SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V13I2P110

Volume 13 Issue 2, 133-146, February 2026

Original Article

Hyper_ALO_HybML: A Radiomics-Driven Ensemble Framework with Improved Ant Lion Optimization for Osteosarcoma Classification

Guntoju Kalpana Devi[1], Mahesh Babu Arrama[2], Jayaprakash Chennoju[3]

[1,2]*Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Hyderabad, Telangana, India*
[3]*Department of ECE, Sir C.R. Reddy College of Engineering, Eluru, Andhra Pradesh, India.*

[1]*Corresponding Author : gkalapana15@gmail.com,*

Received: 19 November 2025 Revised: 20 December 2025 Accepted: 21 January 2026 Published: 20 February 2026

Abstract *- This study presents Hyper_ALO_HybML, a hybrid machine learning framework enhanced through Improved Ant Lion Optimization (ALO) for accurate Osteosarcoma Classification from H&E-stained histopathological images. The radiomics-based texture, shape, and intensity features were processed with three ensemble classifiers, including Random Forest, XGBoost (eXtreme Gradient Boosting), and LightGBM, all tuned using the modified ALO to enhance feature selection ability, learning stability, and anti-overfitting. The dataset is composed of 1,144 images classified into Non-Tumor, Viable Tumor, and Necrotic Tumor. The experimental results demonstrate that the system with ALO-optimized XGBoost obtains the best performance, achieving an accuracy of 93.29%, a precision of 94.08%, a recall of 91.94%, and an MCC of 0.8947, respectively. LightGBM also holds similar generalization, with 92.71% accuracy, and Random Forest stops at 85.42%. Confusion matrix review reveals that gradient boosted models (optimized) deliver the most stable predictions among the classes. Results indicate that the integration of radiomics with ALO-guided hyperparameter optimization helps enhance Osteosarcoma detection and can aid in more reasonable histopathological decision-making. This hybrid approach offers a feasible route toward improved early diagnosis, minimizing errors in manual interpretation as well as enhancing computer-aided pathology.*

Keywords *- Osteosarcoma, Histopathology, Radiomics, Ensemble learning, XGBoost, LightGBM, Ant Lion Optimization, Hyperparameter tuning.*

1. Introduction

People globally have about 100 metabolic disorders [1]. Cancer is deadly and a significant danger to humanity. Cancer is characterized by aberrant cell proliferation and rapid spread. Some cancer-causing variables are avoidable, while others are not. Preventable causes include smoking, drinking, and UV exposure. However, genetic variables are unavoidable [2]. These genetic variables are inherited or caused by lifestyle behaviors like smoking. Additionally, aging is the most unpreventable cause. Bone cancer is uncommon and may start with swelling or soreness. Dysplastic bone cells may travel to other regions of the body and cause bone cancer. Bone cancer cases are anticipated to rise by 3900 in 2022, with 2100 fatalities [3]. Bone cancer is more common in those under 20. Therefore, radiography, X-ray, CT, MRI, ultrasound, and phototherapy or techniques based on PET are the most popular medical imaging technologies [4, 5]. DTBV is a deep-based transfer bone cancer diagnosis system with VGG16 feature extraction and employs X-ray as an image modality. Medical X-ray images are valuable for research and diagnosing disease, because they are low-cost and ubiquitously available. [6]. Surgery, radiation therapy, chemotherapy, and targeted medicines have all been recommended in recent years to manage the disease. Bone cancer is complex, and its detection is challenging [7]. Due to expensive equipment and mishandling details, manual evaluation of malignant photos is inefficient and inaccurate. Cancer research increasingly uses Machine Learning (ML) to analyze data and automatically extract key information [8]. Traditional approaches are accurate, but they lack the precision to identify fine-grained characteristics, quicker real-time detection on a big medical dataset, and flexibility and customization to adjust the algorithm, particularly with ensemble machine learning. Integrating ensemble processes into machine learning algorithms is not well tested. Given this, this work's contributions are as follows:

- The proposed method has been applied by online integrating anywhere modified antlion optimization algorithm with a hyper-parameter tuning function for classifying the Bone Cancer.

- Modified ant lion optimization-assisted RF, XGBoost, and Light GBM improve feature selection, enhance generalization, and mitigate overfitting, leading to more reliable predictions in histopathological image classification.

Osteosarcoma is a highly aggressive bone malignancy and requires accurate histopathological examination for diagnosis and treatment planning, and most existing approaches depend on deep complex architectures with a considerably high computational demand, or do not capture slight radiomic patterns that are important for multiclass classification. Moreover, the existing works focus on a single type of classifier or use some optimization skills, but fail to efficiently integrate the feature selection and ensemble learning.

2. Related Works

This section evaluates the current state of deep learning applications in bone cancer detection, focusing on important methodologies, their applications, and performance across various imaging modalities [9].

Whig et al. (2025) [10] are among the early works on bone cancer detection with classical machine learning methods. They show that manually detected characteristics of Musculoskeletal images can be appropriately classified by standard algorithms, incl. support vector machines or decision trees. Even if this study shows the possibility of automated bone cancer classification, it raises key limitations, which are feature engineering sensitivity and limited generalization on heterogeneous datasets. This work paves the way to the transition to stronger Deep Learning methods. Borji et al. (2025) [11] take this direction further by introducing a novel Hybrid Deep Learning Model designed for Osteosarcoma Histopathology Images. Their approach utilizes convolutional blocks for feature extraction and fine-tuning of classification layers to improve representation of cellular morphology as well as tissue-level patterns. The authors present better diagnostic performance over single CNN models, and note the significance of hybrid architectures in overcoming the complexity of histopathological data. Nevertheless, the method remains mainly concerned with image-level performance and is little devoted to clinical interpretability.

Muthupriya and Aarthy (2025) [12] enhance the search for osteosarcoma detection by introducing discriminative patch-based analysis and GAN-screened imaging. Their approach is centred on generative adversarial networks employed to enhance the quality of data and suppress noise, where patch-level discrimination assists in finer localisation of tumour-affected areas. The results from this study demonstrate the potential of data augmentation and region-specific learning, particularly in cases where only a few annotated medical images are available. Beyond diagnosis, Coudray et al. (2025) [13] design quantitative and morphology-based Deep Convolutional Neural Networks for predicting the survival of osteosarcoma treatment in neoadjuvant and metastasis environments. Their results are an important first step in prognostic modeling, suggesting that deep learning can capture patterns of survival directly from histological images. Unlike previous studies focused on classification, this study emphasizes the implications for AI in outcome prediction but relies on extensive, well-annotated data sets to maintain accuracy.

Cè et al. (2025) [14] present a thorough review of multimodal osteosarcoma imaging throughout the stages, ranging from initial diagnosis up to radiomics-based analysis. Their work demonstrates how the fusion of radiological imaging, radiomics features, and AI-based analysis can provide a comprehensive treatment paradigm. The combination of multimodal data is considered a promising avenue, but issues concerning data fusion, standardization, and reproducibility have yet to be solved. Papageorgiou et al. (2025) [15] narratively review the applications of artificial intelligence in primary malignant bone tumor imaging. Their contributions distill the latest advancements in detection, classification, segmentation, and prognosis, and critically address obstacles to clinical translation. Challenges highlighted are dataset bias, poor external validation, and low explainability. These key themes emphasize the requirement of interpretable AI mechanisms that can be effectively applied across populations.

