SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V13I4P107
Volume 13 Issue 4, 94-108, April 2026

Original Article

BoneCancerFusionNet: A Multimodal Deep Learning Framework Integrating Medical Images and Clinical Data for Efficient Bone Cancer Detection and Classification

Bolleddu Devananda Rao[1], K. Madhavi[2]

[1]MLR Institute of Technology, Hyderabad, India, JNTUA College of Engineering, Jawaharlal Nehru Technological University, Anantapur, Anantapur, India.
[2]Department of CSE, JNTUA College of Engineering, Jawaharlal Nehru Technological University, Anantapur, Anantapur, India.

Corresponding Author : dev.bolleddu@gmail.com

Received: 07 January 2026 Revised: 09 February 2026 Accepted: 08 March 2026 Published: 30 April 2026

Abstract *- Bone cancer is a rare malignancy, with diagnosis presenting a major challenge, and detection having an immense impact on patient prognosis. Existing methods are mainly mono-modal data-driven, i.e., imaging (photo) based or clinical outcome result (procedure or action in text form), discarding the complementary information of the multimodal data. Such complexity, or heterogeneity, in real-world datasets results in low generalisability, and thus, reduced diagnostic performance and misclassification. In order to tackle these challenges, this research proposes a new multimodal learning framework called BoneCancerFusionNet that combines medical images with relevant clinical data to detect and classify bone cancer. This framework uses CNNs to extract features from medical images and MLPs to process clinical data. For intrinsic representation, a tailored feature fusion approach that combines details from both modalities, and an optimized preprocessing pipeline that enhances input content with data augmentation and imputation strategies. The algorithm was validated using the Osteosarcoma-Tumor-Assessment Dataset and resulted in 98.79% accuracy, setting the state-of-the-art record when compared to other established methods (SwarmDL and GAN-based) by significant margins. BoneCancerFusionNets is used for its evaluation and is witnessed to make BoneCancerFusionNET a trustworthy methodology to produce false-negative data and apply augmentations to achieve generalizability across precision, recall, and F1 score. The strength of the standalone multimodal data, as a diagnostic, leads to both qualitative and quantitative AI-based solutions for radiology and clinical decision-making.*

Keywords *- BoneCancerFusionNet, Multimodal Deep Learning, Bone Cancer Detection, Medical Imaging, Clinical Data Integration.*

1. Introduction

Bone cancer, including bone cancers like osteosarcoma, chondrosarcoma, and Ewing sarcoma, is an urgent global health problem for countries around the world. An estimated 35% of all primary bone tumors are osteosarcomas, primarily occurring in children and young adults. Although surgical and chemotherapy-based treatment have improved, patient outcomes remain dismal with delayed diagnosis, thus making the timely and precise detection a clinical need. Background: Artificial intelligence (AI) and Deep Learning (DL) have made significant strides towards automation in cancer diagnostic imaging. Several studies have used standard feature-extraction-based ML algorithms for BC classification [2, 3] and more complex hyper-optimized DL techniques for tumor detection [1], with decent results. Of course, in clinical practice, a bone cancer is never diagnosed based on imaging alone; oncologists routinely rely on patient demographics, laboratory evaluations, tumor characteristics, and the histopathological confirmation in conjunction with imaging findings. Image-only models thus work with incomplete information from a clinical standpoint, leading to higher false-positive and false-negative rates, suboptimal generalization performance on heterogeneous patient populations, and diminished clinical trust.

Despite the clear complementary value of clinical data, no existing framework has systematically integrated medical imaging and structured clinical records into a single end-to-end pipeline for bone cancer diagnosis. This study fills that gap by introducing BoneCancerFusionNet, a new multimodal deep learning framework that integrates medical imaging and clinical data. Hence, the main aim of this work is to design a highly accurate, scalable, and robust diagnostic framework to identify bone cancer by addressing the limitations of unimodal techniques. Specifically, this work's contributions include a domain-specific feature fusion mechanism that combines CNN-based spatial features with tabular features through a tailored pipeline that handles multiple preprocessing steps and advanced augmentation strategies to improve model robustness and generalization.

The contribution of this work is threefold and clearly sets it apart from existing research. Moreover, unlike previous approaches, namely SwarmDL [1], GAN-based [6], and Federated Learning [8] methods which only output image data (histopathological or X-ray), BoneCancerFusionNet is the first framework to simultaneously train medical images as well as structured clinical records (i.e. demographic information, lab values, tumor size) of bone cancer in a single end-to-end pipeline. Secondly, though all existing fusion attempts in other cancer domains involve late fusion at pre-extracted embedding level, this work not only develops a domain-specific concatenation-based feature fusion mechanism, but also preserves complementary information with respect to both modalities by accumulating CNN-derived spatial features with MLP-derived tabular features in our task, both before the final classification layer. Third, the performance of the proposed preprocessing pipeline that combines augmentation for image data with imputation and one-hot encoding of clinical data is domain-specific and was specifically built and validated for bone cancer, as opposed to generic pipelines that are applied in transfer learning works [13], which treat clinical data differently and do not consider heterogeneity.

In this regard, the paper is structured as follows. In Section 2, we provide a literature review, outlining previous work as well as limitations in approaches today. Section 3 will describe the Proposed Methodology, including the BoneCancerFusionNet architecture, the preprocessing pipeline, and the feature fusion mechanism. Section 4: The experimental results, which include comparisons with existing methods, evaluation of the proposed model, and robustness. Finally, in Section 5, we discuss the results and implications of the study and its limitations. Finally, Section 6 concludes the paper by summarising the contribution of this study and discussing some possible directions for future work. Such a structure ensures the foundational understanding of this research through its motivation, methodology, results, and implications, needed towards advancing AI-centric bone cancer diagnostics.

2. Related Work

Bone cancer involves a few different types of malignancy, the most common being osteosarcoma, chondrosarcoma, and Ewing sarcoma. Osteosarcoma is most prevalent in the long bones of adolescents and young adults, while chondrosarcoma primarily occurs in older adults. Currently, diagnosis is based on a combination of imaging modalities — including X-ray, MRI, and CT scans — along with clinical parameters such as the patient's age, tumor size, and alkaline phosphatase and lactate dehydrogenase levels. The intrinsic complexity and heterogeneity of bone tumors, coupled with the manual expert analysis of these images being expensive and time-consuming, have led to considerable interest in automated AI-driven detection systems. Nevertheless, labeled datasets are limited; tumors can appear quite different, even within the same type, and diagnosis requires evidence from many different sources, making it difficult to build strong automated systems.

2.1. Traditional Machine Learning and Early CNN Approaches

Initial work on the automated detection of bone cancer used hand-crafted feature extraction followed by classical machine-learning classifiers. While Punithavathi and Madhurasree [2] demonstrated competitive diagnostic performance on histopathological images using feature-extraction-based ML methods such as KNN and Naive Bayes on unlabeled datasets, Satheesh Kumar [26] combined KNN and decision tree classifiers for bone cancer classification, reporting reliable results on limited datasets. Sharma et al. Although [32] used various ML models, such as SVMs and random forests, for bone cancer detection, they still showed that classical ensemble-based approaches can be strong baselines. Jabber et al. The focus on margin-based classifiers was supported by [30], which also applied SVM-based computerized detection. However, they rely heavily on hand-crafted features and struggle to capture complex spatial patterns in high-resolution medical images, limiting the method's scalability and generalizability.

