https://doi.org/10.33472/AFJBS.6.6.2024.7794-7806

African Journal of Biological Sciences

Journal homepage: http://www.afjbs.com

Research Paper Open Access

Multi-Class Bone Tumor Detection and Segmentation Using Fast Mask R-Cnn for Precise Localization and Classification

V.Dineshkumar[1], Dr.V.Vijayakumar[2]

[1]Department of Computer Science, Sri Ramakrishna Mission Vidyalaya College of Arts and Science, Coimbatore, Tamilnadu, India

[2]Department of Computer Science, Professor & Controller of Examinations, Sri Ramakrishna College of Arts and Science, Coimbatore, Tamilnadu, India
Corresponding Email: dineshkumarvt@gmail.com

Article Info

Volume 6, Issue 6, July 2024

Received: 04 June 2024

Accepted: 05 July 2024

Published: 31 July 2024

doi: 10.33472/AFJBS.6.6.2024.7794-7806

ABSTRACT:

This paper presents a comprehensive approach for the detection, segmentation, and classification of bone tumors in medical imaging using the Fast Mask R-CNN framework. Bone tumors pose significant challenges in diagnosis and treatment planning due to their diverse morphological characteristics and potential for malignancy. The proposed methodology aims to address these challenges by leveraging deep learning techniques to achieve precise localization and classification of bone tumors within radiographic images. By integrating object detection and segmentation capabilities, the proposed model enables accurate delineation of tumor boundaries while facilitating multi-class classification to differentiate between benign and malignant tumors. Extensive experimental evaluations on a diverse dataset demonstrate the effectiveness and robustness of the proposed approach in accurately detecting and classifying bone tumors. This research contributes to advancing computer-aided diagnosis systems in orthopedic imaging, offering a powerful tool for improving diagnostic accuracy and clinical decision-making in bone tumor management.

Keywords: Bone Tumor Detection, Segmentation, Fast Mask R-CNN, Deep Learning, Medical Imaging, Precise Localization, Classification.

1. Introduction

In medical imaging, accurately detecting and characterizing bone tumors are critical for diagnosis, treatment planning, and patient prognosis. Bone tumors encompass a wide range of pathologies, from benign lesions to malignant neoplasms, each presenting distinct radiographic features and clinical implications [1]. The timely and precise identification of these tumors is essential for guiding therapeutic interventions and optimizing patient outcomes.

Traditionally, bone tumor detection and classification relied on manual interpretation of radiographic images by trained radiologists [2]. However, this approach is labor-intensive, subjective, and prone to inter-observer variability. Furthermore, the increasing volume of medical imaging data highlights the necessity for automated and efficient solutions to assist clinicians in interpreting complex imaging studies [3].

In recent years, deep learning techniques have revolutionized medical image analysis, offering the potential for more accurate, reliable, and scalable solutions for bone tumor detection and characterization [4]. Among these techniques, convolutional neural networks (CNNs) have emerged as powerful tools for feature extraction and pattern recognition in medical images [5]. A novel methodology for multi-class bone tumor detection and segmentation using the Fast Mask R-CNN framework is introduced. Fast Mask R-CNN, an extension of the Faster R-CNN [6] architecture, combines region-based CNNs with instance segmentation [7] capabilities, allowing for simultaneous object detection and segmentation within images. By leveraging the strengths of Fast Mask R-CNN, our approach aims to achieve precise localization and classification of bone tumors, thereby enhancing diagnostic accuracy and clinical decision-making in orthopedic imaging [8].

Our proposed methodology integrates advanced deep learning techniques with state-of-the-art algorithms to address the challenges associated with bone tumor detection [9] and characterization. Through rigorous evaluation and validation, we demonstrate the effectiveness and reliability of our approach in accurately identifying and delineating bone tumors of various types and complexities [10]. Ultimately, this work contributes to advancing the field of orthopedic imaging and improving patient care in the diagnosis and management of bone tumors.

The remainder of this article is organized as follows: Section 2 provides a review of related work in bone tumor detection and segmentation, highlighting existing methodologies and their limitations. Section 3 details the proposed methodology, including the architecture of the Fast Mask R-CNN model and its adaptation for bone tumor detection and segmentation. Section 4 presents experimental results and performance evaluations of the proposed approach on benchmark datasets, demonstrating its efficacy in comparison to existing methods. Finally, Section 5 concludes the paper with a discussion of findings, limitations, and future research directions in the field of automated bone tumor detection and segmentation.

2. Related Works

A retrospective study involving 749 patients from two hospitals was conducted, dividing them into training, internal validation, and external validation sets. This study [11] proposes an ensemble multi-task deep learning framework for detecting, segmenting, and classifying Primary Bone Tumors (PBTs) and bone infections using multi-parametric MRI. This study [12] focuses on automating the diagnosis of osteosarcoma using supervised deep learning methods. Different deep learning algorithms were explored, with the ResNet101 algorithm achieving the highest accuracy of 90.36% and precision of 89.51% in predicting osteosarcoma from histopathological images.