Alsubai et al. (2024) [16] introduce the Group Teaching Optimization-based Deep Learning Model for histopathology images of Osteosarcoma. They focus on learning from an optimization-driven) point of view for convergence and classification accuracy. The results show the effectiveness of metaheuristic optimization in deep learning pipelines to improve accuracy, more so with complex histopathology problems. Nonetheless, the investigation continues to be limited to classification accuracy only without more comprehensive clinical or interpretability examination. Liu et al. (2024) [17] deal with osteosarcoma detection in the practical and translational sense by presenting an AI-guided diagnostic product design based on CA-MobileNet V3 applied to microscopy images. Their work emphasizes lightweight architectures that can be applied directly in a clinical setting, taking constraints like computational efficiency and scalability into account. Although the accuracy obtained by the proposed model is promising, based on a single compact architecture, it might fail to capture textural variations at multi-scale levels, which are common in osteosarcoma tissues.

McGenity et al. (2024) [18], in the setting of a systematic review and meta-analysis on AI in digital pathology, bring another higher-level perspective. Their results suggest most AI models have high diagnostic accuracy in cancers as a whole, for both bone malignancy and other cancer types.

Crucially, the study highlights fundamental limitations, including data bias, absence of external validation, and heterogeneity in reporting standards. Our study highlights the importance of rigorous evaluation and transparent model design in osteosarcoma studies. Zhao et al. (2024) [19] Machine learning based CT detection of breast cancer metastasis bone tumors enriches the domain of bone oncology. While not unique to Osteosarcoma, their study serves as an example of the potential utility of CT-based machine learning for bone tumor detection and emphasizes the increasing role of radiological imaging in addition to histopathology.

Vezakis et al. (2023) [20] compared several deep learning methods of detecting and diagnosing Osteosarcoma. Their systematic study exposes the advantages and limitations of several CNN architectures, thus providing practical guidelines for model choice. Despite excellent empirical comparisons, the study does not offer an explanation or downstream clinical tasks, such as prognosis. Aziz et al. (2023) [21] present a hybrid architecture that combines deep feature extraction with a multilayer perceptron classifier. Their work connects deep learning with classical neural networks and outperforms traditional classification methods. These findings demonstrate the importance of feature-level fusion in our work, which also concentrates on image-level classification rather than predictions at the patient level.

Despite advancements in osteosarcoma diagnosis by applying machine learning and deep learning, there remain significant challenges. The existing research is focused on deep end-to-end models, for example, CNNs 3-6, while the practical implementation of such a resource-intensive method (i.e., requiring large amounts of labeled data and computational resources) is almost impossible in a clinic. Furthermore, machine learning techniques generally use suboptimal hyperparameters, which can cause challenges like overfitting and unstable convergence. Additionally, metaheuristic optimization is not well-represented among histopathology applications, with limited fusion implementation for several ensemble learners. Additionally, in many studies, the required multiclass classification is not considered, concentrating instead on binary decision-making, with an apparent detriment for clinical applications. Addressing this gap directly motivates the proposed Hyper_ALO_HybML framework.

3. Proposed Methodology

The entire framework of the proposed Hyper_ALO_HybML for classifying osteosarcoma histopathology images into three classes, i.e., Non-Tumor, Viable Tumor, and Necrotic Tumor, is explained in this section. A novel pipeline, including preprocessing, radiomics feature extraction, and ensemble learning with improved meta-heuristic optimization, is proposed in Figure 1.

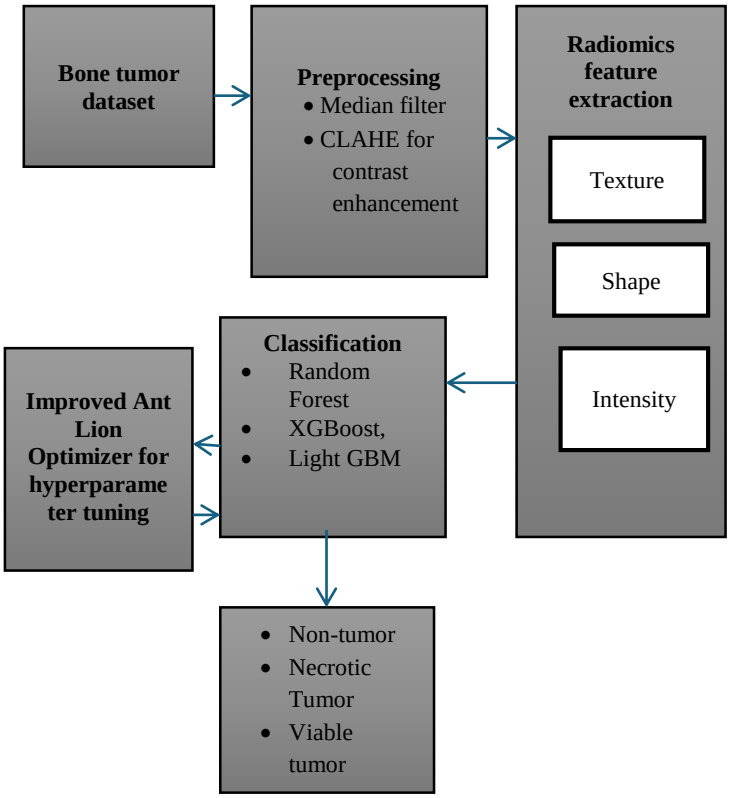


Fig. 1 Block diagram for osteosarcoma detection

The process begins with a bone tumor dataset consisting of histopathological images, which undergo pre-processing using a median filter for noise reduction and CLAHE for contrast enhancement [22]. The next step is to characterize tumor properties based on texture, shape, and intensity using radiomics feature extraction. The three machine learning models, Random Forest, XGBoost, and LightGBM, are then utilized to classify these extracted characteristics. Optimal parameter selection is ensured by using hyperparameter optimization with the Improved Ant Lion Optimizer to enhance the performance of these classifiers. Lastly, the trained model classifies the Tumor into one of three groups: Tumor, necrotic Tumor, or non-tumor.

3.1. Dataset

The dataset includes H&E-stained osteosarcoma histology [23]. In Dallas, UT Southwestern Medical Center clinical scientists collected the data and archived samples from 50 children who were treated at Children's Medical Center, Dallas, between 1995 and 2015. Pathologists selected four of 50 patients according to tumor type after surgery. They are labelled for the primary cancer type as Non-Tumor, Viable Tumor, and Necrosis. Two doctors did the annotation. All the images were split between two pathologists for annotation. Only one pathologist annotated each picture. The collection consists of 1144 images, at a resolution of 1024x1024 (10X) with the following distribution: 345 (30%) viable tumor slides, 263 (23%) Necrotic Tumor Images, and 536 (47%) Non-Tumor Images. The Sample Histopathological bone cancer image is shown in Figure 2.

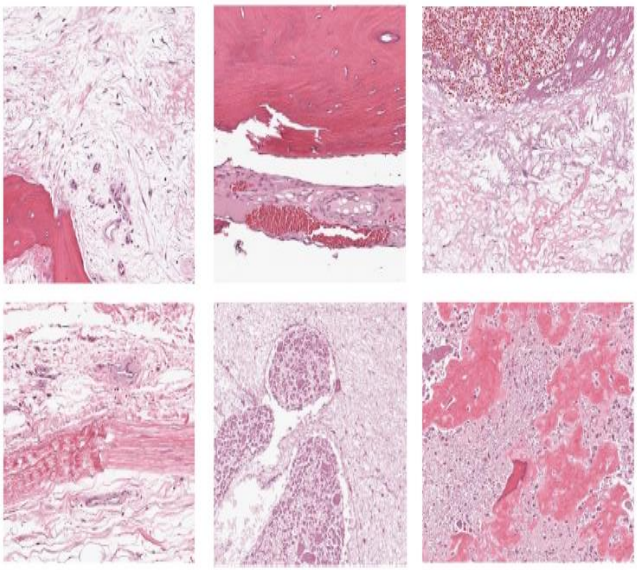


Fig. 2 Sample bone cancer histopathological images

3.2. Preprocessing of the Dataset

A Median Filter (MF) is an impulse noise remover that substitutes each pixel with the median intensity in its neighborhood [24]. This preserves edges while eliminating bright or dark outliers. The study assumes an input picture x of size M×N.