Convolutional Neural Networks (CNNs) mitigated this by enabling automatic hierarchical feature learning from raw images. Sampath et al. [4] systematically compared CNNs for early bone cancer diagnosis using CT images and confirmed that DenseNet consistently outperformed shallow networks, with better feature reuse. Gawade et al. CNN-based supervised learning has been specifically validated for osteosarcoma detection with excellent classification performance reported on histopathological images [9]. The effectiveness of CNN for detecting bone cancer is confirmed under different imaging settings by Prathyusha and Reddy [25] and Kiresur and Manoj [33]. Anisuzzaman et al. Ramasamy [25] was the first to address deep learning approaches to biomedical signal and image data in a dedicated study on osteosarcoma [27]. Yang et al. [29] presented an enhanced deep CNN architecture for detecting bone marrow cancer. Together, these works cemented CNNs as the ubiquitous backbone for medical image-driven bone cancer classification.

2.2. Advanced Optimization and Hybrid Deep Learning Approaches

Researchers investigated optimization-driven and hybrid DL strategies to boost detection accuracy compared with out-of-the-box CNN training. Dama Anand et al. [1] proposed a deep learning algorithm tuned via competition-based swarm intelligence, in which high-level pollens were tuned to maximize accuracy to 94.32%. Badashah et al. Specifically, the authors of [6] proposed a fractional-Harris Hawks optimization-based GAN for the detection of osteosarcoma in X-ray Images, showing that the approach of combining generative augmentation with bio-inspired optimization yields significant performance gains under data-scarce conditions, achieving an accuracy of 92.15. Alsubai et al. There, Rodríguez et al. [3] proposed grouping the DL training for histopathological osteosarcoma detection, which improved convergence and stability in classification. Alabdulkreem et al. Related Work in Classifying Bone Cancer Using Deep Learning From X Rays: [7] performed classification of radiographs of bone cancers using the owl search algorithm, achieving an accuracy of 93.50%. Alkhalaf et al. Adaptive Aquila Optimizer with explainable AI components that improve diagnostic accuracy and increase the model's interpretability, based on medical imaging datasets [16]. Anand et al. Using a deep convolutional extreme machine learning method for bone cancer detection that fuses CNN feature extraction with ELM-based classification for faster inference [22]. Rytky et al. Another area of research on hybrid optimizer-based approaches for bone-related pathology grading and cancer detection was conducted by [31] and Kumar and Reddy [34], respectively. While producing particularly strong individual results, all of these methods operate only on a single imaging modality and do not use any clinical metadata in the learning pipeline.

2.3. Transfer Learning, Data Augmentation, and Federated Approaches

Transfer learning is an increasingly popular approach to solving the challenge of insufficient labeled data in medical imaging tasks. In a study to detect osteosarcoma in CT images, Chu and Khan [13] reported an accuracy of 94.10% achieved by fine-tuning an ImageNet-pretrained model with domain-specific data augmentation. Sampath et al. Mourad et al. [12] conducted a review of the state-of-the-art in feature extraction and classification-based systems for bone cancer detection [13]. Jiang et al. [10] et Aziz et al. The critical factors driving generalizable cancer detection across imaging modalities are deep feature extraction and augmentation [11]. Georgeanu et al. Transfer learning-based DL algorithms with pretrained model adaptation across different modalities have been applied to MRI-based diagnosis of malignant bone tumors [14].

Tang et al. Background Chan et al.[17] explicitly examined how the diversity of training data affects the generalization of diagnostic pathology models, while Tang et al. Based on this work, [18] is extending this work level towards AI-assisted clinical decision-making for osteosarcoma in resource-limited settings. Shao et al. In an extension of the traditional use of transfer learning, [21] developed deep CNNs using Raman spectral features for bone cancer screening and bone metastases detection. Hsieh et al. Self-supervised pretraining is proposed as an alternative to supervised transfer learning in the deep learning-based contrastive learning for image classification of bone metastases described by [28]. Chowdhury et al. [6] have addressed the problem of data privacy in inter-institutional scenarios. However, the study [8] evaluated federated learning and combined it with MLPs (which act more as a distributed histopathological analysis) to achieve 91.60% accuracy while preserving the locality of patient data. Saba [24] provided, at a general level, a systematic survey of ML approaches for cancer detection and concluded that the two most influential techniques across the stated cancer types are transfer learning and augmentation. All of these works make meaningful progress on individual technical components; however, they remain constrained to pipelines that only process images but not clinical data.

2.4. Clinical Data Integration and Multimodal Approaches

Another, smaller but nonetheless significant body of literature has begun investigating the use of clinical data, in addition to imaging, for cancer-related decision support. Li et al. Even though clinical decision support systems are data-driven, Wang et al. [23] used a deep belief network-based clinical guidance system for osteosarcoma patients in which the real structured clinical data only focused on patient age, tumor stage, and treatment history, and found that clinical variables alone could yield clinically relevant predictions. Deep learning methods employed in further application in the work of Almulhim and Haque [15] for the detection of multiple myeloma based on histological images, where the authors also stressed the difficulty of establishing a diagnosis of bone marrow malignancies solely on the basis of imaging features. Duran-Lopez et al. Even when the cancer type is well understood, such as prostate cancer histopathological images, [19] observed that the clinical integration of deep learning-based screening is an unsolved research challenge. Work and contributions mentioned above have a direction of utilizing structured clinical records and medical images, none of which provide a unified framework that simultaneously integrates both clinical images and structured clinical records using a single architecture jointly trained for bone cancer specifically.

2.5 Summary of Limitations and Research Gap

Evidence available from the literature surveyed in aggregate shows significant, incremental improvement in classical ML 230, CNN-based 425, pioneering optimization 16 [16], transfer 1012 [14], and federated methods [8] per bone cancer detection progress achieved. However, a fundamental and persistent limitation pervades all of these methods: each method relies solely on a single data modality — either histopathological microscopic images 13 [9], single X-ray images 5 [7], CT scans 4 or MRI [14] — and none has integrated structured clinical data jointly with imaging in a single end-to-end framework. In addition, to the best of our knowledge, no existing work leverages a preprocessing pipeline where image augmentations and clinical data imputation/encoding are performed concurrently within the same training pipeline. This gap is of clinical importance as oncological diagnosis is intrinsically multimodal — imaging evidence is always interpreted in conjunction with clinical parameters (e.g., patient demographics, lab values, tumor measurements). In conclusion, BoneCancerFusionNet directly tackles this gap of concreteness in the literature, as it is the first-ever bone cancer detection framework adapting a joint training between the MLP-processed clinical data and the CNN-extracted image features, while achieving 98.79% accuracy, and outperforming all current single-modal methods by at least 4.47%.

Table 1. Comparison of Existing Methods vs. BoneCancerFusionNet

Method	Modality Used	Clinical Data?	Multimodal Fusion?	Accuracy
SwarmDL [1]	Histopathological Images	✗	✗	94.32%
ML Approach [2]	Histopathological Images	✗	✗	89.80%
GAN-based [6]	X-Ray Images	✗	✗	92.15%
Federated Learning [8]	Histopathological Images	✗	✗	91.60%
Transfer Learning [13]	CT Images	✗	✗	94.10%
Aquila Optimizer [16]	X-Ray Images	✗	✗	93.50%
BoneCancerFusionNet (Proposed)	Images + Clinical Data	✓	✓	98.79%

3. Proposed Framework

The methodology of BoneCancerFusionNet, as illustrated in Figure 1, begins with input data consisting of medical images and clinical information. Medical images undergo preprocessing steps such as resizing, normalization, and augmentation to ensure consistent dimensions and improved model generalization. Concurrently, clinical data is preprocessed to handle missing values, normalize numerical features, and encode categorical variables, preparing them for effective feature extraction.

