clinical deployment of intelligent healthcare imaging systems requires transparency and interpretability because healthcare decisions directly affect patient outcomes [58]. However, many deep learning systems behave as black-box models with limited interpretability. Explainable artificial intelligence (XAI) frameworks improve transparency by identifying diagnostically important image regions and decision pathways. Common explainability methods includes Grad-CAM, saliency maps, SHAP analysis, and attention visualization. Grad-CAM generates class-discriminative localization maps using gradient information [59]. Similarly, SHAP provides feature attribution analysis for improving model interpretability [60]. Future healthcare AI systems should therefore focus on transparent clinical decision-making, trustworthy healthcare intelligence, explainable multimodal systems and interpretable optimization frameworks.

6.5. Privacy, security, and federated learning

Healthcare imaging datasets contain highly sensitive patient information, thereby creating major concerns related to privacy preservation, cybersecurity, healthcare data sharing and regulatory compliance. Federated learning enables collaborative distributed model training without centralized healthcare data sharing [27]. Instead of transferring patient data, only model parameters are exchanged between healthcare institutions. The federated learning framework is represented by (26).

$$w^{t+1} = \sum_{k=1}^{K} \frac{n_k}{n} w_k^t \tag{26}$$

Future healthcare AI systems are expected to increasingly employ federated learning, privacy-preserving optimization, secure multimodal healthcare intelligence, blockchain-enabled healthcare analytics, and trustworthy uncertainty-aware decision systems [26], [27].

6.6. Multimodal and personalized healthcare intelligence

Future intelligent healthcare imaging systems are expected to move toward multimodal and personalized healthcare analytics by integrating medical imaging, electronic healthcare records, genomic information, clinical laboratory data, physiological monitoring systems. Multimodal healthcare intelligence significantly improves diagnostic accuracy, disease prediction, personalized treatment planning, and clinical decision support [39], [40]. The multimodal fusion framework is represented as (27).

$$F_{fusion} = \sum_{i=1}^{N} w_i F_i \tag{27}$$

Recent transformer-based multimodal architectures have demonstrated strong capability for learning cross-modal relationships and contextual healthcare representations. Future research should therefore focus on multimodal transformer systems, personalized healthcare AI, adaptive clinical intelligence and optimization-aware multimodal learning.

6.7. Discussion

Although hybrid deep learning and optimization frameworks achieved remarkable progress in healthcare imaging systems, several critical challenges continue to limit real-world clinical deployment. Computational complexity, limited annotated datasets, optimization instability, interpretability limitations, and privacy concerns remain major research problems. Future intelligent healthcare imaging systems are expected to increasingly focus on lightweight architectures, explainable healthcare AI, federated learning, multimodal healthcare intelligence and trustworthy optimization-aware systems [26], [27], [40], [58]. The convergence of deep learning, optimization theory, multimodal analytics, and explainable clinical intelligence therefore represents one of the most promising research directions for next-generation healthcare imaging systems.

7. Conclusion

Medical image segmentation and classification play a vital role in modern healthcare by enabling accurate disease diagnosis, treatment planning, and clinical decision support. The rapid advancement of deep learning has significantly improved the capability of automated medical image analysis systems across various imaging modalities, including MRI, CT, X-ray, ultrasound, retinal, and histopathological imaging. Deep learning architectures such as convolutional neural networks, U-Net variants, residual networks, and vision transformers have demonstrated remarkable success in extracting complex image features and improving diagnostic performance. This review comprehensively examined the integration of deep learning models with metaheuristic optimization techniques for medical image segmentation and classification. Optimization algorithms have proven effective for hyperparameter tuning, feature selection, architecture optimization, and segmentation refinement. The combination of deep representation learning and global optimization enables hybrid frameworks to achieve enhanced accuracy, robustness, and convergence performance. Despite significant progress, challenges related to data scarcity, class imbalance, computational complexity, interpretability, privacy preservation, and clinical deployment remain important research concerns. Future developments are expected to focus on explainable artificial intelligence, federated learning, multimodal healthcare analytics, and trustworthy optimization-driven frameworks. Overall, hybrid deep learning and metaheuristic optimization approaches offer substantial potential for developing reliable, scalable, and clinically deployable next-generation medical imaging systems.

