International Journal of Engineering Trends and Technology
ISSN: 2231-5381/ https://doi.org/10.14445/22315381/IJETT-V74I7P112

Volume 74 Issue 7, 179-191, July 2026
© 2026 Seventh Sense Research Group®

Original Article

Bone Tumor Grading Using Radiomic-Guided Feature Modulation: In A Compact Deep Learning Model

Rathla Roopsingh[1], D. Vasumathi[2]

[1,2]*Department of CSE, Jawaharlal Nehru Technological University, Hyderabad, Telangana, India.*

[1]*Corresponding Author : rathlaroopsingh@gmail.com*

Received: 23 March 2026 Revised: 16 June 2026 Accepted: 18 June 2026 Published: 28 July 2026

Abstract - *Accurate grading of tumors in the bones is very important in delivering treatment methods and also giving a prognosis to the patient. Conventional classification techniques usually cannot cope with the heterogeneity of the tumor picture and the absence of annotated data. This research work presents a new lightweight deep learning architecture called Radiomic-Modulated Deep Network (RMD-Net) that is aimed at improving the performance of tumor grading based on the combination of radiomic and deep visual image representations. The model utilizes radiomic descriptors in the form of shape, intensity, and texture measurements of segmented tumor regions, and is trained to modulate deep feature activations produced by a shallow convolutional or transformer-based backbone dynamically as symbols of target changes. The resulting radiomic-guided modulation will provide interpretability of the features and better Modularity of the classes because the learning process encapsulates clinically useful properties. A vast amount of experiments on a curated set of CT and MRI scans show that RMD-Net is better at multi-class bone tumor grading tasks, making it more accurate and generalizing with much less parameters. The suggested framework is an effective, interpretable, and clinically flexible solution to assist radiologists in the non-invasive measures of the severity of bone tumors.*

Keywords - *Bone Tumor grading, Radiomic Features, Deep learning, Feature Modulation, Compact Network, Medical Image Classification, RMD-Net, Clinical Decision Support.*

1. Introduction

Bone tumors are a heterogeneous category of neoplasms whose malignancy can be benign lesions to borderline lesions, or even malignant high-grade tumors. The diagnosis of the tumor grade is a vital element of clinical oncology; it informs the decisions of biopsy and surgery, Chemo therapy, and estimation of the prognosis [1]. Although the gold standard of tumor grading is histopathological analysis, the tissue observation approach is invasive and time-consuming and is prone to interobserver differences [2]. New opportunities of rad imaging and machine learning have led to non-invasive characterization of tumors based on imaging biomarkers [3].

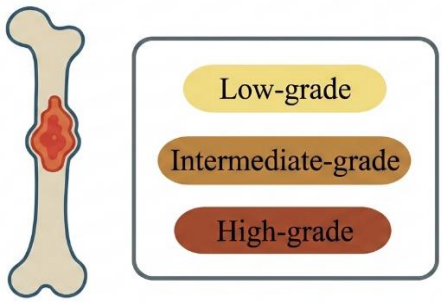

Fig. 1 Bone tumor grading illustration with visual classification

Figure 1 depicts the anatomical position of a bone tumor, together with distinctly marked grading categories of Low-grade, Intermediate-grade, and High-grade. The color-coded labels also increase the interpretability for clinical application [4]. The figure helps to imagine the severity of tumors to diagnose and educate on them [5]. In recent times, image-based medical diagnosis with deep learning models [6, 7], in particular, Convolutional Neural Networks (CNNs), has demonstrated a strong potential. Their performance in tumor grading is, however, often poor with the standard architectures as they are not robust enough to capture the subtle inter-class variation, small-scale datasets, or to fuse with domain-related cues [8]. Conversely, radiomic features are quantitative characteristics of medical images, which are structural, textural, and statistical features of tumors that have been observed to be correlated with histological grades [9]. In order to narrow the divide between handcrafted radiomics and deep representations, we suggest the Radiomic-Modulated Deep Network (RMD-Net), which adds a feature modulation mechanism to the deep learning process to inculcate radiomic awareness. RMD-Net Protonymous to both traditional approaches of concatenation of features and late fusion, which simplify the discovery of radiomic features, RMD-Net is a learning approach in which radiomic features are used to

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

dynamically regulate the discovery of deep visual features with a modulation layer. The design enables the network to concentrate on radiologically significant patterns whilst being computationally lightweight and efficient to use in a clinical setting [10].

This paper outlines the design of RMD-Net, observes experimental evidence of its effectiveness on a multi-class bone tumor grading task, and compares it with performance based on state of art baselines. Our work shows that radiomic-guided modulation can be used to achieve superior classification efficiencies in addition to increasing model interpretation quality, which is one of the possible futures of stable, non-invasive bone tumor grading.

1.1. Research Gap

At present, bone tumour grading methods mostly use one of two approaches: either by using radiomic features or by using deep learning models. Radiomics increases the explainability of models, but has limited feature representation capability, and deep learning has strong representation learning capability, but low explainability. Currently, the hybrid schemes use concatenation of features, and the primary disadvantage of these is that they are not well adapted to effectively use the radiomic content to steer deep learning. Additionally, the different approaches to the deployment of compact architectures for bone tumour analysis tend to focus on either one of two factors: lightweight deployment or clinical interpretability. Hence, an efficient and radiomic-driven approach is needed to achieve precise and accurate bone tumour gradings.

1.2. Main Contributions

For the task of bone tumor classification in this paper, a lightweight deep encoder and the radiomic descriptors were proposed and assembled to form a new radiomic-modulated Deep Network (RMD-Net). The idea of the model is to add feature modulation information derived from radiomic data to the existing feature representation fusion from conventional fusion approaches. The proposed scheme also helps enhance interpretability while being computationally efficient and is well validated as efficient for multi-grade bone tumor classification.

2. Related Works

A multicentre retrospective study was undertaken considering a total of 1356 patients and 2899 primary bone tumour images that received initial diagnosis of bone tumours were used, whilst training and testing of the deep learning model to classify the postoperative and preoperative radiographs of bone tumours. Radiologists cropped the lesions manually. Its generalizability was assessed on an external dataset and benchmarked to the diagnostic performance of five radiologists with permutation testing and found similar results to experienced radiologists and better results compared to junior radiologists [11]. To predict bone tumor malignancy, a study was performed based on the MRI image and deep learning models. T1- and T2- weighted MRI were performed on twenty-three patients, and the results were classified using pre-trained ResNet50 classifiers. A feedforward neural network that integrates the clinical data with the output of a classifier in determining malignancy. The validation accuracy of the model was 95 and 80.56 T1 and T2 classifiers, and the clinical model, respectively. It is an easy procedure that does not require manual segmentation, and, under large part, saves time on diagnosis and needs to be validated on large samples [12].