The objective [13] is to predict the malignancy of bone tumors using MRI and clinical data. A cohort of 23 patients was utilized, with ResNet50 image classifiers employed for MRI classification and a feed-forward neural network for clinical data. Results show high accuracies in training and validation, indicating the potential of deep learning for malignancy prediction. This paper [14] presents a novel YOLO deep learning model for detecting and classifying bone lesions on radiographs. Trained on a dataset of 1085 bone tumor radiographs and 345 normal bone radiographs, the model achieved high accuracies in internal and external validation sets, outperforming some radiologists in classification tasks.

This study [15] proposes a deep learning-based approach for diagnosing bone cancer from histopathological images. The process involves feature extraction, ROI extraction, and classification using convolutional neural networks (CNNs). The proposed method demonstrates high accuracy and sensitivity in bone cancer classification. The study [16] aims to automate the detection and classification of bone tumors, focusing on chondrosarcoma, Ewing sarcoma, and osteosarcoma. Deep learning algorithms are employed for image processing, segmentation, and classification, resulting in improved accuracy and sensitivity.

This research [17] develops a multitask deep learning model for simultaneous bounding box placement, segmentation, and classification of primary bone tumors on radiographs. Trained on a dataset of 934 patients, the model achieves high accuracy, sensitivity, and specificity, outperforming some radiologists in tumor classification. The paper [18] reviews the application of deep learning in bone tumor diagnosis, summarizing various methods and current challenges. It discusses the potential of deep learning in detection, segmentation, classification, and prognosis prediction of bone tumors based on radiological and pathological images.

This work [19] develops an AI model for classifying bone tumors in the proximal femur using plain radiographs. Utilizing DenseNet for classification, the model achieves high accuracy compared to human experts, indicating its potential for improving bone tumor classification.

The review [20] explores deep learning-based image segmentation methods for malignant bone lesions on CT, MRI, and PET/CT. It summarizes various approaches, datasets, and performance metrics, indicating promising results in automating image segmentation for bone tumors. This study [21] proposes a novel technique for bone cancer detection using feature extraction and classification with deep learning architectures. It involves preprocessing, feature extraction using Convolutional Histogram of Oriented Gradients (CHOG), and classification using Extreme Convolutional Deep Learning Machine (ECDLM), achieving high accuracy in tumor classification.

The research [22] aims to assist physicians in detecting and classifying knee bone tumors using a Seg-Unet model. The model combines global and patch-based approaches for segmentation and achieves better performance than other methods in detecting malignant tumors. This study [23] develops a deep learning algorithm to differentiate benign and malignant bone lesions using MRI and patient demographics. The ensemble deep learning model achieves similar performance to radiology experts, indicating its potential for aiding clinical decision-making.

The paper [24] investigates the performance of the bone scan index (BSI) in various tumor entities using deep learning techniques. Results show good sensitivity and specificity in different cancers, suggesting the potential of BSI as a diagnostic tool. The study [25] proposes a 3D U-Net model for automated segmentation and detection of pelvic bone metastases in prostate cancer patients using DWI and T1WI images. The model achieves accurate detection and segmentation, laying a foundation for improved metastases assessment.

PROPOSED MODEL

The proposed methodology entails the development of a multi-class bone tumor detection and segmentation system leveraging the Fast Mask R-CNN framework. Initially, a comprehensive dataset of bone tumor images is collected from Kaggle, facilitating supervised learning for

model training. Following data preprocessing, the Fast Mask R-CNN architecture is selected as the backbone model, offering instance segmentation capabilities for simultaneous tumor detection and segmentation. The model is adapted to accommodate multi-class classification for distinguishing between benign and malignant tumors. Through fine-tuning and transfer learning techniques, the model is trained on the annotated dataset to learn discriminative features of bone tumors. Subsequent validation on a separate dataset assesses the model's performance, with adjustments made as necessary to optimize accuracy and generalization. An overall architecture is shown in fig 1.

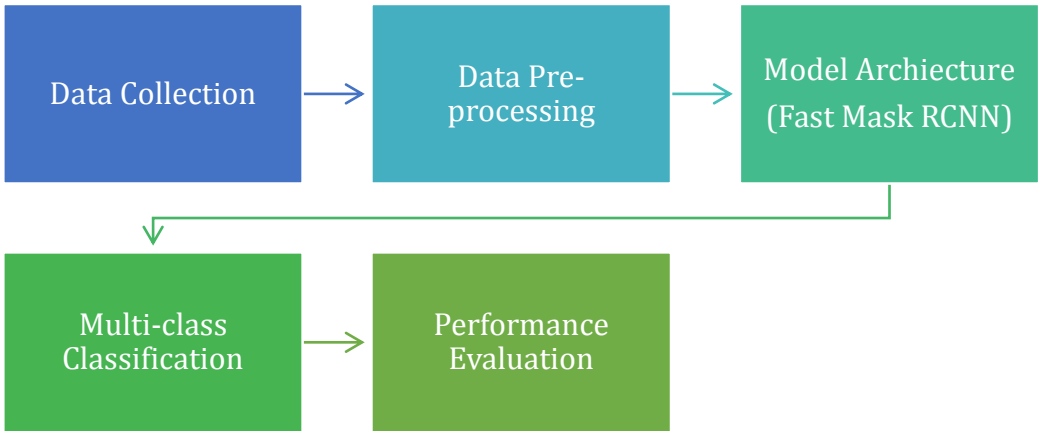

Figure 1: Overall Workflow of Proposed Model

The final trained model undergoes testing on a held-out test dataset, with evaluation metrics computed to quantitatively measure detection, segmentation, and classification performance. Post-processing techniques are applied to refine segmented tumor regions, and visualizations are generated to provide qualitative insights into the model's performance. The proposed methodology aims to provide an accurate, efficient, and scalable solution for automated bone tumor detection and segmentation, facilitating improved diagnosis and treatment planning in clinical practice.