Where $x_{i,j}$represents the pixel location (i,j)

Where $(i,k) \in A = \{1,2,\dots M\} \times \{1,2 \dots N\}$. Below is a quick MF filter evaluation; To depict a window of size (2W+1)×(2W+1) placed at $(i,j), x_{i,j}(W)$ is defined as Equation 1.

$$x_{i,j}(W) = \{(a,b): |a-i| \leq W, |b-j| \leq W, (a,b) \in A\} \quad (1)$$

The median value of all pixel values is$x_{i,j}(W)$ arranged in ascending order $x_{i,j}^{med}(W)$. The working premise of MF (Median Filter) is given in Equations 2 and 3.

$$x_{i,j} = x_{i,j}^{med}(W) \quad (2)$$

$$\text{Here, } (i,j) \in A = \{1,2,\dots M\} \times \{1,2 \dots N\} \quad (3)$$

The image's irrelevant pixels are generally 0 or 255. MF may eliminate noise surrounded by irrelevant pixels or bright spots by utilizing the median value of the filter window pixel values instead of the current pixel. When MF is tiny, it may preserve the light spot's edge, contour, and other features since its size never changes throughout image processing. Imagine an image having a grayscale range of $[0, n-1]$ and a gray level of m. Gray-level PDF is $p_r(r)$. Then, the kth gray level probability is given in Equation 4.

$$p_r(r_k) = \frac{n_k}{n} \quad (4)$$

Where $k = 0,1,2 \dots m-1$ r_k is the k th grayscale.

The Cumulative Distribution Function (CDF) in Equation 5 $T(r_k)$ is then

$$g = T(r_k) = \sum_{j=0}^{k} p_r(r_j) = \sum_{j=0}^{k} \frac{n_j}{n} \quad (5)$$

Where, $0 \leq g \leq 1$. The steps of CLAHE are

1. Divide the image into non-overlapping and continuous parts. An 8x8 area size is typical.
2. Clip the histogram of each section using the threshold. The CLAHE (Contrast Limited Adaptive Histogram Equalization) method limits magnification by clipping the histogram with a threshold before computing the CDF. The transformation function slope is likewise limited.
3. Reallocate and uniformly distribute clipped pixel values below the histogram.
4. Equalize all areas' local histograms.
5. Reconstruct pixel values using linear interpolation.

CLAHE enhances local contrast by performing histogram equalization for small image patches and controlling noise

amplification by a predefined clipping limit. [25]. This enhances the visibility of Tumor Nuclei, Necrotic Regions, and Connective Tissue Patterns.

3.3. Radiomics Feature Extraction

The radiomic model uses quantitative medical imaging. The emerging field of radiomics employs many autonomously produced data characterization techniques to put imaging data into a high-dimensional, mineable feature space [26]. These imaging parameters may capture tumor phenotypes and have predictive power and clinical importance across diseases. Extraction of textural features on the seven extracted ROIs (four lumbar: L1-L4 and three femoral: trochanteric, intertrochanteric, neck). All ROIs were manually drawn by a radiologist with 10 years of experience (A. A.), who specializes in Bone Cancer analysis and reporting. ROI sizes varied by area, averaging 2 ± 0.77 (SD) cm (range: 1 to 3 cm).

The pixel size was 1 mm^2. Each ROI yielded 54 texture characteristics. These were 12 shape, 20 texture, and 8 intensity histogram characteristics. On T1W, T2W, and CE-T1W images, intraosseous, extraosseous, and total lesion ROIs were isolated separately for texture characteristics. Two-dimensional processing calculated texture characteristics using 2D pixels and their x- and y-directions. A 2D ROI needs 16 pixels. Texture features were continually computed from ROIs for all chosen slices and averaged. Thus, the extracted feature is Equation 6, where RFS is the Radiomic Feature Score.

$$RFS = w_1.F_{shape} + w_2.F_{intensity} + w_3.F_{texture} \quad (6)$$

The Denotation of the feature set and its description are shown in Table 1.

Table 1. Denotation of the feature set and its description

Feature Set	Description
Intensity	Mean Variance Skewness Kurtosis
shape	Perimeter Area Eccentricity
Texture - Gray level run-length matrix	RLNonUni GLevNonU LngREmph ShrtREmp
Gray-level co-occurrence matrix	Contrast Correlat SumOfSqs InvDfMom

Radiomics is an approach of transforming image regions into a set of quantitative values that characterize intensity, texture, and shape. The ROIs corresponding to Osteosarcoma were manually contoured by an experienced radiologist in the lumbar and femoral areas in the present study. Fifteen features derived were: eight intensity-based ones [i.e., Mean, Variance, Skewness, and Kurtosis], twelve shape-based (Area, Perimeter, Eccentricity, and Compactness), and thirty-tour texture-based calculated from Gray Level Co-Occurrence Matrix (GLCM), Gray Level Run Length Matrix (GLRLM) [27]. The Textural Parameters include Contrast, Correlation, Energy, Inverse Difference Moment (IDM), Short Run Emphasis (SRE), Long Run Emphasis (LRE), and Non-Uniformity. The parameters provide a good coverage of intra-tumoural Heterogeneity, Necrosis Patterns, and the differences in morphology between tumour vs. non-tumour tissue.

3.4. Ensemble Classification

The processed radiomic features were classified using three complementary ensemble models: Random Forest (RF), XGBoost (XGB), and LightGBM (LGBM). Hyperparameters for each model were optimized using the Improved Ant Lion Optimizer (ALO) method, used to create a bone cancer prediction model in Figure 3.

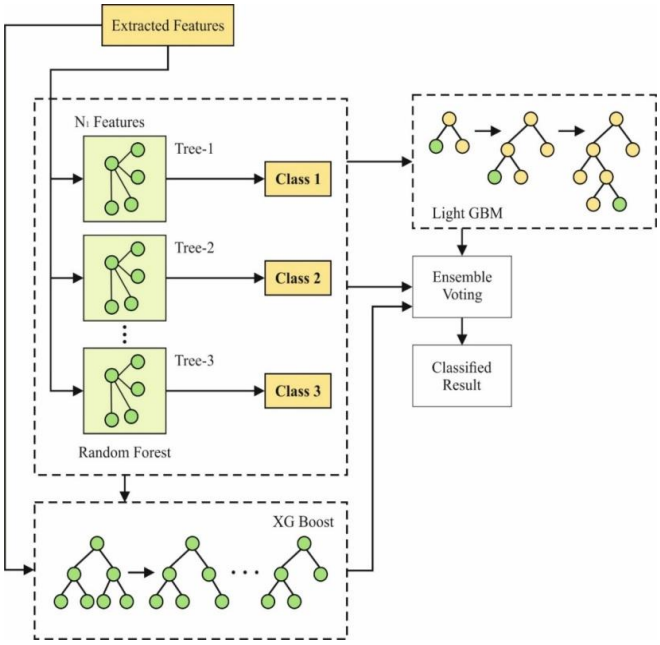


Fig. 3 Architecture diagram of hybrid machine learning classifiers

3.4.1. Random Forest

Random forest is a well-known Machine Learning method for classification, regression, and others [28]. Through the bagging-based method, the random forest algorithm clusters the training data set repeatedly for training based on the original data, and then models each cluster respectively to produce a decision model; finally , it summarizes all decision behaviors to generate the ultimate model. Based on the voting mechanism, the random forest algorithm predicts and outputs

the classification with the most votes. The error from rearranged feature inputs may be used to determine feature significance in the RF. Definition of feature importance: RF_a with the following Equation 7.

$$RF_a(x_j) = \frac{1}{T}\sum_{RFS\in C}\frac{1}{|RFS|}\left(\sum_{i\in RFS} I(h_k^{x_j}(i) \neq y_i - I(h_k(i) \neq y_i)\right) \quad (7)$$

Where y_i, i is the label category for the i[th] Out-Of-Bag (OOB) data, $I(x)$ is the indicator function, and $h_k(i)$ is the prediction sample label function. The dataset RFS sample is used $h_k^{x_j}(i)$ as the classification label.