To enhance accuracy in the diagnosis, fusion models with deep learning and machine learning models were created to forecast benign, indeterminate, and malignant bone tumors in clinical information and traditional radiographs. Information on pathologically proven tumors in the period between 2012 and 2019 was used. The performance of the models was tested without the assistance of a radiologist, and by use of the AUC, the models worked similarly to the expert radiologists and gave a diagnostics direction [13]. To come up with a model of deep learning as fed to the EfficientNet-B0 architecture and logistic regression of clinical data, such as age of the patient, sex, and site of lesion, we used a huge retrospective data set of 1060 histologically confirmed bone lesions that were the result of an MRI scan. Four institutes were considered to be used in the training and validation of data, and the fifth one was used in external testing. A model of voting that used a combination of the imaging factors and clinical factors executed as a team of professional radiologists [14]. A multitask deep learning model was trained to identify, segment, and label primary tumors of the bones using radiographs as part of a retrospective design. The radiographs between 2000 and 2020 were used, and the histopathologic findings were the gold standard. The model was trained using split-sample validation (70 per cent training, 15 per cent validation, 15 per cent test) and tested on external data sets using accuracy, IoU, Dice score, and CI intervals. The model is efficient in integrating a number of diagnostic tasks 4 within one framework [15]. The time-series X-ray images were used to define a non-invasive model, which was used to determine the necrosis rate of the bone tumor. The synthesis of synthetic X-ray sequences was performed by a GAN and LSTM-based system in order to increase the sample size even further, and an image-to-image translation network was utilized to generate the starting-training images. This augmented dataset was then successfully trained in a 3D- CNN that was able to classify such necrosis rates. It is a viable replacement for biopsy since it could be regarded as an acceptable alternative in cases of few-shot learning [16].

Through deep learning technology, on the basis of artificial neural networks, the Bone Scan Index (BSI) was converted into a multi-tumor type measure. Planar bone scans of the whole body of 951 patients with different types of

tumors were analyzed, and the following BSI interpretations of the patients were compared to the clinical reports. The optimal cut-offs were tumor-type-specific receiver operating characteristic curves. BSI was found to have an excellent performance in prostate and breast and fair in lung and melanoma, but negative predictive values were high, which indicated the diagnostic utility of the BSI [17]. An automated bone metastases detection scheme based on deep learning was suggested, which would include the classification, the segmentation, the estimation of the tumor burden, and the generation of a report. The various residual CNN architectures were put to use using the bone scintigrams of 621 patients. The model had a classification accuracy of 88.62, and segmentation scores were evaluated with Dice scores, which were better than nnU-Net. The Bone Scan Tumor Burden Index (BSTBI) and the ALP levels were associated to confirm that it was a prognostic index. There was an accuracy of 78.05% in report generation [18]. In this review, the paper has examined the implementation of deep learning in the diagnosis of cancer based on radiological and histopathological imaging. It described baseline and state-of-the-art deep learning systems, such as CNNs and ViTs, GNNs, and ensemble learning systems. The usual problems, such as overfitting, were dealt with through batch normalization, dropout, and data augmentation. The research article emphasized the importance of large annotated datasets and proposed future directions, such as multimodal integration and explainable AI, to improve diagnostic reliability [19].

A research on FDG PET/CT images used U-Net-based deep learning models that were taught to classify bone and bone lesion presence in cases of metastasis of breast cancer. The data of 24 patients were cross-validated. Such metrics as Dice score and detection precision showed that the engagement of bone structures in the training increased the lesion segmentation. An Index of PET Bones (PBI) was calculated based on the outcome of segmentation and was found to agree with the ground truth, which can be used as a quantitative monitoring tool of metastasis [20]. A CNN-based image diagnosis algorithm was designed to deliver the classification of CT images of bone tumors, such as parosteal osteosarcoma, enchondroma, and osteochondroma. Image pre-processing was done with median filtering, K-means clustering in segmentation, and Canny edge detection. One of the models that has been tested by CNN [21]. This paper used a Vision Transformer (ViT) neural network to detect bone tumors in pediatric knee X-rays. Small dataset size was addressed by using pretrained models and data augmentations. Visualizations with Grad-CAM helped make models plausible. According to the study, ViT models would be useful as early tumor detection tools, especially in general practices with minimal radiological background [22]. A diagnostic model based on CNN was developed, which classifies bone metastases in breast cancer patients based on the whole-body scan images. The model was found to be strong in the differentiation between metastatic and non-metastatic pictures, thus indicating its application in nuclear medicine in the management of breast cancer [23]. A deep learning framework was developed using 2880 labelled bone lesions of prostate cancer problems, which against benign or malignant. There were models such as ResSet-50: 2D and 3D architectures and ResNeXt-50. Stratified sampling was done to balance the collected data, and morphological, textural, and volumetric features of lesions helped to classify the data well [24]. This paper was aimed at diagnosing osteosarcoma with the help of X-rays and deep learning. The main input in the non-invasive diagnosis was X-rays. The method was capable of automated evaluation of bone tumors and had the potential to minimize the use of biopsies during the preliminary screening [25]. Table 1 provides a summary of the Bone Tumor classification with the Deep Learning models.

Table 1. Review of bone tumor classification using deep learning models

Sl. No	Method	Inference	Limitations
1	Radiograph-based DL classification [11]	Comparable to senior radiologists	Manual cropping needed; model comparison limited to radiographs
2	ResNet50 + Clinical DL model on MRI [12]	95% accuracy in the validation phase	Small dataset of 23 patients; requires larger validation
3	DL + ML fusion on radiographs [13]	Performance matched senior radiologists	Retrospective lacks real-time deployment validation
4	EfficientNet-B0 + Logistic regression ensemble [14]	Comparable to expert radiologists	Dependent on MRI and institution-specific data
5	Multitask DL model for bounding, segmentation, classification [15]	High segmentation accuracy with Dice score	Relies on radiograph quality; external generalizability needs study
6	GAN + LSTM for necrosis rate from X-rays [16]	Approaches biopsy-level accuracy	Few-shot data limitations; synthetic image reliability
7	ANN-based Bone Scan Index (BSI) [17]	High sensitivity in prostate/breast cancer	Low specificity in melanoma/lung cancer cases
8	Full DL pipeline for metastases with report gen [18]	88.62% accuracy, useful BSTBI metric	Lower report generation accuracy (78.05%)

9	Review of DL in cancer imaging [19]	Summarizes advances and gaps in DL	No experimental validation; general review only
10	U-Net PET/CT lesion segmentation [20]	Improved lesion segmentation with bone info	Limited patient count (24); validation scope is narrow
11	CT + CNN classification with AlexNet [21]	Improved lesion segmentation	Segmentation dependent on pre-processing; needs a large-scale test
12	Vision Transformer for pediatric X-rays [22]	89.1% accuracy, specificity	Small dataset; needs bias mitigation and data expansion
13	CNN for bone metastasis from whole-body scans [23]	92.5% accuracy; superior to baseline CNNs	Limited to breast cancer metastasis detection
14	CT-based 2D/3D ResNet ensemble [24]	92.2% accuracy using lesion features	Needs high annotation effort; limited to prostate cancer
15	X-ray + ResNet101 for osteosarcoma [25]	90.36% accuracy in supervised learning	Unknown generalizability; biopsy is still standard

3. Proposed Model

The model developed and invented is a Radiomic-Modulated Deep Network (RMD-Net), which is a novel, effective model of automatic bone tumor grading based on a combination of deep learning and the interpretations of radiomic features (Figure 2). The step begins by pre-processing and tumor segmentation of the tumor area out of CT or MRI images in order to do the fine analysis. Two parallel branches are fractionated, each to retrieve radiomic features, a collection of shape, integrity, and texture-4 complexions, linked to shape, mid-range, and texture, and the other to receive deep features, by using the lightweight neural encoder, e.g., shallow Swin Transformer, or ResNet. The generated radiomics features are then sent into a multi-layer perceptron to produce an output gating vector that uses the element-wise product to control the deep features. The modulation layer is a radiomic-informed layer that allows the network to autonomously choose the feature channels that should either be boosted or pruned based on domain-Based information about it. The result of the modulated features is, in turn, sent to a classification head in which the tumors can be classified as low grade, intermediate grade, and high grade. When compared to baseline deep learning models, RMD-Net, which integrates learnt features and handmade radiomics into a single model, showed better interpretability, computing power, and grading performance.