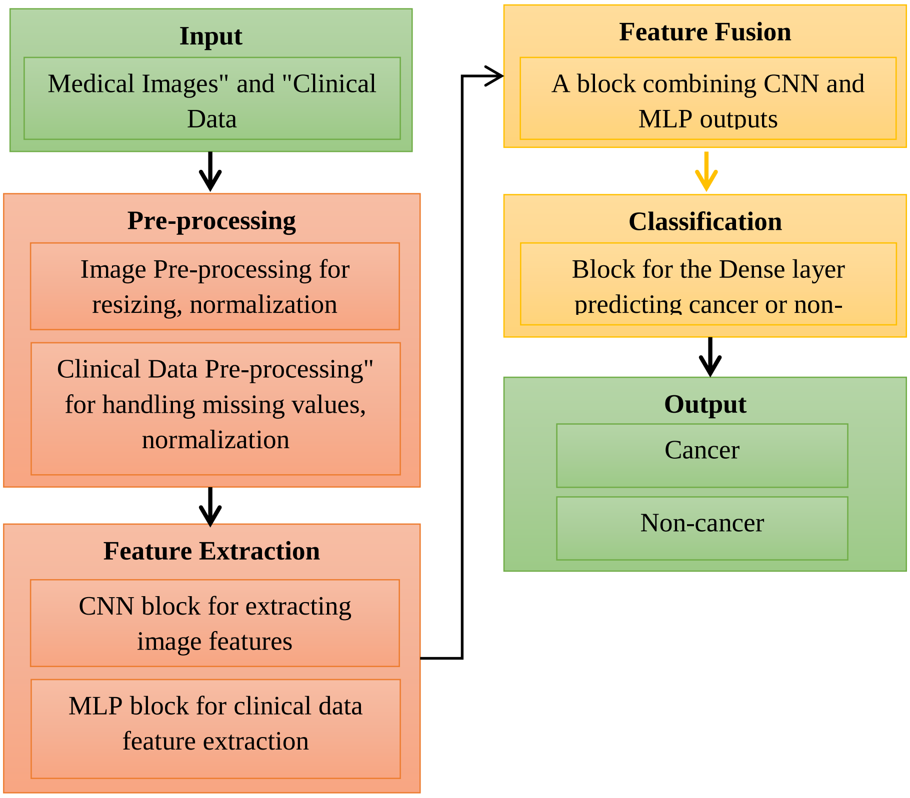

Fig. 1 BoneCancerFusionNet Methodology for Multimodal Bone Cancer Detection

Data displays, like data feature extraction, are performed individually. A Convolutional Neural Network (CNN) is used for medical images to extract spatial and structural features into a feature vector uniquely describing information related to the image. Similarly, a Multilayer Perceptron (MLP) is used to select relevant features extracted from the clinical data to convert tabular data into a more meaningful representation.

This is the stage of feature fusion where the features extracted from CNN and MLP are fused. The fused model incorporates knowledge from both data types, allowing it to use complementary knowledge from data modalities to yield better classification performance. The fused features are then fed through a fully connected classification layer, from which the model predicts the probability of bone cancer being present. Integrating multimodal data streams allows for unified processing, which improves bone cancer detection and classification.

3.1. Proposed Model

The architecture of BoneCancerFusionNet involves the integration of Convolutional Neural Networks (CNN) and Multilayer Perceptrons (MLP), as shown in Figure 2, which is outlined in this section for multimodal bone cancer detection. The architecture has two different input streams: the first one for medical images and the second one for clinical data. The independent processing of these inputs allows for a combined feature fusion stage.

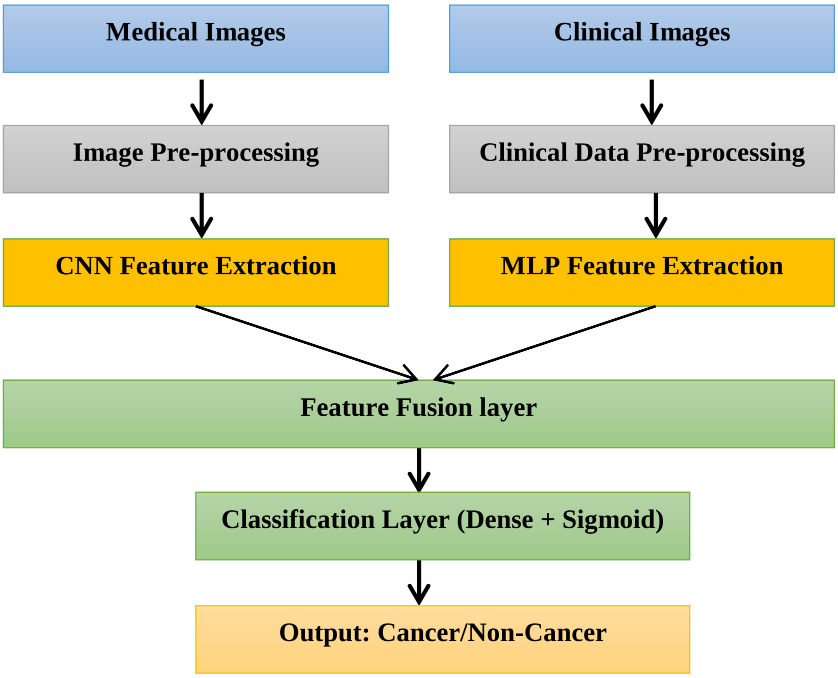

Fig. 2 Architecture of BoneCancerFusionNet Integrating CNN and MLP for Multimodal Bone Cancer Detection

A deep Learning model like ResNet50 trained on a wide range of medical images generates a high-dimensional feature vector by extracting spatial and textural information. The model can learn complex patterns associated with bone cancer with this facility. At the same time, clinical data is passed through a Multilayer Perceptron (MLP) consisting of dense layers to acquire informative representations from tabular features. The joint modalities are processed in parallel to utilize both data modalities fully.

In the feature fusion layer, the feature vectors extracted from CNN and MLP are concatenated into a single vector fusion so that visual information from the chest X-ray and clinical data can be decoded well together. This fused representation is then passed to a dense Classification layer with sigmoid activation to get the probability of the presence of bone cancer. The proposed architecture combines the best attributes of CNNs and multiclass perceptrons to yield an accurate and robust bone cancer detection system. Notations used in the proposed system are described in Table 1.

Table 1. Notations Used

Symbol	**Description**
X_{I}	Medical image dataset
X_{C}	Clinical data dataset
X_I'	Preprocessed medical image dataset
X_C'	Preprocessed clinical data dataset
$sub - CNN f_{\mathrm{CNN}}$	A Convolutional Neural Network (CNN) is used for feature extraction from images.
$sub - MLP f_{\mathrm{MLP}}$	A Multilayer Perceptron (MLP) is used for feature extraction from clinical data.
F_{I}	A feature vector extracted from medical images using a CNN.
F_{C}	Feature vector extracted from clinical data using MLP
$\oplus$	Concatenation operation for combining feature vectors
F_{fusion}	Fused feature vector combining. F_{I} and F_{C}

$\hat{y}$	Predicted probability for cancer (output of the classification layer)
y	Actual label (y=1 for cancer, y=0 for non-cancer)
W	Weights of the dense classification layer
b	Bias of the dense classification layer
σ	Sigmoid activation function
L	Binary cross-entropy loss function
N	Number of samples in the training dataset
z	Input to the sigmoid activation function