3.4.2. XGBoost

The ensemble machine learning based gradient boosting system XGBoost (XGB) uses decision trees [29]. The excellent processing speed and scalability of this Machine Learning method should provide worthwhile results. Regression and classification employ XGB. This method repeatedly identifies weak classifiers to achieve accurate classification. It creates personalized decision trees by defining threshold values and reducing residuals while building trees using gradient descent. Regression trees are weak learners with gradient boosts that map input datasets with continuous marks to one leaf. The prediction target output difference is used to minimize L1 and L2 using a convex loss function. The training repeatedly builds new trees to forecast prior tree residues or errors, which are then merged to make the final prediction. This is mathematically expressed as Equation 8.

$$F_m(X) = F_{m-1} + \propto_m h_m(X, r_{m-1}) \quad (8)$$

Where, $\propto_m$ and r_m are regularization parameters, residuals from the i[th] tree are calculated, and h_m It is a trained function to forecast residuals. r_m using X for the i tree.

3.4.3. LightGBM

LightGBM generates more complicated trees using leaf-wise splits, lowering loss and improving accuracy [30]. Gradient-based One-Sided Sampling (GOSS) is used for partitioning. This approach excludes data points with minor gradients and uses the remaining for information gain estimation and tree building. The loss function gradients are calculated mathematically in Equation 9.

$$|g_{bc}| = |g_{bc}| + \tau \times max(|g_{bc}|, \forall) \quad (9)$$

In classification, $|g_{bc}|$ represents the loss function gradient, τ controls one-sided sampling, and ∀ sets the gradient threshold. The suggested method uses the remaining information to build a conclusion and eliminates minor gradients. Information gain may be important in resource-constrained contexts. It also takes less training and memorization.

3.4.4. Improved Ant Lion Optimizer for Hyperparameter Tuning

The Improved ALO algorithm advances the traditional Ant Lion Optimizer (ALO) by enhancing the balance between exploration and exploitation. It incorporates several key strategies: adaptive shrinking of the random walk boundaries to optimize search capabilities, a modified elitism strategy to retain better solutions, and boundary-preserving mechanisms to avoid invalid parameter values, thereby improving overall performance in optimization tasks.

Ant lions naturally impact ant walks. When the ant enters the trap, the ant lion throws sand to lower it. Let I be the sliding ratio and represent it mathematically expressed in Equations 10 and 11.

$$c_i^t = antlion_j^t + c^t \quad (10)$$

$$d_i^t = antlion_j^t + d^t \quad (11)$$

Where I id $c^t = \frac{c^t}{I}$ and $d^t = \frac{d^t}{I}$. After that, the ant lion's position is updated. This updating mechanism uses a mathematical model to randomly change option parameters expressed in Equations 12,13,14, and 15. The trap ants' status determines the ant lion's location in this model.

$$\begin{cases} c_i^t = antlion_j^t + c^t \\ d_i^t = antlion_j^t + d^t \end{cases} if\ option > 0.75 \quad (12)$$

$$\begin{cases} c_i^t = antlion_j^t - c^t \\ d_i^t = antlion_j^t - d^t \end{cases} if\ option > 0.5 \quad (13)$$

$$\begin{cases} c_i^t = -antlion_j^t + c^t \\ d_i^t = -antlion_j^t + d^t \end{cases} if\ option > 0.25 \quad (14)$$

$$\begin{cases} c_i^t = -antlion_j^t - c^t \\ d_i^t = -antlion_j^t - d^t \end{cases} otherwise \quad (15)$$

The following mathematical Equation 16 prevents ants from fleeing the search space bounds, unlike the original antlion optimization technique.

$$ant_i^t = k\ if\ \left(ant_i^t > b_{up}\right) OR (ant_i^t < b_{low}) \quad (16)$$

The mathematical Equation 17 expression k permits ants to remain in the search space at random outside the boundary.

$$k = b_{low} + rand \times (b_{up} - b_{low}) \quad (17)$$

Where $rand$ is a random integer in the [0, 1] interval b_{low}, lower limit v, and higher bound. The original antlion optimization technique groups ant and antlion populations by

cost value and takes as many as the population size as antlions. The proposed method removes cost sorting and replaces ants with antlions for each ant-antlion pair if their cost values are superior. This indicates antlions ate ants.

3.4.5. Pseudocode for Ensemble Classification with Antlion Optimizer

Initialize RF model

$RF_{pred} \leftarrow RF.pred(X_{test})$

Initialize XGB model

$XGB_{pred} \leftarrow XGB.pred(X_{test})$

Initialize the LGBM model.

$LGBM_{pred} \leftarrow LGBM.pred(X_{test})$

Initialize hyperparameter tuning

start ants and antlions randomly within the search space

Evaluate fitness $F(x)$ for each ant and antlion

For each ant i, compare fitness with the assigned $antlion\ j$

It $ant\ i$ has better fitness than $antlion\ j$

Replace $antlion\ j$ with $ant\ i$

End For

Return Elite Antlion (optimal solution)

$P_{final} \leftarrow (w_1 * RF_{pred}) + (w_2 * XGB_{pred}) + (w_3 * LGBM_{pred})$

3.5. Performance Analysis

The Model simulations in a Google Colab free environment (8 GB RAM, 80 GB HDD, cloud GPU). Even with computationally expensive hyperparameter adjustments, this configuration made model training and assessment fast. The interpretability and repeatability of the findings hinge on the chosen hyperparameters and the computing environment.

The accuracy metric assesses image classification model performance. Precision and recall measures assess the model's false positive and negative elimination. The f1-score assesses precision, recall, and balance. It is useful when both measures are equally important.

Model specificity measures its capacity to recognize benign cases. Specificity is calculated by dividing True Negatives (TN) by the sum of true negatives and false positives.

The performance metric is given in Equations 18, 19,20, and 21.

$$Accuracy = \frac{TP+TN}{TN+TP+FP+FN} \quad (18)$$

$$precision = \frac{TP}{TP+FP} \quad (19)$$

$$recall = \frac{TP}{TP+FN} \quad (20)$$

$$f1-score = 2 \times \frac{recall*preciosn}{recall+precision} \quad (21)$$

4. Result and Discussion

In this section, the performance of three optimal classifiers, including RF, XGB, and LGBM, is re-evaluated on another dataset related to osteosarcoma histopathology. The model's performance was evaluated by accuracy, precision, sensitivity, specificity, F1-score, and MCC, and the confusion matrix was also exploited to understand misclassification.

4.1. Performance of Random Forest (RF)

The RF classifier exhibits moderate capability in classifying the three tissue classes. The confusion matrix shows that the non-tumor class is better detected compared to the Viable Tumor and Necrotic Tumor classes.

A Viable Tumor was misclassified as non-tumor multiple times, indicating that RF has difficulty discriminating between subtle textural differences of Tumor and Normal Tissues.

The diagonal values in Figure 4, Such as 53, 151, and 89, indicate correct classifications for each category, with the model performing best in identifying Non-Tumor Cases. 13 cases were misclassified as Non-Tumor, and another 13 as Viable-Tumor. 8 cases were incorrectly predicted as Viable-Tumor, and 2 as Necrotic-Tumor. 13 cases were misclassified as Non-Tumor, and 1 as Necrotic-Tumor. The Analysis of accuracy and computation for RF is shown in Table 2.