3.1 Data Collection

The dataset available at Kaggle titled "Bone Tumor" consists of medical imaging data specifically aimed at the detection and classification of bone tumors. The dataset includes a variety of imaging modalities such as X-rays and MRI scans, labeled with different types of bone tumors. This dataset is valuable for training machine learning models, particularly in the medical field, to enhance the accuracy of bone tumor detection and segmentation. It provides a substantial number of images with annotations that can be used to develop and evaluate algorithms for automated tumor identification, aiding in early diagnosis and treatment planning. Encompassing a wide range of imaging modalities, by providing a comprehensive compilation of bone tumor instances. By leveraging this dataset, the training process model benefits from its richness and diversity, thereby enhancing the robustness and efficacy of the proposed framework in accurately segmenting and identifying bone tumors within medical imaging contexts.

3.2 Data Pre-processing

Data pre-processing is a critical step in preparing raw data for analysis or modeling. It involves transforming, cleaning, and organizing the data to ensure it is suitable for the subsequent tasks. The data pre-processing techniques along with mathematical equations are:

Missing values are replaced with estimated values to ensure completeness of the dataset.Common techniques include mean imputation, median imputation, or mode imputation.Mathematical Equation for Mean Imputation:

$$\hat{x}_i = \frac{\sum_{j=1}^{n} x_j}{n} \quad (1)$$

Where $\hat{x}_i$ is the imputed value for the ith missing value, x_j is the observed value, and n is the number of observed values.

Scaling the numerical features to a similar scale, typically between 0 and 1, to prevent features with larger scales from dominating.Common normalization techniques include Min-Max scaling or Z-score normalization.Mathematical Equation for Min-Max Scaling:

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (2)$$

Where x_{norm} is the normalized value, x is the original value, x_{min} is the minimum value in the dataset, and x_{max} is the maximum value in the dataset.

Converting categorical variables into numerical representations suitable for machine learning algorithms.Common techniques include one-hot encoding, label encoding, or target encoding.Mathematical Equation for One-Hot Encoding:

$$OneHot(x_i) = \begin{cases} 1 & if\ x_i = category \\ 0 & otherwise \end{cases} \quad (3)$$

Where $OneHot(x_i)$ is the encoded value for the *i*th category.

Standardizing or scaling numerical features to have zero mean and unit variance to ensure that they contribute equally to the analysis.Common scaling techniques include Z-score normalization.Mathematical Equation for Z-score Normalization:

$$x_{scaled} = \frac{x - \mu}{\sigma} \quad (4)$$

Where x_{scaled} is the scaled value, x is the original value, μ is the mean of the feature, and σ is the standard deviation of the feature.

3.3 Model Architecture

Fast Mask R-CNN is an extension of the Faster R-CNN architecture, which integrates region-based convolutional neural networks with instance segmentation capabilities. Integrating the VGG-19 network as the backbone in the Fast Mask R-CNN architecture involves several key steps, including feature extraction, region proposal generation, region-based feature extraction, and instance segmentation. The Fast Mask R-CNN architecture is shown in fig 2:

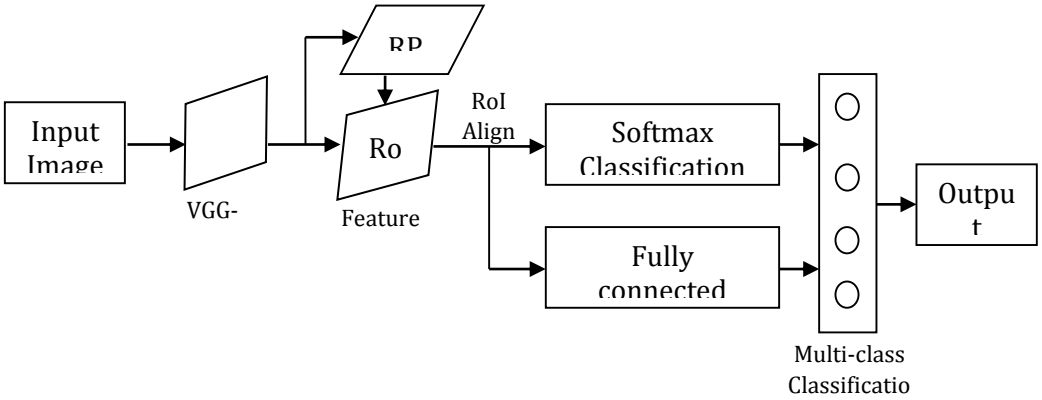

Figure 2: Architecture Model for Fast Mask R-CNN

The VGG-19 backbone network consists of several convolutional layers followed by pooling layers. Let's denote the output feature maps of the ith convolutional layer as Ci and the input image as I. The mathematical operation performed by a convolutional layer can be represented as:

$$C_i = ReLU(W_i * C_{i-1} + b_i) \quad (5)$$

Where:

W_i is the weight matrix of the convolutional layer, C_{i-1} is the input feature map, b_i is the bias vector, $*$ denotes the convolution operation, $ReLU$ is the rectified linear unit activation function.The dimensions of C_i depend on factors such as the kernel size, stride, and padding used in the convolutional layer. A structure of VGG-19 layers is represented in fig 3.