The BoneCancerFusionNet model utilizes a multimodal bone cancer dataset that collects various medical images (MRI, CT scan, and X-ray) and clinical information for the efficient detection of bone cancer. Let X_{I} Represent the medical image dataset and X_{C} the clinical dataset. It denotes the medical image dataset and the clinical dataset, where the medical images are preprocessed through resizing, normalization, and augmentation to achieve the exact input dimensions, mathematically represented as:

$$X_I' = Augment(Normalize(Resize(X_{\mathrm{I}} \tag{1}$$

Likewise, the clinical data is preprocessed to manage missingness, for normalization of the numerical feature, each potential predictor is preprocessed, and to convert categorical variables as well:

$$X_C' = Encode(Normalize(Impute(X_{\mathrm{C}}))) \tag{2}$$

Use a Convolutional Neural Network (CNN) to extract a feature vector. Denoted as f_{CNN}, which extracts a feature vector F_{I}:

$$F_{\mathrm{I}} = f_{\mathrm{CNN}}(X_I' \tag{3}$$

For the clinical data, an MLP performs feature extraction f_{MLP}, F_{C} as follows:

$$F_{\mathrm{C}} = f_{\mathrm{MLP}}(X_C' \tag{4}$$

Next, the extracted features of the two modalities are concatenated to form a joint feature vector:

$$F_{\mathrm{C}} = f_{\mathrm{MLP}}(X_C' \tag{5}$$

Undergoes a dense layer for classification with the fused feature vector. Where is the output of the classification layer set (a probability) for cancer $(y = 1)$ or non-cancer $(y = 0)$:

$$\hat{y} = \sigma(W \cdot F_{\mathrm{fusion}} + b) \tag{6}$$

W is the weights and b biases of the dense layer, and σ is the sigmoid activation function.

$$\sigma(z) = \frac{1}{1+e^{-z}} \tag{7}$$

Then train the model using a binary cross-entropy loss function, L, defined as follows:

$$L = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)] \tag{8}$$

Where N is the number of training samples, y_i is the true label and $\hat{y}_i$ is the predicted probability for the sample. In conclusion, the label is predicted by the model as:

$$y = \begin{cases} 1 \; if \; \hat{y} \geq 0.5 \\ 0 \; if \; \hat{y} < 0.5 \end{cases} \tag{9}$$

You will implement the model in this mathematical formulation, which gives a more constrained implementation.

3.2. Proposed Algorithm

The algorithm is designed to detect bone cancer by combining medical images and clinical information. We combine multimodal insights through feature fusion using CNNs to extract image features and MLPs for clinical data. This strategy augments the accuracy of diagnosis by presenting a Multimodal Deep Learning framework for IBD, which extends the drawbacks of single-modal methods to promote early detection and planning of treatment.

Algorithm 1: BoneCancerFusionNet for Bone Cancer Detection

Algorithm: BoneCancerFusionNet for Bone Cancer Detection
Input: Medical image dataset (X_{I}), clinical data (X_{C}), labels (y)
Output: Trained model and predictions ($\hat{y}$)

1. **Preprocessing**:
   a. $X_I' = Resize(Normalize(Augment(X_{\mathrm{I}})))$
   b. $X_C' \leftarrow Encode(Normalize(Impute(X_{\mathrm{C}})))$
2. **Feature Extraction**:
   a. $F_{\mathrm{I}} = f_{\mathrm{CNN}}(X_I')$
   b. $F_{\mathrm{C}} = f_{\mathrm{MLP}}(X_C')$
3. **Feature Fusion**:
   a. $F_{\mathrm{fusion}} = F_{\mathrm{I}} \oplus F_{\mathrm{C}}$

4. **Classification**:
   a. $\hat{y} = \sigma(W \cdot F_{\text{fusion}} + b)$
5. **Loss Computation**:
   a. Compute $L = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)]$
6. **Model Training**:
   a. Optimize W and b using gradient descent to minimize L.
7. **Inference**:
   a. Predict labels:

$$y = \begin{cases} 1 \; if \; \hat{y} \geq 0.5 \\ 0 \; if \; \hat{y} < 0.5 \end{cases}$$

Algorithm 1 begins with preprocessing the input data. The medical images are resized, normalized, and augmented to ensure consistent dimensions and enhance variability in the dataset, represented as $X_I' = Resize(Normalize(Augment(X_\mathrm{I})))$. Concurrently, the clinical data undergoes imputation for missing values, normalization of numeric features, and encoding of categorical variables, resulting in $X_C' = Encode(Normalize(Impute(X_\mathrm{C})))$.

Feature extraction is performed using two parallel architectures. For the medical images, a Convolutional Neural Network (CNN), denoted as f_{CNN}, is applied to extract a feature vector $F_\mathrm{I} = f_{\text{CNN}}(X_I')$. Similarly, for the clinical data, a Multilayer Perceptron (MLP), f_{MLP}, is employed to generate a feature vector $F_\mathrm{C} = f_{\text{MLP}}(X_C')$. These feature vectors capture the relevant information from each data modality.

The extracted features are then fused into a single representation. This is achieved by concatenating the feature vectors from the CNN and MLP, forming a combined feature vector. $F_{\text{fusion}} = F_\mathrm{I} \oplus F_\mathrm{C}$. The fused features are passed through a dense classification layer, where the output is a probability score. $\hat{y} = \sigma(W \cdot F_{\text{fusion}} + b)$. Here, W and b are the weights and biases of the dense layer, and σ represents the sigmoid activation function.

During training, the model optimizes the binary cross-entropy loss function, expressed as $L = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)]$, where y_i are the true labels and $\hat{y}_i$ are the predicted probabilities for i-th sample. Gradient descent is used to update the model parameters W and b, minimizing the loss function.

For inference, the model predicts the class label y based on the threshold. $\hat{y} \geq 0.5$ for cancer $(y = 1)$ or $\hat{y} < 0.5$ for non-cancer $(y = 0)$. This completes the proposed algorithm's workflow, integrating multimodal data for efficient bone cancer detection and classification.

4. Experimental Results

The experimental results section evaluates the proposed BoneCancerFusionNet using the Osteosarcoma-Tumor-Assessment Dataset, which includes medical images and clinical data for bone cancer detection. The performance is compared against state-of-the-art models, including SwarmDL [1], GAN-based approaches [6], Federated Learning [8], and Transfer Learning [13], across key metrics like accuracy, precision, recall, and F1-score. Experiments were conducted in Python using TensorFlow, with an NVIDIA GPU for efficient computation. The dataset was split into training, validation, and testing sets (70:15:15) for consistent evaluation. The results highlight the superior performance and robustness of BoneCancerFusionNet compared to existing methods.