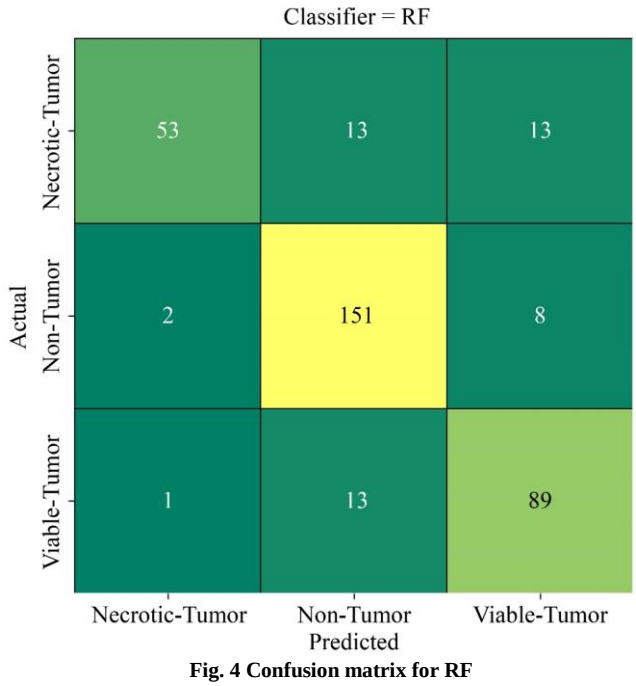


Fig. 4 Confusion matrix for RF

Table 2. Analysis of accuracy and computation for RF

Number of epochs	n_estimators	criterion	max_leaves	Accuracy	computation
1	45	entropy	18	0.822	1.44
2	46	entropy	18	0.81	3.78
3	46	entropy	18	0.83	4.97
4	46	entropy	18	0.86	6.34
5	46	entropy	18	0.857	5.82
6	46	entropy	18	0.854	4.95
7	46	entropy	18	0.86	4.85
8	46	entropy	18	0.86	4.91
9	46	entropy	18	0.86	4.70
10	46	entropy	18	0.86	2.59

Our tenfold results indicate that RF gives accuracy in the range of 0.81 to 0.86, with the best obtained value being 0.86. While the model enjoys entropy-based splitting and a fairly stable number of estimators, there is not much improvement in the overall result with more iterations. This is consistent with its lower precision and recall in the comparative experiment, stipulating that RF may be worse suited for fine-grained histopathological classification than boosters.

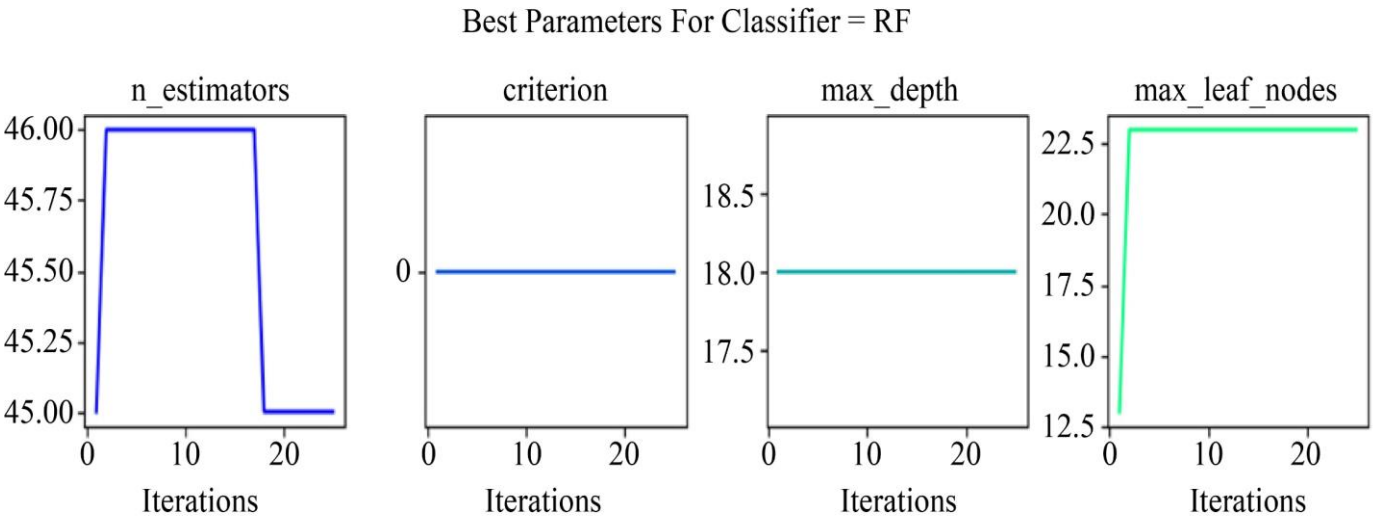


Fig. 5 Best parameter for RF

The hyperparameter tuning for the Random Forest (RF) classifier in Figure 5 identified the optimal values for key parameters, ensuring a balance between accuracy and efficiency. These tuned parameters collectively enhance the RF classifier's predictive accuracy while minimizing unnecessary computational overhead.

4.2. Performance of LGBM (LightGBM)

LightGBM also performs well and remains competitive with XGBoost. The confusion matrix indicates strong classification of Viable Tumor and Non-Tumor samples, with few misclassifications in the Necrotic Tumor class. This suggests that LightGBM efficiently captures both local and global radiomic patterns.

The confusion matrix of Figure 6 describes three classes of Tumor Conditions (Necrotic-Tumor (NT), Non-Tumor (Non-NT), and Viable-Tumor). Block 157 indicates accurately categorized Non-Tumor Samples. The 96th block represents appropriately categorized Viable-Tumor instances. Number 2 represents the number of misclassified Non-Tumor Cases that were incorrectly predicted as Viable-Tumor, as shown in Table 3.

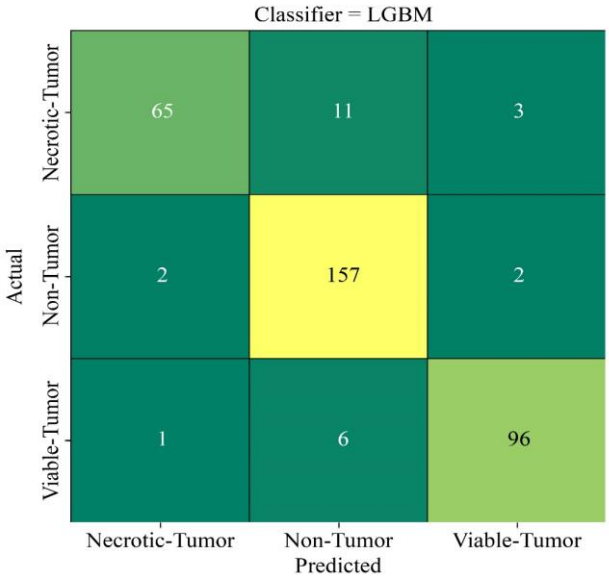


Fig. 6 Confusion matrix for LGBM

Table 3. Analysis of accuracy and computation for LGBM

Number of epochs	n_estimators	max_depth	max_leaves	Learning_rate	Accuracy	computation
1	54	5	2	0.00945	0.8688	1.58
2	100	10	7	0.006390	0.46	2.33
3	81	7	4	0.007665	0.89	4.482
4	81	7	4	0.00766	0.895	4.857
5	81	7	4	0.00766	0.8892	4.465
6	84	8	5	0.0819	0.8950	4.3291
7	87	8	5	0.08537	0.9067	4.5444
8	87	8	6	0.00853	0.9008	4.9087
9	87	8	5	008537	0.9067	4.8156
10	89	8	5	0.008659	0.9154	5.2023