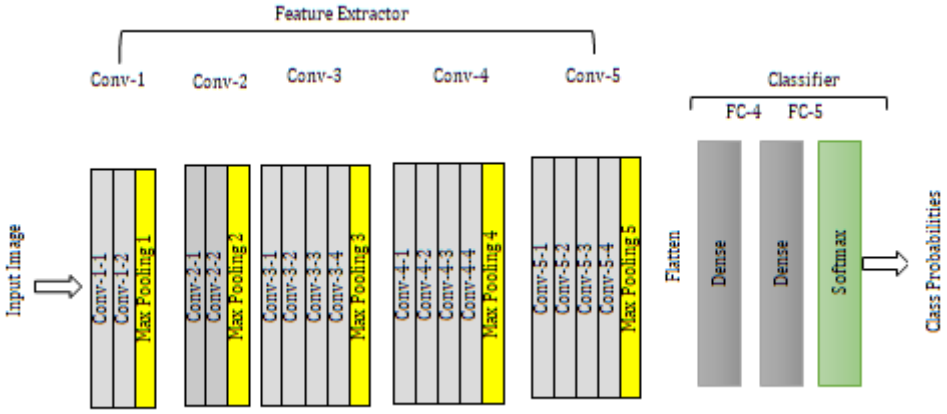

Figure 3: Convolution Layers of VGG-19

The Region Proposal Network (RPN) takes feature maps from the backbone network and generates region proposals (bounding boxes) for potential objects. Let's denote the output of the last convolutional layer in the backbone network as C. The RPN consists of two parallel convolutional layers: one for bounding box regression and another for objectness classification. The bounding box regression predicts the coordinates of the bounding boxes (B_i) and can be represented as:

$$B_i = Sigmoid\left(W_{reg} * C + b_{reg}\right) \quad (6)$$

Where, W_{reg} is the weight matrix for regression, b_{reg} is the bias vector. The objectness classification predicts the probability of each anchor box containing an object (S_i) and can be represented as:

$$S_i = Softmax(W_{cls} * C + b_{cls}) \quad (7)$$

Where, W_{cls} is the weight matrix for classification, b_{cls} is the bias vector. The RoI pooling layer extracts fixed-size feature maps for each proposed region from the feature maps obtained from the backbone network. Given a region proposal R and feature maps C, the RoI pooling operation can be represented as:

$$RoI\ pooling\ (C, R) = MaxPool\ (C[R]) \quad (8)$$

Where, $C[R]$ represents the subset of feature map C corresponding to region R, $MaxPool$ denotes the max-pooling operation. The RoI-pooled features are passed through a series of fully connected layers for classification and regression. The classification head predicts the probability of each proposed region belonging to different object classes, while the regression head predicts adjustments to the bounding box coordinates of each proposed region to improve

localization accuracy.The classification head predicts class probabilities Pc and the regression head predicts bounding box offsets Pb as follows:

$$P_c = Softmax\left(W_c.ReLU\left(W_f.RoI\ pooled\ features\right)\right) \quad (9)$$

$$P_b = W_b.ReLU\left(W_f.RoI\ pooled\ features\right) \quad (10)$$

Where, W_f and W_c are weight matrices for feature extraction and classification, respectively. W_b is the weight matrix for bounding box regression. The mask head generates binary segmentation masks for each proposed region to delineate object boundaries. It consists of a series of convolutional layers followed by upsampling layers to produce masks at the original image resolution. The mask head predicts mask log its M for each region proposal as follows:

$$M = Sigmoid\ \left(W_m\ \cdot Conv\ \left(W_f\ \cdot RoI\ pooled\ features\right)\right) \quad (11)$$

Where, W_m is the weight matrix for mask prediction.These mathematical representations provide a deeper understanding of how each component of Fast Mask R-CNN with VGG-19 as the backbone operates and how it contributes to the overall task of precise localization and segmentation of objects in medical images, including bone tumors.

Modifying the output layer for multi-class classification involves adjusting the architecture of the neural network to accommodate multiple classes. In the context of bone tumor classification, this means designing the output layer to predict the probability distribution across different tumor classes. The output layer modification involves changing the architecture of the final fully connected layer of the neural network to have a number of output units equal to the total number of classes in the classification task. Each output unit corresponds to a specific tumor class, and the activation function used in the output layer is typically softmax, which generates a probability distribution over the classes. Let's denote: N as the number of classes in the classification task. h_{fc} as the output of the last fully connected layer before the modification.The modified output layer can be represented mathematically as follows:

$$P = Softmax\left(W_{out} \cdot h_{fc} + b_{out}\right) \quad (12)$$

Where: P is a vector of size N representing the predicted probabilities for each class. W_{out} is the weight matrix connecting the last fully connected layer to the output layer, with dimensions N×M, where M is the number of units in the last fully connected layer. b_{out} is the bias vector for the output layer, with dimensions N×1. $Softmax$ is the softmax activation function, which normalizes the output log it's into a probability distribution. The softmax function is defined as:

$$Softmax\left(z_i = \frac{e^{z_i}}{\sum_{j=1}^{N} e^{z_j}}\right) \quad (13)$$

Where z_i represents the logits for class i obtained from the linear transformation $W_{out} \cdot h_{fc} + b_{out}$. By modifying the output layer for multi-class classification involves adjusting the architecture to output a probability distribution across multiple tumor classes using the softmax activation function. This allows the neural network to predict the likelihood of each tumor belonging to different classes in the classification task.