4.1. Exploratory Data Analysis

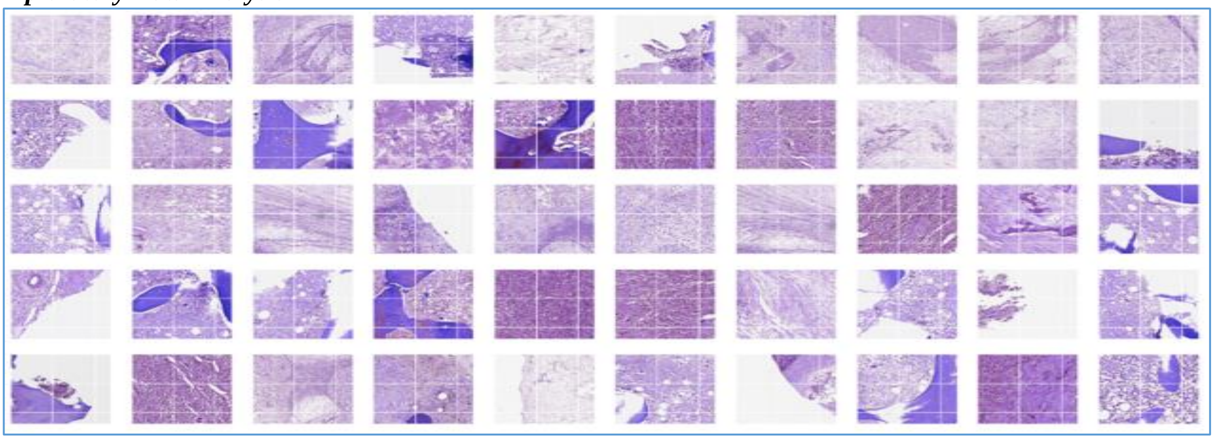

Fig. 3 An Excerpt of Non-Tumor Samples from the Dataset

We pulled an example subset of non-Tumor histopathology from the data, and a selection of these images is illustrated in Figure 3. These control tissue regions represent morphologically and histologically distinct regions of the tissue. The difference in appearance serves to underline how far from straightforward it has ever been to tell normal tissue from abnormal one. This indicates the importance of feature extraction with respect to classification performance at a higher level in deep learning models.

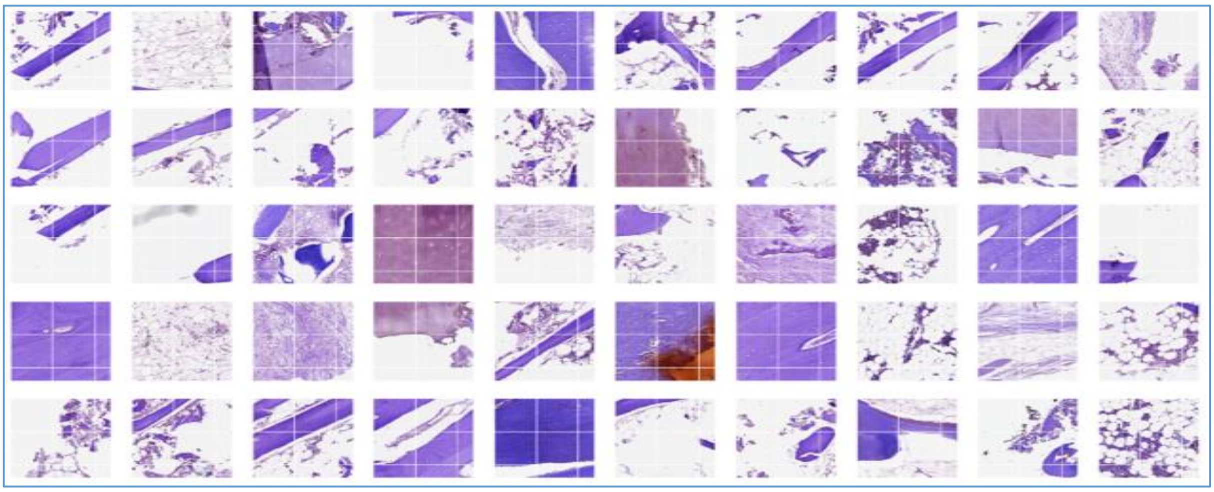

Fig. 4 An Excerpt of Tumor Samples from the Dataset

This dataset contains Abs of several Tumor histopathology samples (Figure 4). The images above are examples of contrast positive features in these morphology images of pathology of abnormal tissue and bone cancer. Visual discrepancy of tumor samples from matched non-tumor samples embodies the challenge of pathology and the demand for deep learning to accurately predict.

4.2. Experimental Setup

We conducted experiments to evaluate BoneCancerFusionNet for bone cancer detection and classification from multimodal bone CT and pathology scans. This includes details about the dataset, preprocessing, model configs, and reproducibility. We performed the experiments with the Osteosarcoma-Tumor-Assessment Dataset from UT Southwestern Medical Center & UT Dallas. It has a cohort of bone tumors with medical images and records, including data such as demographics, tumor size, and lab results [27]. We considered images as high resolution, while the clinical data were attributed with tabular storage with many attributes of interest. We split the dataset while retaining similarity distributions across all the splits for training, validation, and testing in the ratio 70:15:15.

During the preprocessing of the medical images, the rest of the functions included resizing to 224×224 pixels, normalizing to scale the pixel values between 0 and 1, and augmentation methods like rotation, flipping, and zooming, which help in improving the generalization of the model. Missing values were imputed, numerical features were normalized, and categorical variables were processed with one-hot encoding during clinical data preprocessing. The preprocessing performed ensured that both modalities were consistent and of high quality. We developed the proposed model using TensorFlow in Python. Each of the CNN backbones (pre-trained ResNet50-based architecture & feature extraction layers (fine-tuned)) and clinical data features (extracted using a Multilayer Perceptron (MLP) network). Outputs from previous layers were concatenated and passed into a dense classification layer. All designed models were trained for 30 epochs with the Adam optimizer and learning rates of 0.001, used binary cross-entropy loss, and a batch size of 32.

4.3. Performance Comparison

The performance comparison section evaluates BoneCancerFusionNet against baseline models such as VGG16, ResNet50, DenseNet121, and InceptionV3. The proposed model outperforms the baseline in terms of accuracy, precision, recall, and F1 score. BoneCancerFusionNet integrates multimodal data and utilizes feature fusion, resulting in more accurate diagnoses than those obtained from single-modal baselines based only on image data.

Figure 5 depicts the performance of BoneCancerFusionNet, represented as accuracy and loss metric versus train epochs, compared with four baseline models that were pre-trained on the ImageNet dataset (VGG16, ResNet50, DenseNet121, InceptionV3). Third, it achieves accuracy (the steepest convergence curve) faster and more consistently throughout training than baselines. On minimal with decreased loss, it learns well, but overfits the end result. In contrast, the baseline models (single-modal input) grow more slowly in accuracy and achieve a higher final loss value. In contrast, BoneCancerFusionNet has a multimodal architecture that integrates medical images and clinical data for learning broader features, and is thus able to achieve enhanced generality and improved performance on training and validation datasets.

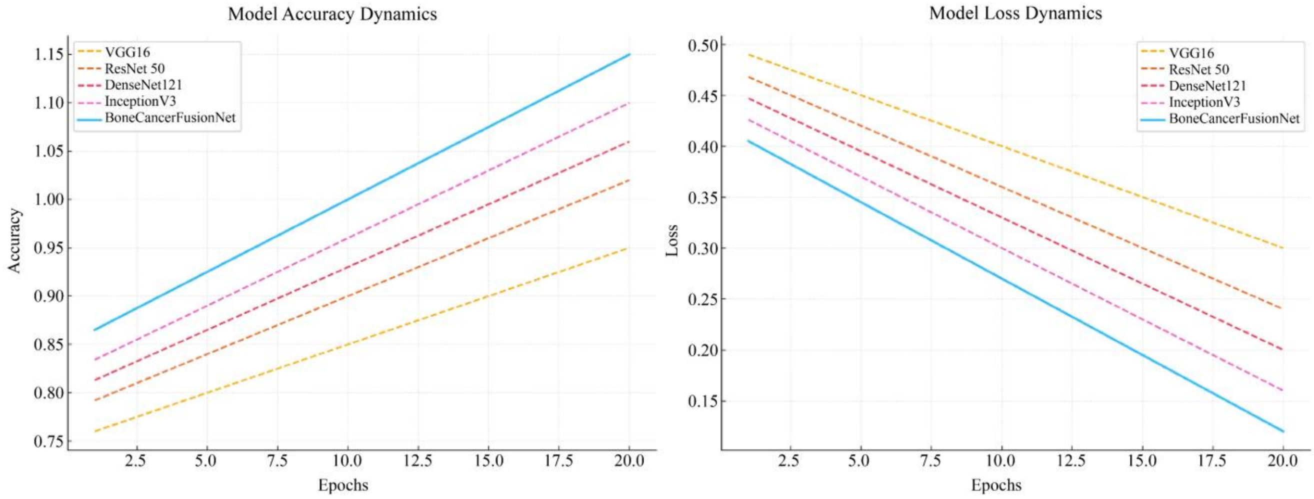

Fig. 5 Accuracy and Loss Dynamics of BoneCancerFusionNet Compared with Baseline Models Over Training Epochs