Declarations and ethical statements

__Conflict of interest:__ The authors declare that there is no conflict of interest.

__Funding statement:__ The authors declare that no specific funding was received for this research.

__Artificial Intelligence usage statement:__ During the preparation of this manuscript, the authors utilized Claude and Grammarly solely for language refinement and grammatical corrections. The authors carefully reviewed and revised the generated content and take full responsibility for the accuracy, integrity, and originality of the final manuscript.

__Availability of data and materials:__ All data and information used in this work were obtained from publicly available published literature and cited sources.

CRediT authorship contribution statement

Potharla Ramadevi: Conceptualization, Data collection, Writing – review & editing. **Bala Krishnama Manohar**:

Data curation & Formal analysis. **N. Bhargavi**: Data curation & Visualization. **Valluru Prathyusha**: Conceptualization, Visualization & Editing. **P. Shailaja**: Data curation & Visualization.

Publisher's note

Krrish Scientific Publications Pvt. Ltd. and the ***Journal of Computing and Data Technology*** remain neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

[1] Litjens G, Kooi T, Bejnordi BE, Setio AA, Ciompi F, Ghafoorian M, Van Der Laak JA, Van Ginneken B, Sánchez CI. A survey on deep learning in medical image analysis. *Medical Image Analysis.* 2017 Dec 1;42:60-88. Available from: https://doi.org/10.1016/j.media.2017.07.005

[2] Shen D, Wu G, Suk HI. Deep Learning in Medical Image Analysis. *Annual Review of Biomedical Engineering.* 2017;19:221–248. Available from: https://doi.org/10.1146/annurev-bioeng-071516-044442

[3] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *In Lecture Notes in Computer Science.* 2015;9351:234–241. Available from: https://doi.org/10.1007/978-3-319-24574-4_28

[4] Long J, Shelhamer E, Darrell T. Fully Convolutional Networks for Semantic Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 2017;39(4):640–651. Available from: https://doi.org/10.1109/TPAMI.2016.2572683

[5] Milletari F, Navab N, Ahmadi SA. V-Net: Fully convolutional Neural Networks for Volumetric Medical Image Segmentation. *2016 Fourth International Conference on 3D Vision.* 2016:565–571. Available from: https://doi.org/10.1109/3DV.2016.79

[6] Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J. UNet++: A nested U-Net architecture for medical image segmentation. *Lecture Notes in Computer Science.* 2018;11045:3–11. Available from: https://doi.org/10.1007/978-3-030-00889-5_1

[7] Schlemper J, Oktay O, Schaap M, Heinrich M, Kainz B, Glocker B, Rueckert D. Attention gated networks: Learning to leverage salient regions in medical images. *Medical Image Analysis.* 2019;53:197–207. Available from: https://doi.org/10.1016/j.media.2019.01.012

[8] He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* 2016:770–778. Available from: https://doi.org/10.1109/CVPR.2016.90

[9] Huang G, Liu Z, van der Maaten L, Weinberger KQ. Densely connected convolutional Networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* 2017:4700–4708. Available from: https://doi.org/10.1109/CVPR.2017.243

[10] Badrinarayanan V, Kendall A, Cipolla R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 2017;39(12):2481–2495. Available from: https://doi.org/10.1109/TPAMI.2016.2644615

[11] Chen LC, Papandreou G, Kokkinos I, Murphy K, Yuille AL. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 2018;40(4):834–848. Available from: https://doi.org/10.1109/TPAMI.2017.2699184

[12] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L, Polosukhin I. Attention is All You Need. *Advances in Neural Information Processing Systems.* 2017;30:5998–6008. Available from: https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html