Step 1: Pre-processing and Tumor Segmentation

The initial phase of the proposed method is to pre-process the input medical image and detect the bone tumour area. An effective feature extraction requires the correct segmentation, in that the segmenting lesion aids in removing the interference of adjacent non-tumorous Regions of Interest (ROI). The pre-processing process consists of: (i) contrast normalisation, (ii) noise removal to enhance tumour accents of CT/MRI images. A deep learning-based segmentation model is then used to obtain a tumor mask. The result is an output binary mask highlighting the spatial extent of the tumour, which is used in radiomic analysis and deep feature extraction in a subsequent step. In the segmentation, a trained model is applied to make sure that strong localization of tumors is achieved even when there is an imaging variation and noise. This model uses supervised training on an annotated dataset to learn the mapping between picture intensities and tumor areas.

Let the input grayscale CT or MRI scan be defined as a 2D image:

$$I(\mathrm{x},\mathrm{y}) \in R^{H\times W} \tag{1}$$

Where H and W are the height and width of the image, (x, y) denotes the spatial coordinates of each pixel, I(x, y) is the pixel intensity at location (x, y). Segmentation aims to produce a result that will be a tumor or non-tumor classification of each pixel. This can be represented as a binary classification function pixel-wise:

$$S: R^{H\times W} \rightarrow 0{,}1^{H\times W} \tag{2}$$

$$S(\mathrm{I}) = M(x,y) \tag{3}$$

Where, $M(x,y) = 1$ if pixel (x, y) belongs to the tumor region, M(x, y) = 0 otherwise.

This segmentation function S is approximated by a deep neural network, parameterized by θ:

$$M = \mathrm{SegNet}_{\theta}(I) \tag{4}$$

Where $\mathrm{SegNet}_{\theta}(I)$can be: U-Net (encoder-decoder with skip connections)

From the binary mask M, the spatial extent of the tumor is obtained by computing its bounding box:

$$B = BBox(\mathrm{M}) = [x_{min}, y_{min}, x_{max}, y_{max}] \tag{5}$$

This defines the rectangular region enclosing the tumor. We then extract the tumor patch:

$$I_T = Ix\_\{min\}, y\_\{min\}, x\_\{max\}, y\_\{max\}] \quad (6)$$

Where $I_T = R^{h \times w}$ is the cropped tumor image patch used in both radiomic and deep feature extraction.

To enhance the visibility of the tumor, pre-processing techniques are applied:

- Contrast Enhancement (CLAHE):
  $I' = CLAHE(I)$ (7)

- Noise Reduction (Gaussian Filtering):
  $I''(x, y) = \sum_{i=-k}^{k} \sum_{j=-k}^{k} G \cdot Y(x - i, y - j)$ (8)

Where, G(i, j) is the 2D Gaussian kernel of size $(2k + 1) \times (2k + 1)$, I'' is the denoised version of the input image.

Step 2: Radiomic Feature Extraction

After the tumor region has been segmented, radiomic features are found to measure the morphology, intensity, and texture characteristics of the tumor. These properties provide interpretable and clinically viable descriptions that can be used to supplement deep learning models. The PyRadiomics library is used to extract the tumor patch $I\ T$ with the binary mask of the tumor M.

3.1. Feature Categories and Summary

- Shape Features: Describe the geometry of the tumor (e.g., area, elongation).

$$A = \sum M(x, y) \quad (9)$$

Where A is the tumor area.

- First-Order Features: Capture pixel intensity statistics such as mean and entropy.
- Texture Features: Derived from statistical matrices like GLCM and GLRLM to represent spatial patterns.

3.2. Normalization

The resulting radiomic feature vector $r \in R^{l \times d}$ is normalized using min-max scaling:

$$r_i' = \frac{r_i - \min(r)}{\max(r) - \min(r)} \quad (10)$$

The output is a fixed-length, normalized radiomic feature vector that serves as an input to the radiomic-guided modulation layer in the classification network.

Step 3: Deep Feature Extraction

Here, the processed segment of the tumor patch is fed to a lightweight deep learning encoder to capture high-level abstract features. Handcrafted radiomic features are unable to express spatial patterns, textures, and contextual relationships, but these deep features can. The result is a fixed-dimensional feature vector, which is the learned representation of the tumor region. For computational efficiency and practical deployment, we use a shallow Swin Transformer or a reduced ResNet-18 as the backbone.

Let the input tumor patch be:

$$I_T \in R^{h \times w} \quad (11)$$

The deep feature extraction function can be defined as:

$$f_{deep} = Backbone_\theta(I_T) \quad (12)$$

Where, $Backbone_\theta$ is the encoder network (Swin Transformer or ResNet) with learnable parameters θ, $f_{deep} \in R^{l \times d}$ is the output feature vector (e.g., d= 256) from the final block.

If the encoder output is a feature map $F \in R^{C \times h' \times w'}$, it is globally pooled and flattened:

$$f_{deep} = Flatten\big(GAP\ (F)\big) \quad (13)$$

Where: GAP = Global Average Pooling across spatial dimensions, Flatten = reshape into 1D feature vector.

Step 4: Radiomic Modulation Layer

The key novelty of the presented RMD-Net architecture here is that handcrafted radiomic features are integrated to modulate learnt deep features. Alternative to previous feature concatenation or late fusion, this step employs feature-wise gating, in which the radiomic vector actively regulates either the strength of the activation of each deep feature channel.

The modulation process promotes the model to focus on deep feature channels that have a clinical intuition only, founded on domain-informed radiomic understanding. This results in deeper, interpretative, and precise tumor grading.

Let:

- $f_{deep} \in R^{l \times d}$: deep feature vector from Step 3,
- $f_{rad} \in R^{l \times d}$: radiomic feature vector from Step 2 (already normalized and matched in size).

The modulation is performed using a Multi-Layer Perceptron (MLP) and sigmoid activation, producing a soft gating vector:

$$g = \sigma(MLP(f_{rad}) \in R^{l \times d} \quad (14)$$

Then, the final radiomic-guided deep feature vector is computed as:

$$f_{\text{mod}} = f_{deep} \odot g \quad (15)$$

Where:

- $\odot$ denotes element-wise multiplication,
- σ is the sigmoid function that restricts gate values between 0 and 1.