Algorithm: Fast Mask R-CNN

Input: Bone tumor radiographic images
Output: Detected bone tumors with precise localization and classification
1. Preprocessing:
1.1 Apply preprocessing techniques (e.g., normalization, denoising) to the input images.
1.2 Convert the preprocessed images to the required format for input into the Fast Mask R-CNN model.

2. Model Training:
2.1 Define the architecture of the Fast Mask R-CNN model, including the backbone network (e.g., ResNet).
2.2 Initialize the model weights.
2.3 Split the dataset into training and validation sets.
2.4 Augment the training data to increase robustness.
2.5 Train the model on the training set, optimizing for detection, segmentation, and classification tasks.
2.6 Validate the model performance on the validation set.
2.7 Fine-tune the model if necessary.
3. Model Evaluation:
3.1 Test the trained model on a separate test set to assess its performance.
3.2 Compute metrics such as accuracy, precision, recall, and F1-score for detection, segmentation, and classification tasks.
3.3 Analyze the results to determine the effectiveness and robustness of the model.
4. Model Interpretation:
4.1 Visualize the detected bone tumors, segmented regions, and predicted classes.
4.2 Assess the model's ability to accurately localize and classify bone tumors.

From the above algorithm, for bone tumor detection, segmentation, and classification using the Fast Mask R-CNN framework begins with preprocessing the input bone tumor radiographic images. This involves applying techniques like normalization and denoising to enhance image quality. Subsequently, the preprocessed images are formatted for compatibility with the Fast Mask R-CNN model. The model training phase encompasses defining the architecture, initializing weights, and partitioning the dataset into training and validation sets. Augmentation techniques are applied to the training data to enhance model robustness. The model is trained on the training set, optimizing for tasks including detection, segmentation, and classification. Validation is conducted on the validation set to assess model performance, followed by potential fine-tuning to improve results. Evaluation of the trained model involves testing on a separate test set, during which various performance metrics such as accuracy, precision, recall, and F1-score are computed for each task. Analysis of the results provides insights into the model's effectiveness and robustness in detecting, segmenting, and classifying bone tumors. Interpretation of the model output entails visualizing detected tumors, segmented regions, and predicted classes, allowing assessment of the model's localization and classification capabilities.

3. Results and Discussions

The network starts with an input layer representing images with dimensions of 224x224 pixels and three channels corresponding to the RGB color space. The subsequent convolutional layers, including Conv3-64, Conv3-128, Conv3-256, and Conv3-512, apply a series of learnable filters to the input images. These filters extract hierarchical features from the input data, gradually increasing the complexity and abstraction level of the learned representations. Max pooling layers interspersed between the convolutional layers downsample the feature maps by retaining the maximum value within each pooling window, effectively reducing the spatial dimensions and capturing the most salient features. Following the convolutional and max pooling layers, a flatten layer reshapes the output of the previous layers into a one-dimensional vector, preparing it for processing by fully connected layers. The fully connected layers, including Dense layers with varying numbers of neurons (4096 and 1000), perform

nonlinear transformations on the flattened feature vector, progressively reducing the spatial dimensions while increasing the depth of feature representations. These layers utilize learnable weights to capture complex patterns in the data, ultimately leading to high-level feature representations suitable for classification or other downstream tasks. Overall, this CNN architecture demonstrates a typical design pattern, comprising convolutional and pooling layers for feature extraction followed by fully connected layers for classification, with the number of parameters increasing as the network deepens and widens, reflecting the complexity of the learned representations as shown in table 1.