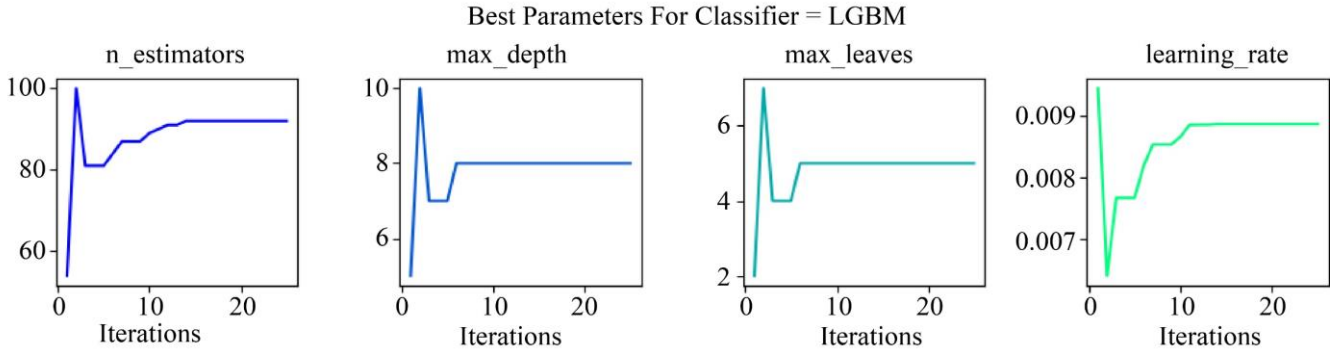


Fig. 7 Best parameter for LGBM

In the evaluation, LightGBM reaches a maximum accuracy of 0.927, with consistent learning rate optimization seen in the hyperparameter plots in Figure 7. Its F1-score (0.927) and MCC (0.8858) are close to those of XGB, demonstrating strong overall generalization. The number of estimators stabilizes around 90-100**,** ensuring sufficient learning without excessive complexity. The maximum tree depth and maximum leaves per tree converge to 7-8 and 5**,** respectively, balancing model performance and overfitting. The learning rate stabilizes around 0.009**,** ensuring controlled updates for better generalization. These optimizations enhance the classifier's efficiency and accuracy.

4.3. Performance of ALO-XGB

XGBoost demonstrates the strongest performance among the three classifiers. The confusion matrix in Figure 8 shows that it accurately identifies all three classes with fewer misclassifications than RF or LGBM. In particular, Non-Tumor and Viable Tumor samples show high correct classification rates, reflecting XGB's ability to separate high-dimensional radiomic features effectively.

The confusion matrix noted that 68 cases were correctly classified, while 6 were misclassified as non-tumor and 5 as Viable-Tumor. The model performed exceptionally well, correctly classifying 157 cases, with only 4 misclassified as Viable-Tumor and none as Necrotic-Tumor. 95 cases were accurately identified, while 7 were misclassified as Non-Tumor and 1 as Necrotic-Tumor.

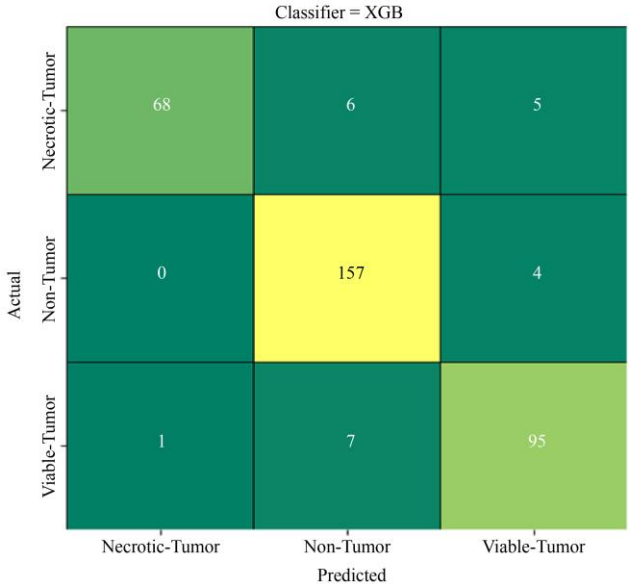


Fig. 8 Confusion matrix for XGB

Table 4. Analysis of accuracy and computation for XGB

Number of epochs	n_estimators	Max_depth	max_leaves	Learning rate	Accuracy	computation
1	100	25	23	0.01	0.830	6.080
2	100	25	23	0.01	0.9271	10.190
3	100	25	23	0.01	0.9271	10.941
4	100	25	23	0.01	0.92711	10.250
5	100	25	23	0.01	0.924	6.29
6	100	25	23	0.01	0.924	6.845
7	100	25	23	0.01	0.9241	9.613
8	100	25	23	0.01	0.9300	9.770
9	100	25	23	0.01	0.9241	10.450
10	100	25	23	0.01	0.9241	9.447217699984321

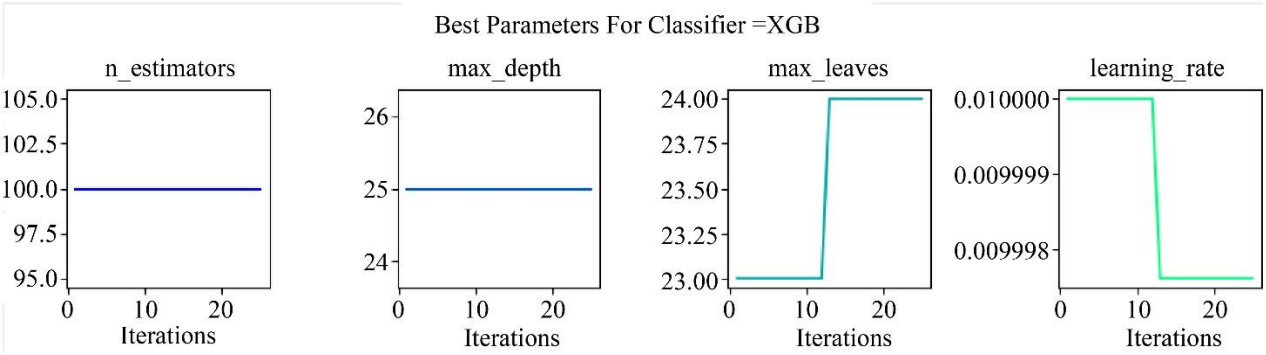


Fig. 9 Best parameter for XGB

XGB keeps the hyperparameters fixed and learned using the improved Ant Lion Optimization (ALO) over 10 iterations. The best result is obtained by our method (Acc. = 0.933) with much better convergence and stability, as shown in the Figure 9. The precision (0.9408), sensitivity (0.9194), specificity (0.9624), and MCC (0.8947) of the model were also higher than those of other models, as presented in Table 4. These results validate that XGB harnesses the boosted decision tree method well, and there is considerable gain from the fine-tuning of learning rate and depth values.

From Figure 9, the XGB classifier shows that the optimal number of estimators is 100**,** indicating a stable model without overfitting. The max_depth is set to 25**,** allowing deeper decision trees for capturing complex patterns. The max_leaves parameter initially starts at 23 and stabilizes at 24**,** optimizing tree complexity for better performance. The learning rate begins at 0.01 but later adjusts slightly below 0.009998**,** ensuring a controlled gradient descent for optimal convergence.

4.4. Comparative Analysis

The comparative evaluation of Random Forest, XGBoost, and LightGBM classifiers is shown in Figure 10, and the comparative evaluation of all classifiers is shown in Table 5. The XGBoost achieves the highest accuracy of 0.933, followed closely by LightGBM at 0.927, while Random Forest lags at 0.854. For sensitivity or recall, XGBoost and LightGBM show similar values of 0.919 and 0.910, respectively, significantly outperforming Random Forest, which scores 0.824. The F1-score further confirms the balanced performance of XGBoost and LightGBM, with scores of 0.933 and 0.927, significantly ahead of Random Forest at 0.854. Lastly, the Matthews Correlation Coefficient further highlights the effectiveness of XGBoost and LightGBM, with scores of 0.895 and 0.886, respectively. Overall, XGBoost demonstrates the best performance across all metrics, followed closely by LightGBM, while Random Forest performs the worst in comparison. Table 6 describes the classified images of Necrotic-Tumor, Non-Tumor, and Viable-Tumor in the analysis of the pre-processed original image.