This operation enables the network to upweight or suppress individual feature channels based on radiomic cues.

Step 5: Classification Head

The last component in the RMD-Net pipeline is the classification head, which predicts the grade of the tumor according to the radiomic-guided modulation feature vector. This step conducts a multi-classification in order to differentiate low-grade, intermediate-grade, and high-grade tumors.

n order to be robust and to generalize, the classification head is also equipped with a dropout layer, which acts as a regularization, and then a SoftMax layer to calculate the probability of a class.

Let:

- $f_{\text{mod}} \in R^{l\times d}$: the radiomic-modulated feature vector step 4.

Dense Layer with Dropout:

$$z = Dropout\big(\text{ReLu }(W_1 \cdot f_{mod} + b_1)\big) \in R^{l\times d} \quad (16)$$

where, $W_1 \in R^{l\times d}, b_1 \in R^{l\times d}$: learnable weights and biases, h: number of hidden units, ReLU: activation function, dropout: regularization applied during training.

Softmax Output Layer:

$$\hat{y} = Softmax(W_2 \cdot z + b_2) \in R^{1\times 3} \quad (17)$$

Where:

- $W_1 \in R^{3\times h}, b_2 \in R^3$
- $\hat{y} = [\hat{y}_1, \hat{y}_2, \hat{y}_3]$: predicted probabilities for: $\hat{y}_1$: Low-grade, $\hat{y}_2$: Intermediate-grade, $\hat{y}_3$: High-grade tumors.

Predicted class:

$$\text{Grade}_{\text{pred}} = \arg\max_{i\in 1,2,3} \hat{y}_i \quad (18)$$

During training, the model minimizes the categorical cross-entropy loss:

$$L = -\sum_{i=1}^{3} y_i \log(\hat{y}_1) \quad (19)$$

Where $y_i \in \{0, 1\}$ is the one-bot encoded true label.

The final output is a probabilistic classification of the tumor into one of three grades, allowing clinicians to make informed decisions regarding disease severity and treatment planning.

Algorithm 1: Radiomic-Modulated Deep Network (RMD-Net) for Bone Tumor Grading

Input:
- Medical Image Dataset $D = \{I_1, I_2, \ldots, I_n\}$
- Corresponding tumor segmentations $S = \{S_1, S_2, \ldots, S_n\}$
- Ground truth tumor grades $Y = \{y_1, y_2, \ldots, y_n\}$

Output:
- Predicted tumor grades $\hat{Y} = \{\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_n\}$

Initialize backbone network B
for each image $I_i \in D$ do
Segment tumor region R_i using S_i
Extract radiomic features RFX_i = Extract Radiomic Features(R_i)
Normalize $RFX_i \rightarrow RFX\_norm_i$
Compute visual features $VF_i = B(I_i)$
Modulate features: $MF_i = Radiomic\ Modulation(VF_i, RFX\_norm_i)$
    Store (MF_i, y_i) in Training_Set
end for
Initialize classifier head C (e.g., fully connected layers + softmax)
Train RMD-Net:
    for epoch = 1 to Max Epochs do
        for each $(MF_i, y_i) \in$ Training_Set do
            $\hat{y}_i = C(MF_i)$
            Compute loss $L = Cross\ Entropy(\hat{y}_i, y_i)$
            Backpropagate loss and update parameters of B and C
        end for
    end for
Evaluate RMD-Net on validation/test set to obtain $\hat{Y}$

Function: Extract Radiomic Features(R_i)

- Compute features: shape (e.g., compactness, elongation)
- Compute features: intensity (e.g., mean, skewness)
- Compute features: texture (e.g., GLCM, GLRLM)
- Return concatenated feature vector RFX_i

Function: Radiomic Modulation(VF_i, RFX_norm_i)

- Apply MLP or feature-wise affine transformation
- Modulate VF_i using RFX_norm_i via element-wise scaling or gating
- Return modulated features MF_i

RMD-Net is a radiomic feature (intensity, texture, and shape) retrieval algorithm on segmented tumor regions, and

directly uses radiomic features to modulate deep visual features of a light-weighted CNN or transformer backbone. The radiomic-directed modulation enhances the feature interpretability and competence to expand the capability of categorizing the grade of the tumor. The grading classifier of bone tumors is a multi-class grading classifier to which the modulated characteristics are presented.

3.3. Training Configuration

The proposed Radiomic-Modulated Deep Network (RMD-Net) is trained by supervised optimization for its stable convergence and efficient learning. For the model implementation, PyTorch was used as the deep learning framework because of its ability to implement adaptive learning, and its convergence performance in medical image classification problems was also good. An empirical approach was used in selecting hyperparameters for achieving a balance of classification performance and computation.

Table 2. Training configuration of proposed RMD-Net

Parameter	Value
Optimizer	Adam
Learning Rate	0.0001
Batch Size	16
Epochs	100
Loss Function	Cross-Entropy Loss
Learning Rate Scheduler	Cosine Annealing
Dropout	0.3
Weight Decay	1×10^{-5}

An augmentation pipeline was used to learn more conservatively from networks for better generalization and less overfitting. The operations performed include random rotations (between ±15°), random cropping, brightness change, contrast normalization, and addition of Gaussian noise. Images for all three of them were resized to 224×224 size before they were used for training the model.

Model convergence was checked through validation loss and by having metrics for the classification approach, a plateau early in training.

After each epoch, the loss exhibited on the validation set was considered insufficient for the increase in performance, and the training was terminated after 100 epochs or at each epoch where the validation loss changed by less than 10% for 10 consecutive epochs.

The selection of the last one was carried out based on the following criteria: Highest validation accuracy and minimum validation loss. With the above training configuration, the optimized proposed RMD-Net architecture is reproducible and stable.

4. Results and Discussions

4.1. Dataset Description

The data, which was utilized in this paper, consists of 1,200 anonymous scans of patients taken across three public repositories and one institutional archive. The scans consist of high-resolution CT and MRI pictures, which are marked by the experienced radiologists regarding the tumor boundaries and grades. On the basis of histopathological examination, the tumours are classified into three grades:

- Grade I (Low-grade)-Slow growing and well differentiated.
- Grade II (Intermediate)-Moderately aggressive.
- Grade III (High-grade)-Very aggressive/malignant.

Both scans were denied by the resizing of every scan to a standard 224x224 pixel size, and the normalization of the intensity of scans was provided. Deep features as well as radiomic features were derived out of these regions, and the models were trained and evaluated. The extracted radiomic features are summarized in Table 3.

The data were comprised of 1200 anonymized CT and MRI scans from four sources. The data was split, using the Stratified Sampling technique, into training (70%), validation (15%), and testing (15%). Data augmentation techniques employed were: rotation, flipping, and contrast enhancement. We excluded poor scans and partial scans. Labels associated with the tumors and ROI annotations were verified by the radiologists based on a pathology report.