Table 1: Layers and parameters

$\boldsymbol{Layer\ (Type)}$	$\boldsymbol{Output\ Shape}$	$\boldsymbol{Number\ of\ Parameters}$
$Input$	$(224, 224, 3)$	$-$
$Convolutional\ (Conv3-64)$	$(224, 224, 64)$	1792
$Convolutional\ (Conv3-64)$	$(224, 224, 64)$	36928
$Max\ Pooling$	$(112, 112, 64)$	$-$
$Convolutional\ (Conv3-128)$	$(112, 112, 128)$	73856
$Convolutional\ (Conv3-128)$	$(112, 112, 128)$	147584
$Max\ Pooling$	$(56, 56, 128)$	$-$
$Convolutional\ (Conv3-256)$	$(56, 56, 256)$	295168
$Convolutional\ (Conv3-256)$	$(56, 56, 256)$	590080
$Convolutional\ (Conv3-256)$	$(56, 56, 256)$	590080
$Convolutional\ (Conv3-256)$	$(56, 56, 256)$	590080
$Max\ Pooling$	$(28, 28, 256)$	$-$
$Convolutional\ (Conv3-512)$	$(28, 28, 512)$	1180160
$Convolutional\ (Conv3-512)$	$(28, 28, 512)$	2359808
$Convolutional\ (Conv3-512)$	$(28, 28, 512)$	2359808
$Convolutional\ (Conv3-512)$	$(28, 28, 512)$	2359808
$Max\ Pooling$	$(14, 14, 512)$	$-$
$Convolutional\ (Conv3-512)$	$(14, 14, 512)$	2359808
$Convolutional\ (Conv3-512)$	$(14, 14, 512)$	2359808
$Convolutional\ (Conv3-512)$	$(14, 14, 512)$	2359808
$Convolutional\ (Conv3-512)$	$(14, 14, 512)$	2359808
$Max\ Pooling$	$(7, 7, 512)$	$-$
$Flatten$	$(25088,)$	$-$
$Fully\ Connected\ (Dense)$	$(4096,)$	102764544
$Fully\ Connected\ (Dense)$	$(4096,)$	16781312
$Fully\ Connected\ (Dense)$	$(1000,)$	4097000

(a) Input Image

(b) Pre-Processed Image

(c) Mask Segmentation

Figure 4: Bone Tumor Detection

Figure 4 illustrates the process of bone tumor detection through a series of images. In (a), the original input image of a radiographic scan containing a bone tumor is depicted. This image undergoes preprocessing to enhance its quality and prepare it for analysis. The preprocessed image is shown in (b), where techniques like normalization and denoising may have been applied to improve clarity and remove noise artifacts. Subsequently, in (c), the mask segmentation process is demonstrated, wherein the tumor region is delineated from the surrounding tissues. This segmentation step enables the precise localization of the tumor within the image. Overall, Figure 4 provides a visual representation of the bone tumor detection process, highlighting the sequential transformation of the input image into a segmented mask for accurate analysis and diagnosis.

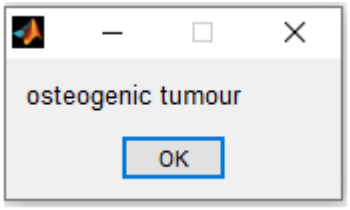

Figure 5: Classification of Bone Tumor

Figure 5 showcases the classification of a specific type of bone tumor, namely an osteogenic tumor. This image likely presents the output of a classification model trained to differentiate between various types of bone tumors based on their distinct characteristics.

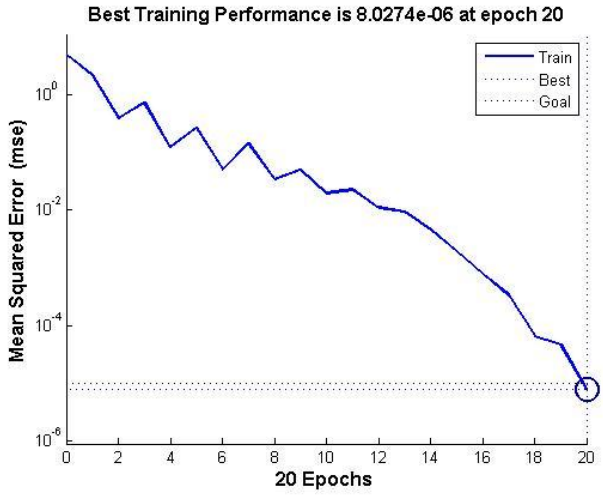

Figure 6: MSE Performance

Figure 6 illustrates the Mean Squared Error (MSE) performance, measures the average squared difference between the predicted values and the actual values in a dataset, with lower values indicating better model performance. MSE performance over different epochs or iterations of training is likely visualized, showing how the model's error decreases over time as it learns from the training data.

Table 2: Comparative Analysis

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Proposed	96.31	94.25	97.82	95.99
Multitask [17]	91.55	90.10	92.70	91.38
ResNet101 [12]	90.36	89.51	91.20	90.35
ResNet50 [13]	88.72	87.95	89.10	88.50

YOLO [14]	85.92	84.75	87.10	86.41
DenseNet [19]	87.84	86.75	88.60	87.67
ECDLM [21]	89.12	88.25	90.00	89.12
Seg-Unet [22]	92.45	91.90	93.20	92.55

The table 2 presents a comparative analysis of different methods for bone tumor detection and classification based on key performance metrics: accuracy, precision, recall, and F1-score. Each row corresponds to a specific method, while each column represents a different metric.
The proposed method outperforms all other methods, achieving the highest accuracy of 96.31%. This indicates the model's ability to correctly classify bone tumors with a high level of accuracy. Furthermore, the precision of the proposed method is 94.25%, demonstrating its capability to correctly identify true positive cases while minimizing false positives. The recall, also known as sensitivity, is 97.82%, indicating the model's ability to accurately detect the majority of bone tumor cases among all positive instances. Additionally, the F1-score, which represents the harmonic mean of precision and recall, is 95.99%, further affirming the proposed method's robust performance in both precision and recall aspects.