[13] Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S, Uszkoreit J, Houlsby N. An image is worth 16x16 words: Transformers for image recognition at scale. *International Conference on Learning Representations.* 2021. Available from: https://arxiv.org/abs/2010.11929

[14] Tajbakhsh N, Shin JY, Gurudu SR, Hurst RT, Kendall CB, Gotway MB, Liang J. Convolutional neural networks for medical image analysis: Full training or fine tuning? *IEEE Transactions on Medical Imaging.* 2016;35(5):1299–1312. Available from: https://doi.org/10.1109/TMI.2016.2535302

[15] Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: An overview and application in radiology. *Insights into Imaging.* 2018;9:611–629. Available from: https://doi.org/10.1007/s13244-018-0639-9

[16] Lundervold AS, Lundervold A. An overview of deep learning in medical imaging focusing on MRI. *Zeitschrift für Medizinische Physik.* 2019;29(2):102–127. Available from: https://doi.org/10.1016/j.zemedi.2018.11.002

[17] Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. Dermatologist-level classification of skin cancer with deep neural networks. *Nature.* 2017;542:115–118. Available from: https://doi.org/10.1038/nature21056

[18] Ardila D, Kiraly AP, Bharadwaj S, Choi B, Reicher JJ, Peng L, Tse D, Etemadi M, Ye W, Corrado G, Naidich DP, Shetty S. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nature Medicine.* 2019;25:954–961. Available from: https://doi.org/10.1038/s41591-019-0447-x

[19] Havaei M, Davy A, Warde-Farley D, Biard A, Courville A, Bengio Y, Pal C, Jodoin PM, Larochelle H. Brain tumor segmentation with Deep Neural Networks. *Medical Image Analysis.* 2017;35:18–31. Available from: https://doi.org/10.1016/j.media.2016.05.004

[20] Kennedy J, Eberhart R. Particle swarm optimization. *Proceedings of ICNN'95 – International Conference on Neural Networks.* 1995;4:1942–1948. Available from: https://doi.org/10.1109/ICNN.1995.488968

[21] Mirjalili S. Grey Wolf Optimizer. *Advances in Engineering Software.* 2014;69:46–61. Available from: https://doi.org/10.1016/j.advengsoft.2013.12.007

[22] Yang XS. Firefly Algorithms for Multimodal Optimization. *Lecture Notes in Computer Science.* 2009;5792:169–178. Available from: https://doi.org/10.1007/978-3-642-04944-6_14

[23] Mirjalili S, Lewis A. The Whale Optimization Algorithm. *Advances in Engineering Software.* 2016;95:51–67. Available from: https://doi.org/10.1016/j.advengsoft.2016.01.008

[24] Storn R, Price K. Differential Evolution – A Simple and Efficient Heuristic for global Optimization over Continuous Spaces. *Journal of Global Optimization.* 1997;11:341–359. Available from: https://doi.org/10.1023/A:1008202821328

[25] Bergstra J, Bardenet R, Bengio Y, Kégl B. Algorithms for Hyper-Parameter Optimization. *Advances in Neural Information Processing Systems.* 2011;24. Available from: https://proceedings.neurips.cc/paper_files/paper/2011/hash/86e8f7ab32cfd12577bc2619bc635690-Abstract.html

[26] Abdar M, Pourpanah F, Hussain S, Rezazadegan D, Liu L, Ghavamzadeh M, Fieguth P, Cao X, Khosravi A, Acharya UR, Makarenkov V. A review of uncertainty quantification in deep learning: Techniques, applications and challenges. *Information Fusion.* 2021 Dec 1;76:243-97. Available from: https://doi.org/10.1016/j.inffus.2021.05.008