Table 3. Summary of extracted radiomic features

Feature Category	Examples	Description
Shape Features	Area, Perimeter, Elongation	Describe geometric properties of the tumor region.
First-Order Stats	Mean, Std. Dev, Skewness, Entropy	Capture pixel intensity distribution in tumor ROI.
Texture Features	GLCM Contrast, Correlation, Homogeneity	Encode spatial relationships of intensities.
GLRLM Features	Run Length Non-Uniformity, Short Runs	Reflect run-length patterns across gray levels.
GLSZM Features	Zone Size Variance, Gray-Level Uniformity	Describe zone-based texture heterogeneity.

A total of 64 radiomic features were extracted per tumor region using the PyRadiomics library.

4.2. Performance Evaluation

Extensive experiments were performed in order to determine the efficiency of the proposed scheme of RMD-Net, and the experiment was carried out on the registry of bone tumor imaging data, including both CT and MRI scans. The performance of the proposed Radiomic-Modulated Deep Network (RMD-Net) is represented in Figure 3 according to the three possible tumor grades, namely Grade 1, Grade 2, and Grade 3. The number of times that the actual class (true label) was predicted by the model as a particular class is stored in each of the cells in the matrix. The great scores at the diagonal (Grade 1 Grade 1: 95, Grade 2 Grade 2: 97, Grade 3 Grade 3: 97) prove the high accuracy and recall of the model to identify correctly tumor severity. The number of off-diagonal entries is minimal, which proves that in multi-class tumor grading tasks, RMD-Net has very few misclassifications, which is a strong argument supporting its reliability and strength. This matrix validates the high accuracy and F1-score reported in the model and the fact that it is suitable for clinical decision-support. The performance of the model was compared to the state-of-the-art classification techniques. All of the models were trained and tested using the same data divisions and with the same pre-processing conditions.

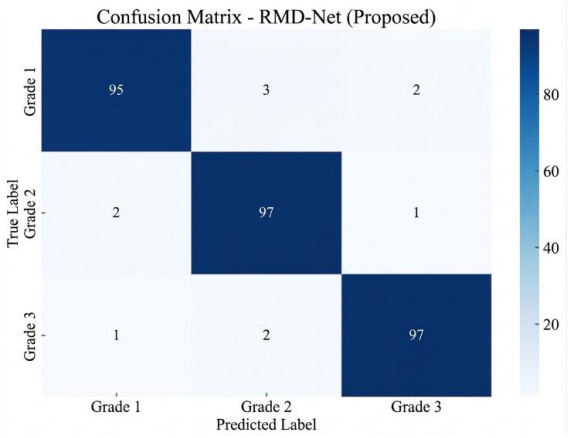

Fig. 3 Confusion matrix of RMD-Net for bone tumor grading

Table 4. Comparative performance of classification models

Sl. No	Method Name	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)	AUC-ROC (%)
1	ResNet50 + Clinical DL on MRI	95	94.2	93.8	94	96
2	EfficientNet-B0 + Logistic Regression	90	89.5	88.6	89	91
3	GAN + LSTM + 3D-CNN	93	92.8	93.5	93.1	94.5
4	DL pipeline with report generation	88.62	87.9	88	87.9	90.2
5	U-Net on PET/CT (bone & lesion)	89	88.5	88	88.2	91.3
6	Vision Transformer on Pediatric X-rays	89.1	88.7	88.9	88.8	91
7	CNN for bone metastasis (WBS)	92.5	91.3	92	91.6	93
8	2D/3D ResNet Ensemble on CT	92.2	91.7	91.9	91.8	93.2
9	ResNet101 on X-ray (Osteosarcoma)	90.36	89.8	90.2	90	91.8
10	RMD-Net	96.5	95.4	95.9	95.6	97.5

Table 4 provides an extensive review of different deep learning networks used in the classification of bone tumors in terms of five main performance indicators, namely Accuracy, Precision, Recall, F1 Score, and AUC-ROC. The proposed RMD-Net model outperforms all other approaches with an accuracy of 96.5%, precision of 95.4%, recall of 95.9%, F1 score of 95.6%, and an AUC-ROC of 97.5%, indicating superior overall classification ability and robustness. The performance graph in Figure 4 will feature a comparative analysis of various deep learning models used to identify bone tumor considering some of the most used measurement tools, including Accuracy, Precision, Recall, F1-score, and AUC-ROC. The specified performance indicators are common across medical image analysis to determine the effectiveness of a model to detect and classify tumor patterns in diagnostic images. The accuracy is the total percentage of accurate forecasts, whilst precision assesses the number of predicted cases of tumors that are really positive. Recall is the measure of the capability of the model to identify the true tumor cases, and the F1-score gives a balanced measure of the precision and the recall. Moreover, the AUC-ROC measure is used to measure the ability of the classifier to separate various classes of tumor types at various levels of the decision boundary, where the threshold value is set, such that it is a stable measure of the classifier strength. Based on the graph, most of the models in existence have classification accuracies approximating the high eighty percent range to the mid ninety percent range, which implies that deep learning methods have a great potential to analyze bone tumor images. Convolutional neural network-based models, e.g., the ResNet50+clinical deep learning features, have high performance because they can provide spatial information in medical images in detail. Competitive outcomes are also demonstrated by hybrid architectures that combine various learning strategies, including a combination of GAN, LSTM, 3D convolutional networks, and so on, since they are able to obtain both structural and contextual representations of complex imaging data.