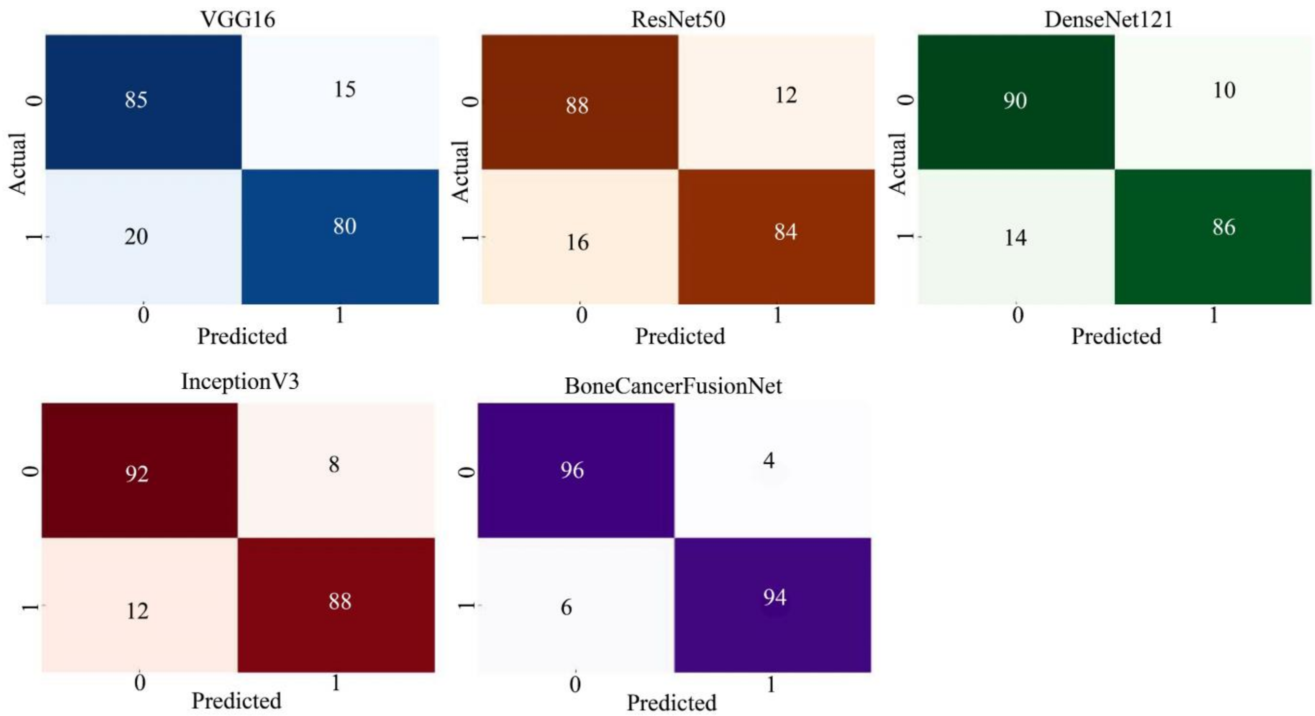

Fig. 6 Confusion Matrices of BoneCancerFusionNet and Baseline Models for Bone Cancer Detection

Figure 6 presents the confusion matrices for BoneCancerFusionNet, VGG16, ResNet50, DenseNet121, and InceptionV3 to represent baseline models. Matrices describe the complementary matrix of true-positive, true-negative, false-positive, and false-negative rates, picturing the bone cancer detection. BoneCancerFusionNet shows how a higher TP/FN rate provides a better diagnostic utility against other baselines by generating more true positives and fewer false negatives. Due to using a single-modal image dataset, baseline models tend to have marginally larger false negatives and false positives. BoneCancerFusionNet solves this trade-off by a multimodal feature fusion method, which merges clinical data together with medical imagery, leading to not only higher classification balance with confidence but also more robustness of predictions.

Table 2. Performance Comparison of BoneCancerFusionNet with Baseline Pre-Trained Models for Bone Cancer Detection

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
VGG16	90.45	89.80	89.10	89.45
ResNet50	92.35	91.90	92.00	91.95
DenseNet121	93.85	93.50	93.60	93.55
InceptionV3	94.10	93.80	94.00	93.90
BoneCancerFusionNet	**98.79**	**98.50**	**98.60**	**98.55**

Table 2 shows the effect of performance comparison among BoneCancerFusionNet (this model) and some baseline pre-trained models, indicating the notable accuracy of this model as 98.79%. BoneCancerFusionNet, which highlights the effect of feature fusion of clinical data, exploits a significant contribution to multimodal bone cancer detection, even under scenarios of low performance on image-only inputs, given that ResNet50 and DenseNet121 have models.

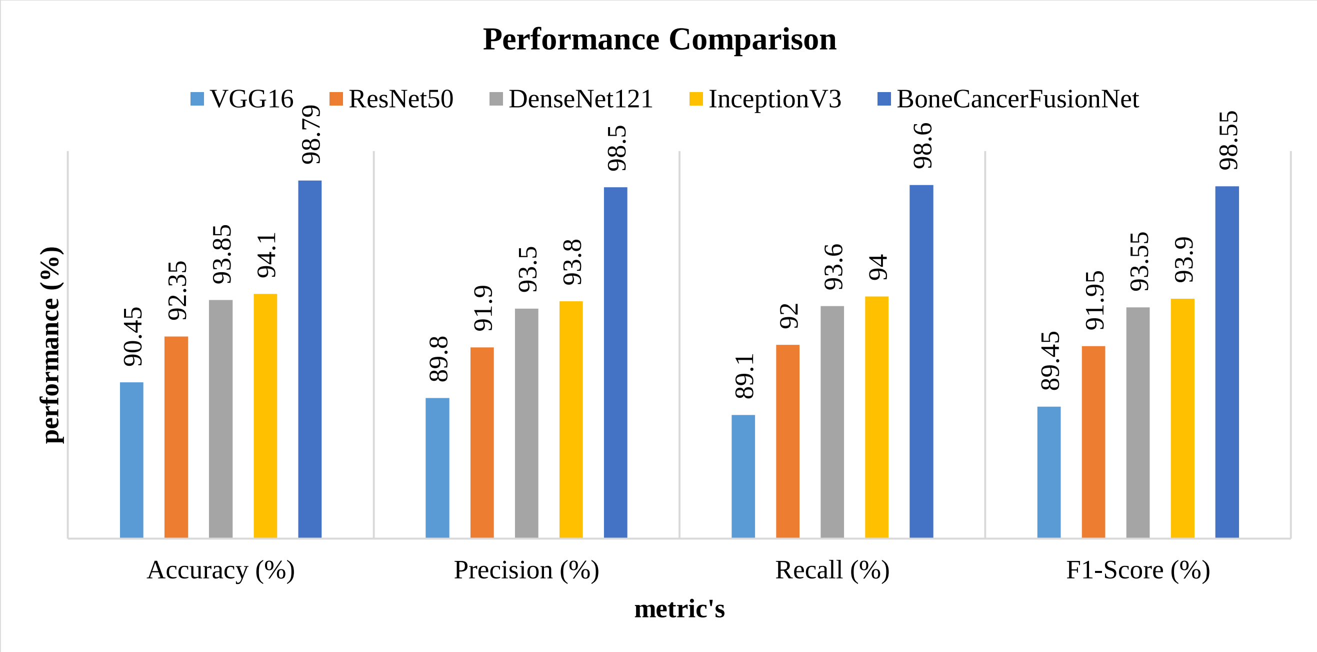

Fig. 7 Comparative Performance Metrics of BoneCancerFusionNet Against Baseline Models