Table 5. The overall comparative evaluation of all classifiers

Classifier	Accuracy	Precision	Sensitivity	specificity	F1-score	MCC
RF	0.8542	0.8695	0.8243	0.9194	0.8542	0.771
LGBM	0.9271	0.9362	0.91	0.9581	0.9271	0.8858
ALO-XGB	0.9329	0.9408	0.9194	0.9624	0.9329	0.8947

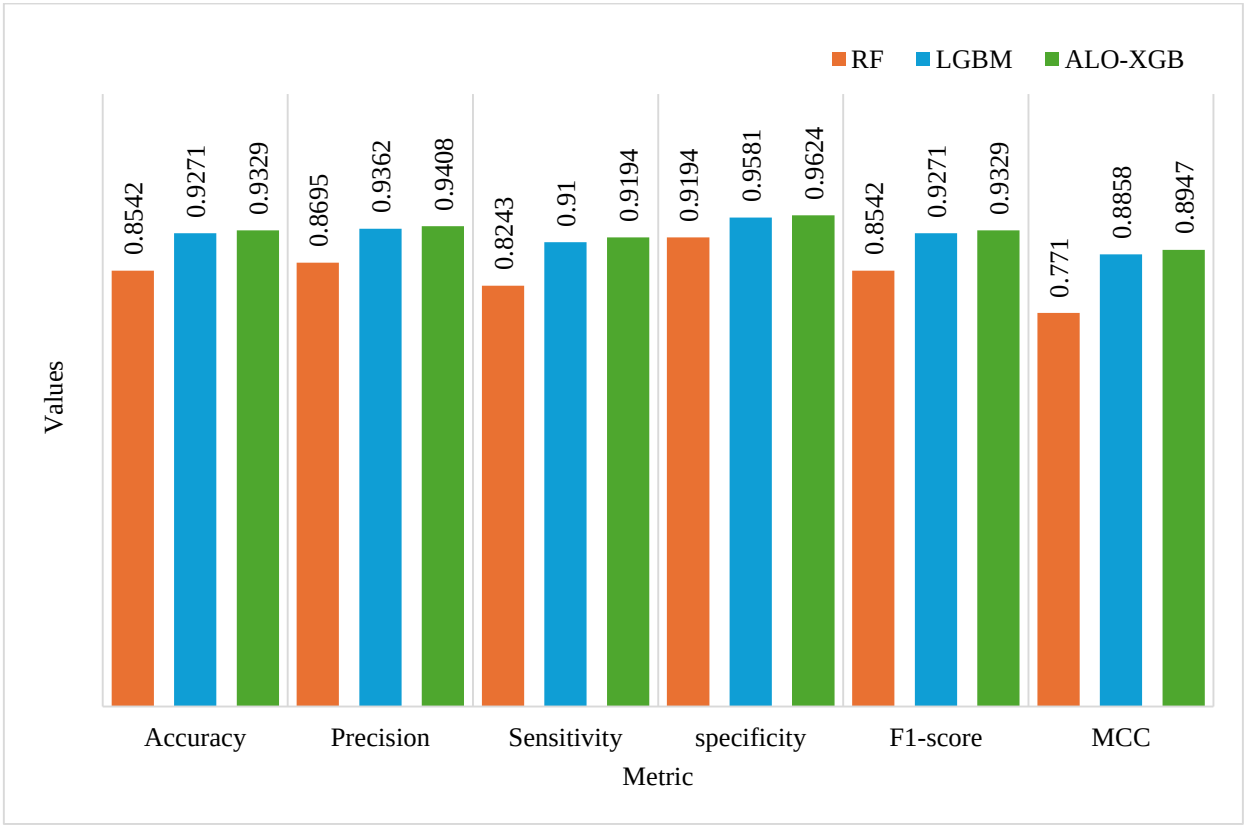
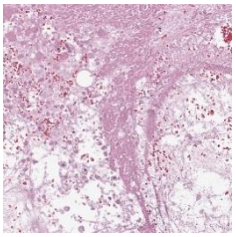
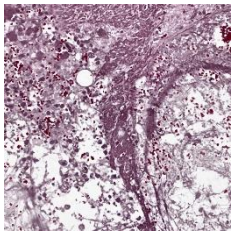
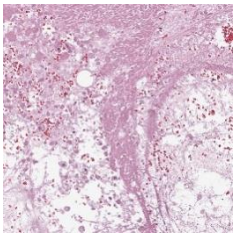
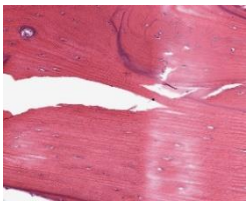
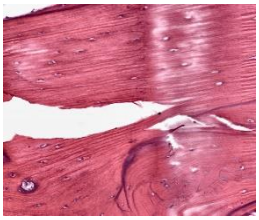
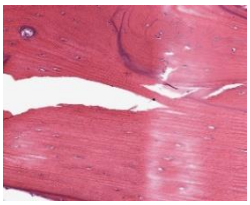
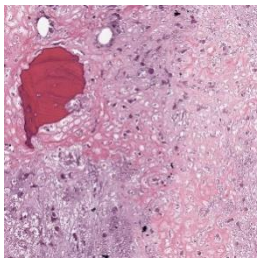
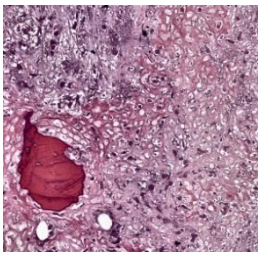
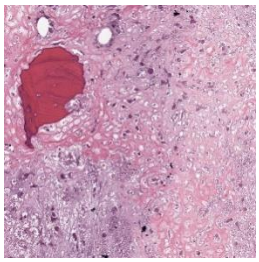


Fig. 10 Overall comparative analysis

Table 6. Classified images of necrotic-tumor, non-tumor, viable-tumor, and analysis of the pre-processed original image

Classes	Raw_image	Pre-processes	Original Image
Necrotic-Tumor			
Non-Tumor			
Viable-Tumor			

The comparative bar graph of Figure 10 indicates that XGBoost is a better model in terms of all evaluation metrics. Its steady performance demonstrates the success of strong ensemble learning combined with ALO-based hyperparameter search. LightGBM comes second with a slightly better score, still far beyond what RF achieves.

The lowest performances are obtained by Random Forest, partly because of the difficulty in modeling complex textural and structural variations of histopathology patterns.

The newly proposed Hyper_ALO_HybML framework provides a fundamentally new design philosophy from others in osteosarcoma detection. Although most of the existing studies are related to end-to-end deep learning architectures or single-model classifiers, the significance of radiomics-driven and optimization-guided ensemble materializes in this work, towards robust multiclass osteosarcoma classification.

The main novelty is that the radiomic features, ensemble models, and metaheuristic optimization methods are not considered as separate steps in isolation. A system that is more stable, interpretable, and less data-dependent than Deep Neural Networks.

5. Conclusion

The performance of the Hyper_ALO_HybML method is significantly better for multiclass classification of Osteosarcoma when using the radiomic features associated with ALO-optimize ensemble models. XGBoost was the best among our tested classifiers, achieving 93.29% accuracy as well as the highest F1 and MCC scores, closely followed by LightGBM. These findings underscore the advantage of using metaheuristic optimization in ensemble learning, particularly when dealing with complex tissue patterns and class imbalance. The approach mitigated overfitting risk, ensured stable convergence, and enhanced the interpretability of challenging histopathological images. With more extensive validation on larger and diverse datasets, this is a promising approach for deployment in clinical settings. Pathologist collaboration and real-world testing should be the critical next steps to fine-tune the system for robust performance in a diagnostic setting.