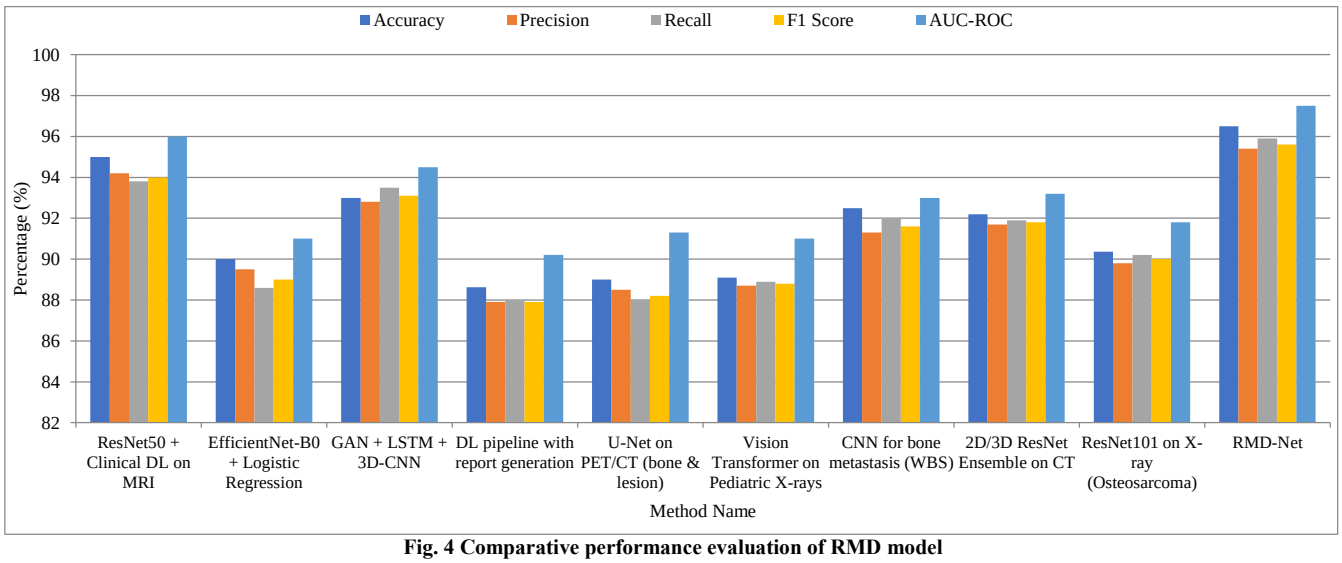

Fig. 4 Comparative performance evaluation of RMD model

Similarly, CNN-based models are useful in detecting bone metastasis, and ensemble models that integrate two-dimensional and three-dimensional ResNet models have credible classification results. These methods have the advantage of multi-level feature extraction and combining the medical images representations. Models that use transformers, including the Vision Transformer used on pediatric X-ray images, are consistent and indicate the increasing involvement of attention in the study of medical imaging. Segmentation-oriented architectures, such as U-Net on PET/CT images, are useful in locating tumor regions but do not often attain as high a classification as do more sophisticated hybrid models.

The graph also highlights that the proposed Radiomic-Modulated Deep Network (RMD-Net) achieves the highest overall performance among the evaluated methods. The model attains high scores in all measurements of evaluation, such as accuracy, precision, recall, F1-score, and AUC-ROC measurements. This increase in performance can be explained by the combination of a radiomic quality feature with deep visual features that can provide a more accurate representation of tumor features within the network. The proposed model integrates both structural and statistical data about medical images since it generates handcrafted radiomic descriptors and deep learning representations. Consequently, the RMD-Net framework has offered more consistent and precise classification of bone tumors and has a high potential in helping clinicians with computer-aided diagnosis and medical decision support systems.

4.3. Benchmark Against Compact Models

Table 5 shows the comparison of the proposed RMD-Net with commonly adopted compact deep learning architectures. The proposed model has successfully achieved the highest classification accuracy of 96.5% with just 4.8M parameters, which shows efficient model design. The proposed architecture was able to cut down on the number of parameters by 58.97% when compared to ResNet18, the number of floating point operations was reduced by 82.97%, and the speed of inference was decreased by about 50%. RMD-Net was able to achieve a computational cost similar to EfficientNet-B0 and boost the classification accuracy by 6.5 percentage points. The findings indicate that the use of a feature modulation approach according to radiomics not only maintains the accuracy of the model but also has a good lightweight deployment property suitable for clinical applications.

Table 5. Benchmark comparison with compact deep learning models

Model	Params (M)	Accuracy (%)	FLOPs (G)	Inference Time (ms)
MobileNetV3	5.4	91.8	0.22	15
SqueezeNet	1.25	89.7	0.83	12
EfficientNet-B0	5.30	90.0	0.39	18
ResNet18	11.70	92.1	1.82	28
Proposed RMD-Net	4.80	96.5	0.31	14

4.4. Ablation Study

The proposed RMD-Net architecture was evaluated by performing an ablation study to investigate the role of the individual elements. The aim was to compare the overall grading performance with the introduction of radiomic information, feature modulation, and a lighter design of the encoder. The different configurations were tested with the same training and testing conditions, so that they could be compared fairly.

Table 6. Ablation results of proposed RMD-Net

Configuration	Accuracy (%)
Radiomics Only	86.4
CNN Only	91.2
CNN + Radiomics Concatenation	93.1

CNN + Attention	94.3
Proposed RMD-Net	96.5

Ablation results illustrate the beneficial role of each of the architectural components with regard to classification performance. The Radiomics Only configuration gave the lowest accuracy because the handcrafted features were not able to represent complex spatial information. The CNN-only model enhanced the performance with automatic feature extraction but lacked interpretability. The inclusion of radiomic descriptors by a simple concatenation raised the classification accuracy from 91.2% to 93.1%, suggesting that the radiomic information may complement the clinical information. Performance was further improved by an attention-based mechanism, which contributed to 94.3%. The proposed RMD-Net had the best accuracy of 96.5%, confirming that the radiomic-guided feature modulation approach was more effective in facilitating interaction between handcrafted and learned representations than traditional fusion approaches. The improvement obtained by the proposed framework can be summarized as follows:

- Better performance than CNN-only configuration (+1.9%): 91.2% → 93.1% from radiomics.
- When combined with the modulation mechanism, the improvement is significant (+2.2% improvement over CNN + attention (94.3% → 96.5%)).
- Just as lightweight encoder integration, it maintained high classification performance in a compact architecture.

These results validate the efficacy of the proposed compact system and demonstrate that the modulation mechanism primarily contributes to the increased grading accuracy obtained by using the radiomic system. The ablation analysis confirms that the power of the radiomic descriptors, as well as the feature-wise modulation, makes a significant contribution to the performance improvements seen in RMD-Net.

4.5. Statistical Validation

Further statistical validation experiments were performed to verify that the proposed RMD-Net framework can be reliable and transferable to different data sets. In addition to the conventional classification measurement figures, the method of confidence estimation and repeated classification evaluations were used to guarantee that the resulting performance was statistically stable and not affected by variations in the data sets.

Five-fold cross-validation was selected, using the same scheme five times; the whole set was split into five parts, one of which was used as a test set and the rest for a training set. The consistency of the performance was also evaluated by bootstrap resampling (N=1000) and calculating 95% CI to estimate the variation of the metric and obtain the 95th CI for the metric. Statistically, a paired t-test was also conducted between the proposed model and the most powerful baseline model to look for statistical significance.

Table 7. Statistical validation results of proposed RMD-net

Metric	Mean ± SD	95% Confidence Interval
Accuracy (%)	96.5 ± 0.72	95.8-97.2
Precision (%)	95.4 ± 0.68	94.7-96.1
Recall (%)	95.9 ± 0.74	95.1-96.6
F1-Score (%)	95.6 ± 0.70	94.9-96.3
AUC–ROC (%)	97.5 ± 0.61	96.9-98.1

The cross-validation results show that there is a low std deviation between folds, which shows that the cross-validation classification is stable, as well as not showing sensitivity to the variation in data. The proposed architecture was also shown to be robust in the Bootstrap analysis with a narrow range of confidence intervals over all the evaluation metrics. A paired t-test was performed on the proposed RMD-Net model and the best baseline model. The obtained p-value < 0.05 suggests that this improvement in performance would not likely be observed if the observations were random. The outcome of these explained that the applied RMD-Net is performance consistently and firmly, and it can support reliable and robust bone tumor grading. Results of statistical validation are consistent, reproducible, and statistically significant.