[27] McMahan B, Moore E, Ramage D, Hampson S, y Arcas BA. Communication-Efficient Learning of Deep Networks from Decentralized Data. *In Artificial Intelligence and Statistics* 2017 Apr 10 (pp. 1273-1282). Pmlr. Available from: https://proceedings.mlr.press/v54/mcmahan17a.html

[28] Touvron H, Cord M, Douze M, Massa F, Sablayrolles A, Jégou H. Training data-efficient image transformers & distillation through attention. *In International conference on machine learning 2021* Jul 1 (pp. 10347-10357). PMLR. Available from: https://proc

eedings.mlr.press/v139/touvron21a.html

[29] Tan M, Le Q. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *In International Conference on Machine Learning (ICML) 2029* May 24 (pp. 6105-6114). PMLR. Available from: https://proceedings.mlr.press/v97/tan19a.html

[30] Oktay O, Schlemper J, Folgoc LL, Lee M, Heinrich M, Misawa K, Mori K, McDonagh S, Hammerla NY, Kainz B, Glocker B. Attention U-Net: Learning Where to Look for the Pancreas. *arXiv preprint arXiv:1804.03999*. 2018 Apr 11. Available from: https://openreview.net/forum?id=Skft7cijM

[31] Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M, Adam H. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *arXiv preprint arXiv:1704.04861*. 2017 Apr 17. Available from: https://doi.org/10.48550/arXiv.1704.04861

[32] Kingma DP, Ba J. Adam: A Method for Stochastic Optimization. *arXiv preprint* arXiv:1412.6980. 2014 Dec 22. Available from: https://doi.org/10.48550/arXiv.1412.6980

[33] Ioffe S, Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *InInternational conference on machine learning* 2015 Jun 1 (pp. 448-456). pmlr. Available from: https://proceedings.mlr.press/v37/ioffe15.html

[34] Akiba T, Sano S, Yanase T, Ohta T, Koyama M. Optuna: A Next-generation Hyperparameter Optimization Framework. *In-Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining* 2019 Jul 25 (pp. 2623-2631). Available from: https://doi.org/10.1145/3292500.3330701

[35] Goodfellow I, Bengio Y, Courville A, Bengio Y. Deep learning. *Cambridge: MIT press*; 2016 Nov 18. Available from: https://www.deeplearningbook.org

[36] LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*. 1998 Nov 30;86(11):2278-324. Available from: https://doi.org/10.1109/5.726791

[37] LeCun Y, Bengio Y, Hinton G. Deep learning. *nature*. 2015 May 28;521(7553):436-44. Available from: https://doi.org/10.1038/nature14539

[38] Bengio Y, Courville A, Vincent P. Representation Learning: A Review and New Perspectives. *IEEE transactions on pattern analysis and machine intelligence*. 2013 Mar 7;35(8):1798-828. Available from: https://doi.org/10.1109/TPAMI.2013.50

[39] Miotto R, Wang F, Wang S, Jiang X, Dudley JT. Deep learning for healthcare: review, opportunities and challenges. *Briefings in Bioinformatics*. 2018 Nov;19(6):1236-46. Available from: https://doi.org/10.1093/bib/bbx044

[40] Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nature medicine*. 2019 Jan;25(1):44-56. Available from: https://doi.org/10.1038/s41591-018-0300-7

[41] Rajpurkar P, Irvin J, Zhu K, Yang B, Mehta H, Duan T, Ding D, Bagul A, Langlotz C, Shpanskaya K, Lungren MP. CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning. *arXiv preprint arXiv*:1711.05225. 2017 Nov 14. Available from: https://doi.org/10.48550/arXiv.1711.05225

[42] Suzuki K. Overview of deep learning in medical imaging. *Radiological Physics and Technology*. 2017 Sep;10(3):257-73. Available from: https://doi.org/10.1007/s12194-017-0406-5

[43] Greenspan H, Van Ginneken B, Summers RM. Guest Editorial Deep Learning in Medical Imaging: Overview and Future Promise of an Exciting New Technique. *IEEE Transactions on Medical Imaging*. 2016 Apr 29;35(5):1153-9. Available from: https://doi.org/10.1109/TMI.2016.2553401