Acknowledgments

I want to express my very great appreciation to My Guide, Mahesh Babu Arrama, and Co-Guide Jayaprakash Chennoju for their valuable and constructive suggestions during the planning and development of this research paper.

References

[1] Khan Muhammad et al., "Deep Learning for Multigrade Brain Tumor Classification in Smart Healthcare Systems: A Prospective Survey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 2, pp. 507-522, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[2] Majidur Rahman et al., "Symptom Wise Age Prediction of Cancer Patients using Classifier Comparison and Feature Selection," *2019 22nd International Conference on Computer and Information Technology (ICCIT)*, Dhaka, Bangladesh, pp. 1-6, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[3] Cancer Stat Facts: Bone Cancer, National Cancer Institute, Surveillance, Epidemiology, and End Results Program, 2021. [Online]. Available: https://seer.cancer.gov/statfacts/html/bones.html

[4] M. Gaume et al., "The Appropriate and Sequential Value of Standard Radiograph, Computed Tomography and Magnetic Resonance Imaging to Characterize a Bone Tumor," *Scientific Reports*, vol. 12, pp. 1-9, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[5] Jiachen Sun et al., "Progress of Photodynamic Therapy Applications in the Treatment of Bone Cancer," *International Journal of Molecular Sciences*, vol. 22, no. 21, pp. 1-38, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[6] P. Naveen, and B. Diwan, "Pre-Trained VGG-16 with CNN Architecture to Classify X-Rays Images Into Normal or Pneumonia." *2021 International Conference on Emerging Smart Computing and Informatics (ESCI)*, Pune, India, pp. 102-105, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[7] M.M. Ranjitha et al., "Bone Cancer Detection Using K-Means Segmentation and Knn Classification," *2019 1st International Conference on Advances in Information Technology (ICAIT)*, Chikmagalur, India, pp. 76-80, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[8] S. Famila et al., "RETRACTED ARTICLE: Transfer Learning Based Feature Extraction with Metaheuristic Optimization Algorithm for Detecting Gastric Cancer Using Optoelectronic Sensor in Endoscope," *Optical and Quantum Electronics*, vol. 56, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[9] K.V. Deepak, and R. Bharanidharan, "A Survey on Deep Learning and Machine Learning Techniques Over Histopathology Image Based Osteosarcoma Detection," *Multimedia Tools and Applications*, vol. 84, pp. 15763-15786, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[10] Pawan Whig et al., *Bone Cancer Classification and Detection Using Machine Learning Technique*, Diagnosing Musculoskeletal Conditions using Artifical Intelligence and Machine Learning to Aid Interpretation of Clinical Imaging, Academic Press, pp. 65-80, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[11] Arezoo Borji et al., "Advanced Hybrid Deep Learning Model for Enhanced Evaluation of Osteosarcoma Histopathology Images," *Frontiers in Medicine*, vol. 12, pp. 1-19, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[12] R. Aarthy, and V. Muthupriya, "Enhanced Detection and Classification of Osteosarcoma Bone Cancer Using Discriminative Patches and GAN-Screened Imaging," *Intelligent Decision Technologies*, vol. 19, no. 6, pp. 4130-4146, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[13] Nicolas Coudray et al., "Quantitative and Morphology-Based Deep Convolutional Neural Network Approaches for Osteosarcoma Survival Prediction in the Neoadjuvant and Metastatic Settings," *Clinical Cancer Research*, vol. 31, no. 2, pp. 365-375, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[14] Maurizio Cè et al., "Multimodal Imaging of Osteosarcoma: From First Diagnosis to Radiomics," *Cancers*, vol. 17, no. 4, pp. 1-35, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[15] Platon S. Papageorgiou et al., "Artificial Intelligence in Primary Malignant Bone Tumor Imaging: A Narrative Review," *Diagnostics*, vol. 15, no. 13, pp. 1-24, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[16] Shtwai Alsubai et al., "Group Teaching Optimization with Deep Learning-Driven Osteosarcoma Detection Using Histopathological Images," *IEEE Access*, vol. 12, pp. 34089-34098, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[17] Qian Liu, Xing She, and Qian Xia, "AI based Diagnostics Product Design for Osteosarcoma Cells Microscopy Imaging of Bone Cancer Patients using CA-MobileNet V3," *Journal of Bone Oncology*, vol. 49, pp. 1-9, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[18] Clare McGenity et al., "Artificial Intelligence in Digital Pathology: A Systematic Review and Meta-Analysis of Diagnostic Test Accuracy," *NPJ Digital Medicine*, vol. 7, pp. 1-19, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[19] Xiao Zhao et al., "Deep Bone Oncology Diagnostics: Computed Tomography based Machine Learning for Detection of Bone Tumors from Breast Cancer Metastasis," *Journal of Bone Oncology*, vol. 48, pp. 1-10, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[20] Ioannis A. Vezakis, George I. Lambrou, and George K. Matsopoulos "Deep Learning Approaches to Osteosarcoma Diagnosis and Classification: A Comparative Methodological Approach," *Cancers*, vol. 15, no. 8, pp. 1-15, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[21] Tarek Aziz et al., "A Novel Hybrid Approach for Classifying Osteosarcoma using Deep Feature Extraction and Multilayer Perceptron," *Diagnostics*, vol. 13, no. 12, pp. 1-27, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[22] Yiran Wang et al., "Comprehensive Diagnostic Model for Osteosarcoma Classification using CT Imaging Features," *Journal of Bone Oncology*, vol. 47, pp. 1-8, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[23] Patrick Leavey et al., "Osteosarcoma-Tumor-Assessment | Osteosarcoma Data from UT Southwestern/UT Dallas for Viable and Necrotic Tumor Assessment (Data Set)," *The Cancer Imaging Archive*, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[24] Kanimozhi Sampath, Sivakumar Rajagopal, and Ananthakrishna Chintanpalli, "A Comparative Analysis of CNN-based Deep Learning Architectures for Early Diagnosis of Bone Cancer using CT Images," *Scientific Reports*, vol. 14, pp. 1-9, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[25] Faikah Zakaria, and Hamizah Emran, "Stress Fracture Detection using Adaptive Histogram Equalization (Ahe) and Contrast Limited Adaptive Histogram Equalization (Clahe): A Comparative Study," *Health Scope*, vol. 3, no. 1, pp. 47-50, 2020. [Google Scholar] [Publisher Link]

[26] Ashish Sharma et al., "Bone Cancer Detection using Feature Extraction based Machine Learning Model," *Computational and Mathematical Methods in Medicine*, vol. 2021, no. 1, pp. 1-13, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[27] Jinlian Jin et al., "Machine Learning Based Gray-Level Co-Occurrence Matrix Early Warning System Enables Accurate Detection of Colorectal Cancer Pelvic Bone Metastases on MRI," *Frontiers in Oncology*, vol. 13, pp. 1-10, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[28] Samir D. Mehta, and Ronnie Sebro, "Random Forest Classifiers Aid in the Detection of Incidental Osteoblastic Osseous Metastases in DEXA Studies," *International Journal of Computer Assisted Radiology and Surgery*, vol. 14, pp. 903-909, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[29] Meiling Chu et al., "A Study on the Development of a Machine Learning-Based Clinical Prediction Model for Bone Health Management in Breast Cancer Patients," *Research Square*, pp. 1-20, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[30] Bassam M. Kanber et al., "LightGBM: A Leading Force in Breast Cancer Diagnosis through Machine Learning and Image Processing," *IEEE Access*, vol. 12, pp. 39811-39832, 2024. [CrossRef] [Google Scholar] [Publisher Link]