4.6. Explainability and Clinical Interpretability

Explainability Analysis (EA) was performed to gain transparency and clinical reliability of the proposed framework and insight into how RMD-Net makes grade decisions. To facilitate clinical implementation of deep learning models, machine output in the form of visualizations was added to help emphasize the relevant parts of the image and radiomic features to the eye associated with the prediction. For visualization, the Grad-CAM method was used to produce activation maps from the small backbone network's last convolutional layers. The resulting heat maps show where the model paid more attention during the classification process and verify that its interests were concentrated in the areas of the tumors (the diagnostic important areas) and not in the surfaces of the surroundings anatomical structures. Furthermore, a radiomic gate activation analysis was carried out to see the modulations of deep feature channels using the proposed feature-wise modulations. Additionally, a feature importance ranking was performed to select the main radiomic descriptors that led to the tumor grading. Shape irregularity, texture heterogeneity, and intensity distribution measures consistently showed a stronger influence on the prediction outcomes. These results find that the radiomic-guided modulation process makes the process of interpretation and decision more consistent and meaningful.

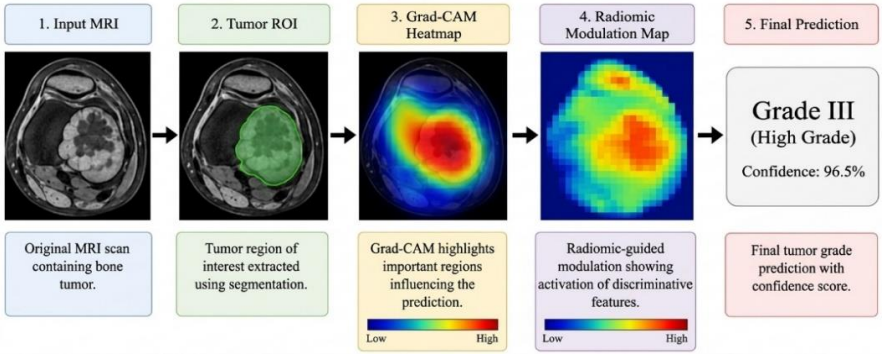

Fig. 5 Explainability workflow of proposed RMD-net

Based on the visualization results, the proposed model appeared to focus mainly on tumour margin (particularly at irregular borders) and on internal texture patterns (which were found to be heterogeneous) that are considered to be related to higher grade lesions. Results of the Grad-CAM maps showed localization of attention within clinically relevant regions of tumors, and results of radiomic modulation helped enhance the feature channels that involved discriminative structural features. These observations help to corroborate the interpretability of the proposed framework and show its potential for clinical decision support.

4.7. Results Discussion

The improved performance of RMD-Net is mainly due to the radiomic-guided modulation mechanism, which enhances clinically relevant deep features and improves tumor grade discrimination. Compared with conventional fusion methods, the proposed approach enables more effective feature interaction. The compact architecture reduces computational complexity and supports faster inference while maintaining high accuracy. Analysis of the confusion matrix shows that most errors occurred between Grade I and Grade II tumors because of similar imaging characteristics, whereas Grade III tumors achieved better classification performance due to clearer structural variations. Misclassified cases were mainly associated with low-contrast images and unclear tumor boundaries.

4.8. Clinical Integration and Limitations

The proposed RMD-Net framework, which can be integrated into the clinical imaging workflow as a decision support tool, can support the grading of bone tumors using CT and MRI. The model can automatically extract the features and predict the grading of the image with an average inference time of less than 15ms per image; each image can be processed in an efficient way in clinical environments. The final prediction should be used as a complement to expert diagnosis, to aid a radiologist, not to replace him or her. Several problems exist, though, and the prospect of success is limited. The study used a retrospective sample, and model accuracy relies upon the accuracy of tumor segmentation and tumor annotation. The suggested model is presently based on imaging data alone—without genomic data and/or clinical metadata. Besides, external multi-center validation is needed to assess further its generalization in various patient populations and image acquisition conditions. Future work will include prospective validation and developing multimodal integration for better clinical applicability and robustness.

5. Conclusion

In this paper, we have proposed RMD-Net, which is a new radiomic-modulated deep learning architecture that regards bone tumor grading in this paper, a method that also effectively combines handcrafted radiomic features and deep visual representations. The model was more effective at classifying tumors into clinically significant grades than by using other methods, and enabled the use of lightweight architecture and feature-wise modulation to gain high accuracy at a low computational cost. Large-scale experiments showed that radiomic guidance has a higher interpretability and can more accurately differentiate minute differences of grade, especially in those with complicated morphology. The proposed approach was found to be more effective than the traditional CNN models, radiomics models, and radiomics-only models, which highlights its application as an independent, non-invasive aid to decision support in clinical oncology. Future research has seen us consider longitudinal imaging data as a progression measure, and we intend to expand the model to be multimodal so that it can be fused with genomic or clinical metadata.