[44] Litjens G, Ciompi F, Wolterink JM, de Vos BD, Leiner T, Teuwen J, Išgum I. State-of-the-Art Deep Learning in Cardiovascular Image Analysis. *JACC: Cardiovascular Imaging*. 2019 Aug;12(8 Part 1):1549-65. Available from: https://doi.org/10.1016/j.jcmg.2019.06.009

[45] Raghu M, Zhang C, Kleinberg J, Bengio S. Transfusion: Understanding Transfer Learning for Medical Imaging. *Advances in Neural Information Processing Systems*. 2019;32. Available from: https://proceedings.neurips.cc/paper/2019/hash/eb1e78328c46506b46a4ac4a1e378b91-Abstract.html

[46] Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L. ImageNet: A large-scale hierarchical image database. *In 2009 IEEE Conference on computer vision and Pattern recognition 2009* Jun 20 (pp. 248-255). Ieee. Available from: https://doi.org/10.1109/CVPR.2009.5206848

[47] Breiman L. Random Forests. *Machine learning*. 2001 Oct;45(1):5-32. Available from: https://doi.org/10.1023/A:1010933404324

[48] Cortes C, Vapnik V. Support-vector networks. *Machine learning*. 1995 Sep;20(3):273-97. Available from: https://doi.org/10.1007/BF00994018

[49] Jain AK. Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*. 2010 Jun 1;31(8):651-66. Available from: https://doi.org/10.1016/j.patrec.2009.09.011

[50] Zadeh LA. Fuzzy sets. *Information and Control*. 1965 Jun 1;8(3 Available from: https://doi.org/10.1016/S0019-9958(65)90241-X

[51] Kohonen T. The self-organizing map. *Proceedings of the IEEE*. 2002 Aug 6;78(9):1464-80. Available from: https://doi.org/10.1109/5.58325

[52] Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning. Available from: https://doi.org/10.1007/978-0-387-84858-7

[53] Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Generative adversarial nets. *Advances in Neural Information Processing Systems*. 2014;27. Available from: https://doi.org/10.1145/3422622

[54] Shorten C, Khoshgoftaar TM. A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*. 2019 Dec;6(1):1-48. Available from: https://doi.org/10.1186/s40537-019-0197-0

[55] Isensee F, Jaeger PF, Kohl SA, Petersen J, Maier-Hein KH. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods*. 2021 Feb;18(2):203-11. Available from: https://www.nature.com/articles/s41592-020-01008-z

[56] Hatamizadeh A, Nath V, Tang Y, Yang D, Roth HR, Xu D. Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images. *In International MICCAI Brainlesion Workshop* 2021 Sep 27 (pp. 272-284). Cham: Springer International Publishing. Available from: https://link.springer.com/chapter/10.1007/978-3-031-08999-2_22

[57] Chen J, Lu Y, Yu Q, Luo X, Adeli E, Wang Y, Lu L, Yuille AL, Zhou Y. TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation. *arXiv preprint arXiv:2102.04306*. 2021 Feb 8. Available from: https://doi.org/10.48550/arXiv.2102.04306

[58] Doshi-Velez F, Kim B. Towards A Rigorous Science of Interpretable Machine Learning. *arXiv preprint arXiv:1702.08608*. 2017 Feb 28. Available from: https://doi.org/10.48550/arXiv.1702.08608

[59] Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *In Proceedings of the IEEE international conference on computer vision* 2017 (pp. 618-626). Available from: https://doi.org/10.1109/ICCV.2017.74

[60] Lundberg SM, Lee SI. A Unified Approach to Interpreting Model Predictions. *Advances in neural information processing systems*. 2017;30. Available from: https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html