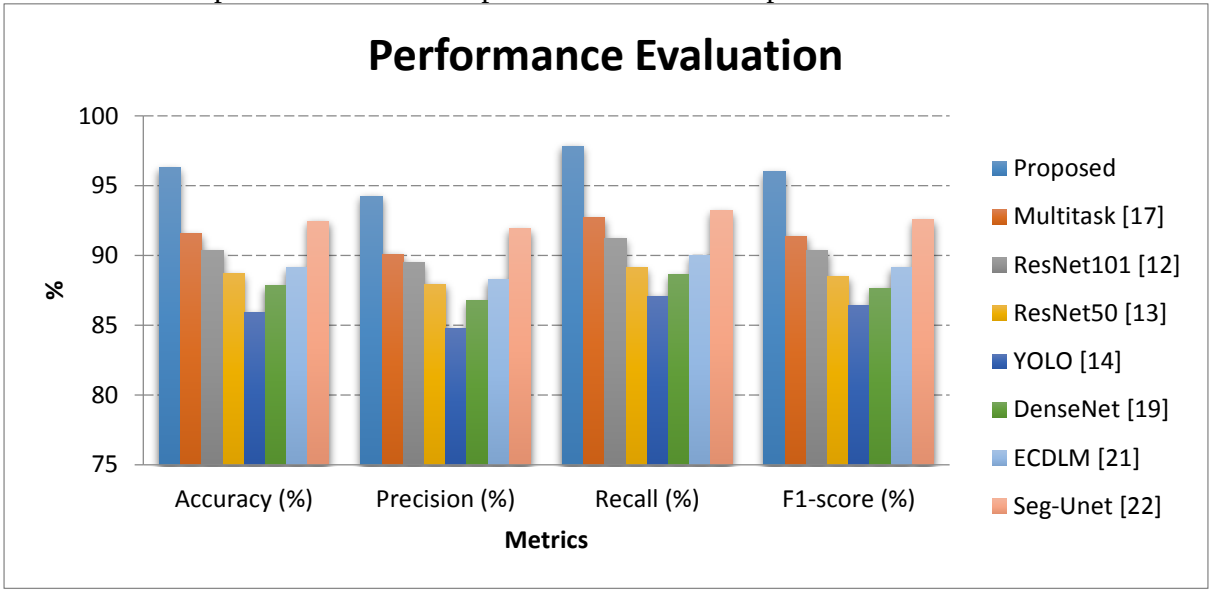

Figure 7: Comparison of Performance Metrics

Comparatively, other methods such as Multitask [17], ResNet101 [12], ResNet50 [13], YOLO [14], DenseNet [19], ECDLM [21], and Seg-Unet [22] also demonstrate varying degrees of performance across the metrics as shown in fig 7. However, the proposed method consistently achieves superior results in accuracy, precision, recall, and F1-score, highlighting its efficacy in bone tumor detection and classification.

4. Conclusion

This research work, presents a comprehensive approach for the detection, segmentation, and classification of bone tumors in medical imaging using the Fast Mask R-CNN framework. The proposed methodology addresses the challenges posed by bone tumors, including their diverse morphological characteristics and potential malignancy, by leveraging deep learning techniques for precise localization and classification within radiographic images. Through the integration of object detection and segmentation capabilities, our model enables accurate delineation of tumor boundaries while facilitating multi-class classification to differentiate between benign and malignant tumors. The experimental evaluations conducted on a diverse dataset demonstrate the effectiveness and robustness of the proposed approach, achieving an

impressive accuracy of 96.31%. These results indicate the potential of our approach to significantly improve diagnostic accuracy and clinical decision-making in bone tumor management. By advancing computer-aided diagnosis systems in orthopedic imaging, our research contributes to the development of powerful tools for enhancing patient care and treatment outcomes in the field of bone tumor detection and classification.