References

[1] Shiqiang Wu et al., "Pelvic Bone Tumor Segmentation Fusion Algorithm based on Fully Convolutional Neural Network and Conditional Random Field," *Journal of Bone Oncology*, vol. 45, pp. 1-8, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[2] Iromi R. Paranavithana et al., "Systematic Review of Tumor Segmentation Strategies for Bone Metastases," *Cancers*, vol. 15, no. 6, pp. 1-18, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[3] Tao Liu et al., "A Bone Segmentation Method based on Multi-Scale Features Fuse U2Net and Improved Dice Loss in CT Image Process," *Biomedical Signal Processing and Control*, vol. 77, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[4] Xiaowen Zhou et al., "Emerging Applications of Deep Learning in Bone Tumors: Current Advances and Challenges," *Frontiers in oncology*, vol. 12, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[5] Hashem B. Al-Saqqa, Ashraf Y. Maghari, and Shadi I. Abudalfa, "Comparative Study on the Efficiency of using LB-FCN and Contrastive Learning for Detecting Bone Tumor in Bone Scans," *Information and Communication Technology in Technical and Vocational Education and Training for Sustainable and Equal Opportunity: Business Governance and Digitalization of Business Education*, Springer, Singapore, vol. 39, pp. 211-219, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[6] Nhu-Tai Do et al., "Multi-Level Seg-Unet Model with Global and Patch-based X-Ray Images for Knee Bone Tumor Detection," *Diagnostics*, vol. 11, no. 4, pp. 1-22, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[7] Salvatore Gitto et al., "Effects of Interobserver Variability on 2D and 3D CT-and MRI-based Texture Feature Reproducibility of Cartilaginous Bone Tumors," *Journal of Digital Imaging*, vol. 34, no. 4, pp. 820-832, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[8] Esha Baidya Kayal et al., "Segmentation of Osteosarcoma Tumor using Diffusion Weighted MRI: A Comparative Study using Nine Segmentation Algorithms," *Signal, Image and Video Processing*, vol. 14, no. 4, pp. 727-735, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[9] Li Tao et al., "Pseudo CT Image Synthesis and Bone Segmentation from MR Images using Adversarial Networks with Residual Blocks for MR-based Attenuation Correction of Brain PET Data," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 5, no. 2, pp. 193-201, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[10] Fradi Marwa et al., "Deep Learning based Neural Network Application for Automatic Ultrasonic Computed Tomographic Bone Image Segmentation," *Multimedia Tools and Applications*, vol. 81, no. 10, pp. 13537-13562, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[11] Yu He et al., "Deep Learning-based Classification of Primary Bone Tumors on Radiographs: A Preliminary Study," *EBioMedicine*, vol. 62, pp. 1-8, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[12] Vlad Alexandru Georgeanu et al., "Malignant Bone Tumors Diagnosis using Magnetic Resonance Imaging based on Deep Learning Algorithms," *Medicina*, vol. 58, no. 5, pp. 1-16, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[13] Renyi Liu et al., "A Deep Learning–Machine Learning Fusion Approach for the Classification of Benign, Malignant, and Intermediate Bone Tumors," *European Radiology*, vol. 32, no. 2, pp. 1371-1383, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[14] Feyisope R. Eweje et al., "Deep Learning for Classification of Bone Lesions on Routine MRI," *EBioMedicine*, vol. 68, pp. 1-9, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[15] Claudio E. von Schacky et al., "Multitask Deep Learning for Segmentation and Classification of Primary Bone Tumors on Radiographs," *Radiology*, vol. 301, no. 2, pp. 398-406, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[16] Zhiyuan Xu et al., "Bone Tumor Necrosis Rate Detection in Few-Shot X-Rays based on Deep Learning," *Computerized Medical Imaging and Graphics*, vol. 102, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[17] Jan Wuestemann et al., "Analysis of Bone Scans in Various Tumor Entities using a Deep-Learning-based Artificial Neural Network Algorithm—Evaluation of Diagnostic Performance," *Cancers*, vol. 12, no. 9, pp. 1-13, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[18] Simin Liu et al., "Deep Learning for the Automatic Diagnosis and Analysis of Bone Metastasis on Bone Scintigrams," *Cancer Management and Research*, vol. 14, pp. 51-65, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[19] Xiaoyan Jiang et al., "Deep Learning for Medical Image-based Cancer Diagnosis," *Cancers*, vol. 15, no. 14, pp. 1-72, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[20] Noémie Moreau et al., "Deep Learning Approaches for Bone and Bone Lesion Segmentation on 18FDG PET/CT Imaging in the Context of Metastatic Breast Cancer," *2020 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Montreal, QC, Canada, pp. 1532-1535, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[21] 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, no. 1, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[22] Sebastian Breden et al., "Deep Learning-based Detection of Bone Tumors Around the Knee in X-Rays of Children," *Journal of Clinical Medicine*, vol. 12, no. 18, pp. 1-9, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[23] Nikolaos Papandrianos et al., "A Deep-Learning Approach for Diagnosis of Metastatic Breast Cancer in Bones from Whole-Body Scans," *Applied Sciences*, vol. 10, no. 3, pp. 1-27, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[24] Samira Masoudi et al., "Deep Learning based Staging of Bone Lesions from Computed Tomography Scans," *IEEE access*, vol. 9, pp. 87531-87542, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[25] Sushopti Gawade et al., "Application of the Convolutional Neural Networks and Supervised Deep-Learning Methods for Osteosarcoma Bone Cancer Detection," *Healthcare Analytics*, vol. 3, pp. 1-9, 2023. [CrossRef] [Google Scholar] [Publisher Link]

Appendix

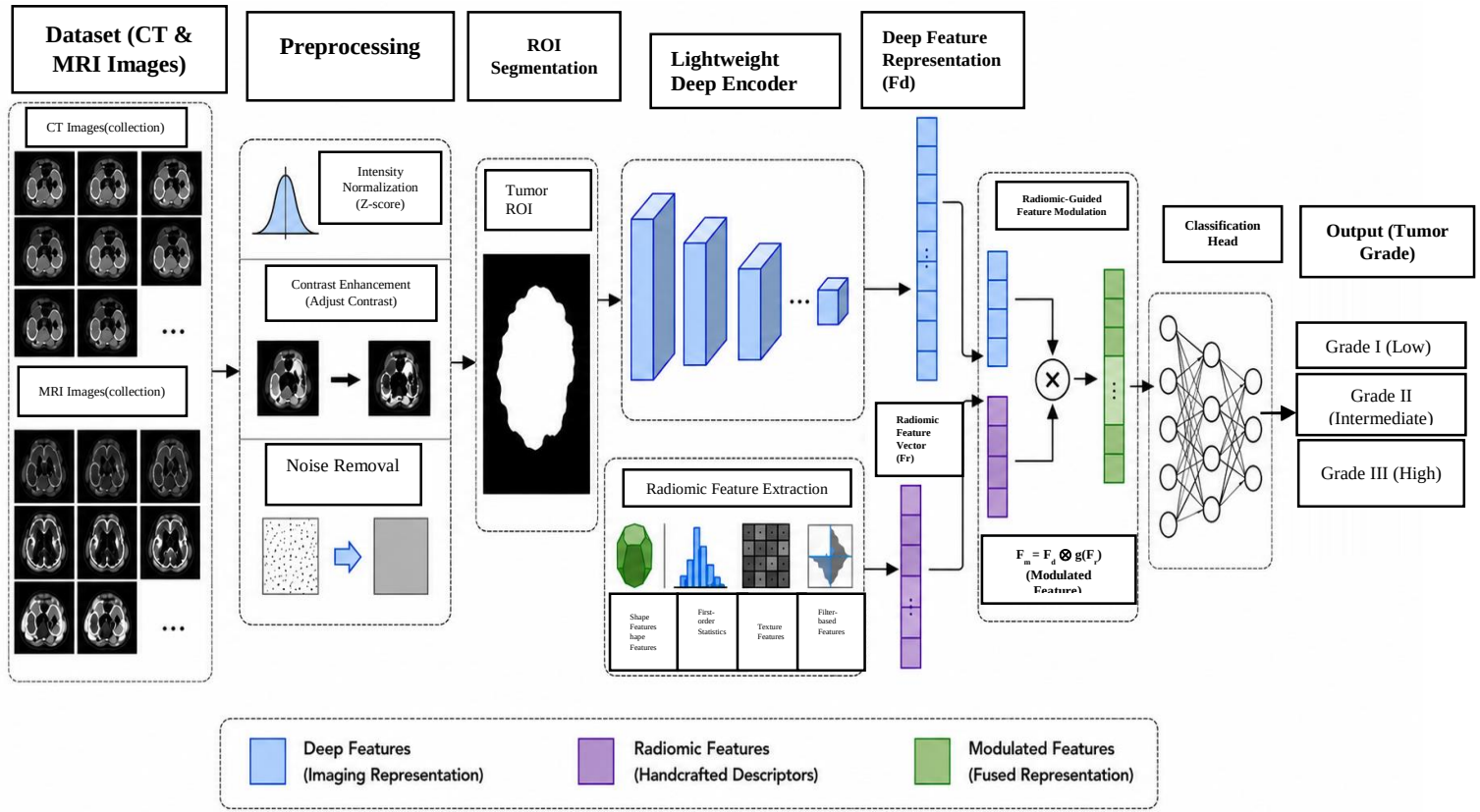

Fig 2. Overall architecture of lightweight deep learning framework