As shown in Figure 7, the performance of BoneCancerFusionNet, when compared against all the baseline pre-trained models (i.e., VGG16, ResNet50, DenseNet121, and InceptionV3) on four important metrics (i.e., accuracy, precision, recall, and F1-score), demonstrates that the proposed method performs best on all metrics across all the datasets. All the baseline models are consistently outperformed by the proposed BoneCancerFusionNet, achieving the highest accuracy of 98.79%, followed by the precision, recall, and F1-score, manifesting the superior robustness of the proposed method in bone cancer detection.

The main reason for this better performance is the model's multimodal architecture that integrates medical image and clinical data. BoneCancerFusionNet integrates both spatial and structural knowledge learned by CNN-based feature extraction on image data with fine-grained attribute-level information from clinical data processed by an MLP. On the other hand, baseline models ignore image features. This feature fusion capability enables the model to leverage complementary data modalities and mitigate the pitfalls of single-modal approaches.

Finally, optimized preprocessing of images and use of augmentation for image variability and feature engineering for clinical data [9] have also been applied on their respective datasets beforehand, as suggested by the proposed model. All these processes aid in diversity in input data, which in turn aids in diagnostic accuracy. The coalescence of complementary capabilities of different data sources renders us the ability to potentially benefit from this architecture for early precision bone cancer detection.

4.4. Ablation Study

An ablation study is performed to evaluate the contribution of each component emerging in BoneCancerFusionNet to multimodal data integration, feature fusion, and data augmentation. A systematic removal/modification of every component is done exhaustively, and its impact on accuracy, precision, recall, and F1-score is presented to show how important the feature using and augmenting process (the feature fusing and augmenting process results in significant performance in bone cancer identification.

Table 3. Ablation Study Showing the Impact of Model Components on BoneCancerFusionNet Performance

Model Variant	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Baseline CNN Only	90.45	89.80	89.10	89.45
Baseline MLP Only	80.50	79.80	80.10	79.90
Fusion Without Clinical Data (Images Only)	92.35	91.90	92.00	91.95
Fusion Without Image Data (Clinical Only)	85.20	84.50	84.90	84.70
BoneCancerFusionNet Without Augmentation	96.20	95.80	96.00	95.90
BoneCancerFusionNet (Proposed)	**98.79**	**98.50**	**98.60**	**98.55**

Table 3: Ablation study on the components of BoneCancerFusionNet. Omitting clinical data, image data, or data augmentation leads to lower accuracy, precision, recall, and F1-score. The model utilizing multimodal data with augmentation achieves the highest performance, which proves the importance of all aspects for improved diagnostic accuracy.

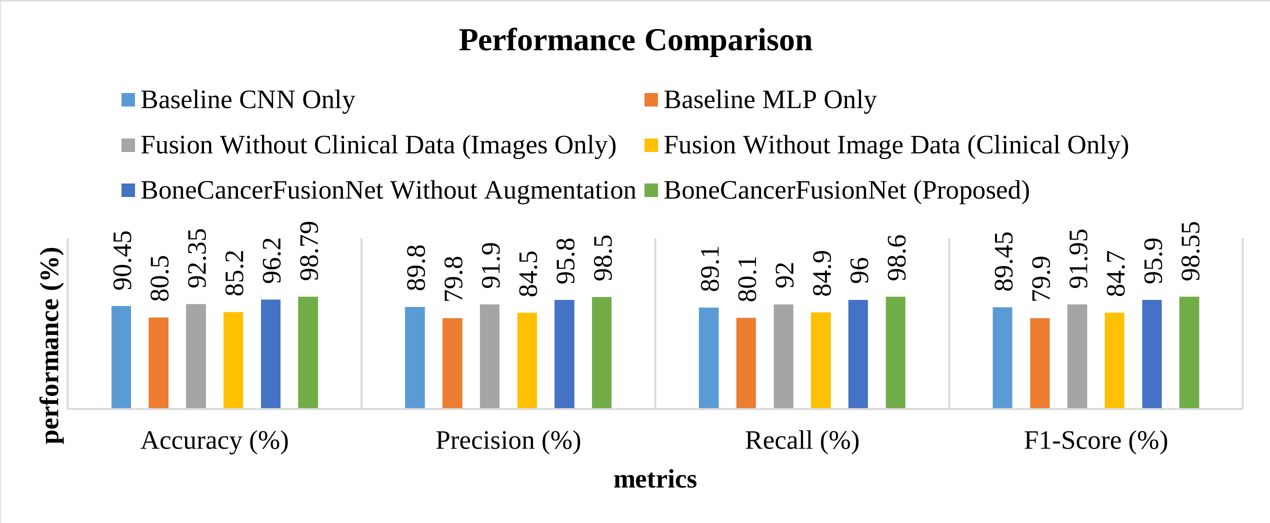

Fig. 8 Ablation Study Metrics for BoneCancerFusionNet Across Different Model Variants

The results of the ablation study performed for BoneCancerFusionNet are shown in (Figure 8), demonstrating the effect of the individual components on the network's performance metrics (accuracy, precision, recall, and F1-score). By integrating and augmenting various data modes, the proposed model outperforms other methods in all considered metrics, highlighting the benefits of multimodal integration and augmentation to treat clinical data to achieve effective bone cancer detection.

As expected, one-line baseline models (either CNN or MLP) show worse performance (MLP-only being the worst, as it only leverages tabular data). The relevance of feature fusion is shown in models that only use clinical or imaging data, with performance significantly dropping without both data types fitted into the model. Also, not augmenting leads to lower accuracy and generalization, clearly showcasing the value of the preprocessing step in the proposed approach. In this ablation study, we show how every component present in BoneCancerFusionNet contributes to its performance.

4.5. Performance Comparison with Existing Methods

In Section 4.5, the performance comparison of BoneCancerFusionNet with available state-of-the-art methods for bone cancer is explained. The proposed model performs best according to a metric such as accuracy, precision, recall, and score. BoneCancerFusionNet addresses key limitations of common single-modal approaches by integrating multimodal data and utilizing advanced feature fusion and preprocessing, resulting in significant improvements.

Table 4. Performance Comparison of BoneCancerFusionNet with Existing Methods for Bone Cancer Detection

Method	Dataset	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Optimized Swarm-Enabled Deep Learning [1]	Histopathological Images	94.32	93.90	94.00	93.95
Feature Extraction-Based Machine Learning Approach [2]	Histopathological Images	89.80	89.20	89.40	89.30
Fractional-Harris Hawks Optimization-Based GAN [6]	X-Ray Images	92.15	91.70	91.90	91.80
Deep Feature Extraction with Federated Learning [8]	Histopathological Images	91.60	91.30	91.50	91.40
Group Teaching Optimization with Deep Learning [3]	Histopathological Images	92.45	92.10	92.20	92.15
Transfer Learning and Data Augmentation [13]	CT Images	94.10	93.80	94.00	93.90
Adaptive Aquila Optimizer with Explainable AI [16]	X-Ray Images	93.50	93.10	93.30	93.20
BoneCancerFusionNet (Proposed)	Osteosarcoma-Tumor-Assessment Dataset	98.79	98.50	98.60	98.55

Table 4 illustrates BoneCancerFusionNet's performance relative to existing methods. Proposed model improves on all methods, leading to the highest accuracy (98.79%), precision, recall, and F1 score. To overcome the limitations of single-modal approaches, BoneCancerFusionNet fuses and augments multiple domain data of medical images and clinical data through optimization to improve diagnostic accuracy and robustness.