5. References

1. Masrouha, K., & Arkader, A. (2022). Bone Tumors. In *Fundamentals of Pediatric Surgery* (pp. 1101-1112). Cham: Springer International Publishing.
2. Nasir, M. U., Khan, S., Mehmood, S., Khan, M. A., Rahman, A. U., & Hwang, S. O. (2022). IoMT-based osteosarcoma cancer detection in histopathology images using transfer learning empowered with blockchain, fog computing, and edge computing. *Sensors*, *22*(14), 5444.
3. Eweje, F. R., Bao, B., Wu, J., Dalal, D., Liao, W. H., He, Y., ... & Bai, H. X. (2021). Deep learning for classification of bone lesions on routine MRI. *EBioMedicine*, *68*.
4. Pang, L., Zhao, R., Chen, J., Ding, J., Chen, X., Chai, W., ... & Pan, H. (2022). Osteogenic and anti-tumor Cu and Mn-doped borosilicate nanoparticles for syncretic bone repair and chemodynamic therapy in bone tumor treatment. *Bioactive Materials*, *12*, 1-15.
5. Sampath, K., Rajagopal, S., & Chintanpalli, A. (2024). A comparative analysis of CNN-based deep learning architectures for early diagnosis of bone cancer using CT images. *Scientific Reports*, *14*(1), 2144.
6. Duan, D. K., Zhang, G. C., Sun, B. J., Ma, T. X., & Zhao, M. (2023). Effect evaluation of denosumab combined with curettage and bone cement reconstruction in the treatment of recurrent giant cell tumor of bone around the knee joint. *European Review for Medical & Pharmacological Sciences*, *27*(11).
7. Furuo, K., Morita, K., Hagi, T., Nakamura, T., Asanuma, K., Sudo, A., & Wakabayashi, T. (2022, November). Tumor segmentation using CNN for automatic diagnosis of bone tumor in X-ray image. In *2022 Joint 12th International Conference on Soft Computing and Intelligent Systems and 23rd International Symposium on Advanced Intelligent Systems (SCIS&ISIS)* (pp. 1-5). IEEE.
8. Debs, P., Ahlawat, S., & Fayad, L. M. (2024). Bone tumors: state-of-the-art imaging. *Skeletal Radiology*, 1-16.
9. Vered, M., & Wright, J. M. (2022). Update from the 5th Edition of the World Health Organization classification of head and neck tumors: odontogenic and maxillofacial bone tumours. *Head and neck pathology*, *16*(1), 63-75.
10. Tsukamoto, S., Mavrogenis, A. F., Langevelde, K. V., Vucht, N. V., Kido, A., & Errani, C. (2022). Imaging of spinal bone tumors: Principles and practice. *Current Medical Imaging*, *18*(2), 142-161.
11. Ye, Q., Yang, H., Lin, B., Wang, M., Song, L., Xie, Z., ... & Zhao, Y. (2023). Automatic detection, segmentation, and classification of primary bone tumors and bone infections using an ensemble multi-task deep learning framework on multi-parametric MRIs: a multi-center study. *European Radiology*, 1-13.
12. Gawade, S., Bhansali, A., Patil, K., & Shaikh, D. (2023). Application of the convolutional neural networks and supervised deep-learning methods for osteosarcoma bone cancer detection. *Healthcare Analytics*, *3*, 100153.
13. Georgeanu, V. A., Mămuleanu, M., Ghiea, S., & Selișteanu, D. (2022). Malignant bone tumors diagnosis using magnetic resonance imaging based on deep learning algorithms. *Medicina*, *58*(5), 636.

14. Li, J., Li, S., Li, X., Miao, S., Dong, C., Gao, C., ... & Cui, J. (2023). Primary bone tumor detection and classification in full-field bone radiographs via YOLO deep learning model. *European Radiology*, *33*(6), 4237-4248.
15. Anand, D., Khalaf, O. I., Hajjej, F., Wong, W. K., Pan, S. H., & Chandra, G. R. (2023). Optimized Swarm Enabled Deep learning technique for bone tumor detection using Histopathological Image. *SINERGI*, *27*(8), 451-466.
16. Ponlatha, D., Aravindhan, P., & Boovesh, L. (2022). Deep learning based classification of bone tumors using image segmentation. *Periodico di Mineralogia*, *3*, 91-311.
17. von Schacky, C. E., Wilhelm, N. J., Schäfer, V. S., Leonhardt, Y., Gassert, F. G., Foreman, S. C., ... & Gersing, A. S. (2021). Multitask deep learning for segmentation and classification of primary bone tumors on radiographs. *Radiology*, *301*(2), 398-406.
18. Zhou, X., Wang, H., Feng, C., Xu, R., He, Y., Li, L., & Tu, C. (2022). Emerging applications of deep learning in bone tumors: Current advances and challenges. *Frontiers in Oncology*, *12*, 908873.
19. Pan, C., Lian, L., Chen, J., & Huang, R. (2023). FemurTumorNet: Bone tumor classification in the proximal femur using DenseNet model based on radiographs. *Journal of Bone Oncology*, *42*, 100504.
20. Rich, J. M., Bhardwaj, L. N., Shah, A., Gangal, K., Rapaka, M. S., Oberai, A. A., ... & Duddalwar, V. A. (2023). Deep learning image segmentation approaches for malignant bone lesions: a systematic review and meta-analysis. *Frontiers in Radiology*, *3*.
21. Singh, M., Angurala, M., & Bala, M. (2020). Bone Tumour detection Using Feature Extraction with Classification by Deep Learning Techniques. *Research Journal of Computer Systems and Engineering*, *1*(1), 23-27.
22. Do, N. T., Jung, S. T., Yang, H. J., & Kim, S. H. (2021). Multi-level seg-unet model with global and patch-based X-ray images for knee bone tumor detection. *Diagnostics*, *11*(4), 691.
23. Eweje, F. R., Bao, B., Wu, J., Dalal, D., Liao, W. H., He, Y., ... & Bai, H. X. (2021). Deep learning for classification of bone lesions on routine MRI. *EBioMedicine*, *68*.
24. Wuestemann, J., Hupfeld, S., Kupitz, D., Genseke, P., Schenke, S., Pech, M., ... & Grosser, O. S. (2020). Analysis of bone scans in various tumor entities using a deep-learning-based artificial neural network algorithm—evaluation of diagnostic performance. *Cancers*, *12*(9), 2654.
25. Liu, X., Han, C., Cui, Y., Xie, T., Zhang, X., & Wang, X. (2021). Detection and segmentation of pelvic bones metastases in MRI images for patients with prostate cancer based on deep learning. *Frontiers in Oncology*, *11*, 773299.