[61] Akuthota S. Enhancing Chronic Obstructive Pulmonary Disease (COPD) Diagnosis through Machine Learning Models Trained on Respiratory Sounds. *In 2025 2nd International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE) 2025* May 7 (pp. 1-6). IEEE. Available from: https://doi.org/10.1109/RMKMATE64874.2025.11042677

[62] Akuthota S. Enhanced Breast Cancer Classification Using Attention-Augmented CNN and Multi-View Learning on the Inbreast Dataset. *In 2025 International Conference on Computing Technologies & Data Communication (ICCTDC) 2025* Jul 4 (pp. 1-5). IEEE. Available from: https://doi.org/10.1109/ICCTDC64446.2025.11159034

[63] Ramadevi P, Das R. An Extensive Analysis of Machine Learning Techniques With Hyper-Parameter Tuning by Bayesian Optimized SVM Kernel for the Detection of Human Lung Disease. *IEEE Access*. 2024 Jul 3;12:97752-70. Available from: https://doi.org/10.1109/ACCESS.2024.3422449

[64] Manohar B, Das R, Ramadevi P, Balusamy B. A hybridized long–short-term memory networks-based deep learning model using reptile search optimization for COVID-19 prediction. *In Emerging Trends and Applications of Deep Learning for Biomedical Data Analysis* 2025 Jan 1 (pp. 49-72). Academic Press. Available from: https://doi.org/10.1016/B978-0-443-26765-9.00003-2

[65] Manohar B, Das R, Lakshmi M. A hybridized LSTM-ANN-RSA based deep learning models for prediction of COVID-19 cases in Eastern European countries. *Expert Systems with Applications*. 2024 Dec 5;256:124977. Available from: https://doi.org/10.1016/j.eswa.2024.124977

[66] Manohar B, Das R.Comparison of Hybrid Artificial Neural Networks With GA, PSO, and RSA in Predicting COVID-19 Cases: A Case Study of India. InMulti-Disciplinary applications of fog computing: responsiveness in real-time 2023 (pp. 207-244). IGI Global Scientific Publishing. Available from: https://doi.org/10.4018/978-1-6684-4466-5.ch011

[67] Fatmir Basholli, Mohammed R. Hayal, Ebrahim E. Elsayed, Davron Aslonqulovich Juraev. Deep Learning for Skin Disease Classification: A Comparative Study of CNN and CNN-LSTM Architectures. *Journal of Computing and Data Technology* . 2025 Jun. 30;1(1):40-9. Available from: https://doi.org/10.71426/jcdt.v1.i1.pp40-49

[68] Aazad SK, Saini T, Ajad A, Chaudhary K, Elsayed EE. Deciphering Blood Cells - Method for Blood Cell Analysis using Microscopic Images. *Journal of Modern Technology*. 2024 Sep. 8;1(1):9-18. Available from: https://doi.org/10.71426/jmt.v1.i1.pp9-18

[69] Kalnoor G, Dasari KS, Suma S, Waddenkery N, B. Pragathi. Enhanced Brain Tumor Detection from MRI Scans Using Frequency Domain Features and Hybrid Machine Learning Models. *Journal of Modern Technology*. 2025 Jan. 8;1(2):141-9. Available from: https://doi.org/10.71426/jmt.v1.i2.pp141-149

[70] Balaji VR, Jahan MI, Sridarshini T, Geerthana S, Thirumurugan A, Hegde G, Sitharthan R, Dhanabalan SS. Machine learning enabled 2D photonic crystal biosensor for early cancer detection. *Measurement*. 2024 Jan 1;224:113858. Available from: https://doi.org/10.1016/j.measurement.2023.113858

[71] Oyinna B, Udo PD, Nurhidayat I, Muslimyar AR. Integrating Data Processing and Advanced Analytics for Scalable Knowledge Discovery in Complex Data Environments. *Journal of Computing and Data Technology*. 2025 Nov. 21;1(2):115-20. Available from: https://doi.org/10.71426/jcdt.v1.i2.pp115-120