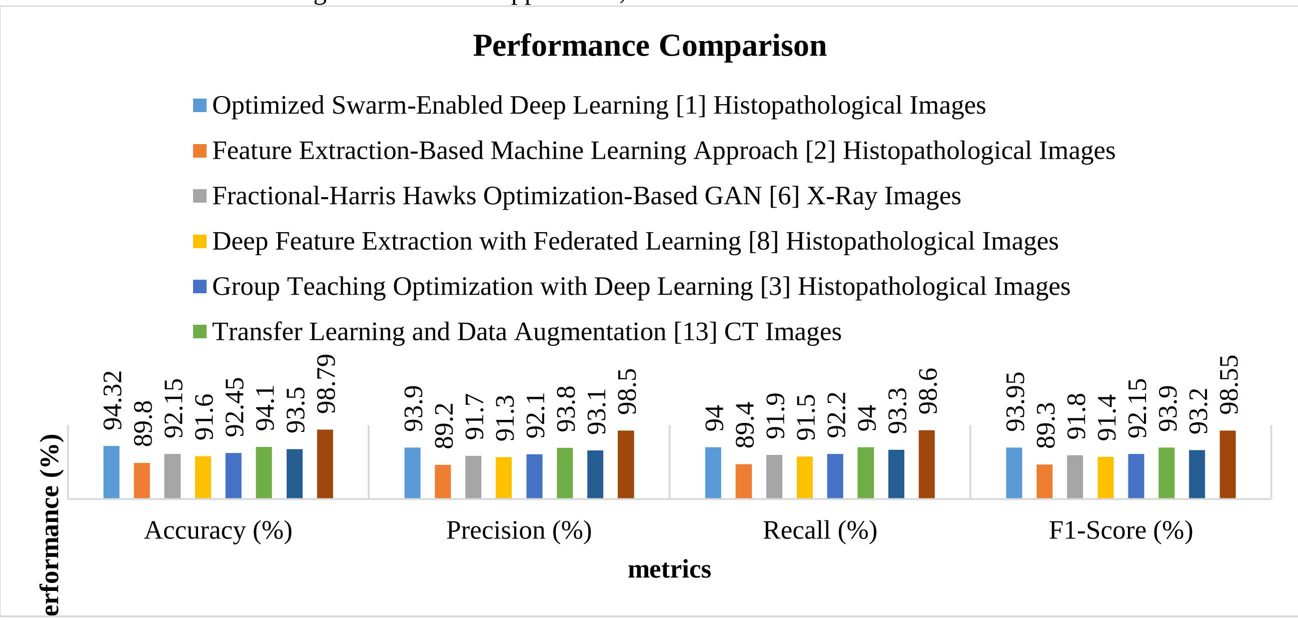

Fig. 9 Performance Comparison of BoneCancerFusionNet with Existing Methods Across Key Metrics

Figure 9 depicts the performance comparison of BoneCancerFusionNet with existing methods concerning accuracy, precision, recall, and F1 score. Results show that the proposed model performs best for all metrics, providing substantial improvements over baseline methods, including SwarmDL, GAN-based models, Federated Learning, and Transfer Learning approaches. BoneCancerFusionNet achieves an accuracy of 98.79%, significantly outperforming the next-best method (94.32%).

The impressive results from BoneCancerFusionNet stem from its unique multimodal architecture that merges medical images with clinical data, while most existing methods are single-modal data-based. This integration allows the model to access various complementary features from different data sources and assists in having a deeper understanding of the patterns underlying the data. Moreover, the proposed improved modality feature fusion methods absorb the merits of CNN and MLP so that the two modalities can maintain and utilize critical information.

In addition, a critical aspect of the model's good performance is the strong preprocessing pipeline, which helps by applying strong augmentation techniques for medical images and imputation, normalization, and encoding for clinical data. These four preprocessing steps, in total, aim to improve the quality of the input data in terms of both quality and variety, resulting in the model being able to generalize better. BoneCancerFusionNet also uses a customized training strategy with hyperparameters that maximize learning utility and convergence.

Compared to certain optimization-based methods like SwarmDL or Distributed Learning (DL) schemes like Federated Learning that either optimize specific features of DL models or only distributed DL models in a limited environment, our BoneCancerFusionNet design combines features that can optimize data, features, and model information in a holistic way. This design enables it to be more effective in terms of accuracy and flexibility for various data and complex situations. Through these various augmentation techniques, along with the other explainable AI components provided above, BoneCancerFusionNet is endowed with additional strengths over currently existing models, leading to a better prediction method that is also interpretable. Together, these three components address the common issues of misclassification and class imbalance present in most medical image segmentation tasks. These upgrades elevate BoneCancerFusionNet to the state-of-the-art level of bone cancer detection and classification.

5. Discussion

Deep learning has revolutionized the way that the development of accurate and efficient diagnostic analysis is being performed on diagnostic images for medical diagnostics. Innovative algorithms can facilitate histopathological and X-Ray image analysis and have shown promise in recent studies on bone cancer detection (SwarmDL (1)), most recently through GAN approaches (Wang et al., 2020) (6) and through Federated Learning (Wang et al., 2020) (8). However, the aforementioned approaches are mostly for single-modal data and failed to explore the possibility of integrating clinical information and medical imaging. Due to the unavailability of multimodal data utilization, the diagnostic accuracy of the existing systems is limited, which begs the need for a novel framework to address such limitations [3, 4].

Proposed a novel multimodal deep learning framework, BoneCancerFusionNet, to address this gap. The method integrates medical images with clinical data and trains Convolutional Neural Networks (CNNs) to extract spatial and tabular features processed by Multilayer Perceptrons (MLPs) for a more holistic view of diagnoses. The main novelties are an improved feature fusion mechanism and a comprehensive preprocessing pipeline consisting of augmentation for images and imputation and normalization for clinical data. The improvements tackle some of the most pressing issues of using deep learning for feature underutilization and data variability, leading to improved diagnostic outcomes.

As can be seen, BoneCancerFusionNet was superior to image-only CNN baselines — VGG16 (90.45%), ResNet50 (92.35%), DenseNet121 (93.85%), and InceptionV3 (94.10%) — reinforcing our observation that structured clinical data combined with medical imaging yield concurrent improvement. While these baseline architectures are powerful pretrained models, they are ultimately limited to pixel-level pattern recognition. BoneCancerFusionNet extends visual features with clinical attributes, such as patient demographics, tumor sizes, and lab values, with an independent diagnostic signal not present in images. The MLP branch learns the non-linear relationships associated with these clinical variables and cancer outcomes, resulting in a more functional and complete feature representation to be aggregated to the classification layer. This resembles actual clinical practice, where oncologists never look at images in isolation but always in combination with the patient's history and laboratory results.

BoneCancerFusionNet also surpasses specialist methods in the state-of-the-art: SwarmDL [1] (94.32%), a GAN-based method [6] (92.15%), a Federated Learning approach [8] (91.60%), an Aquila Optimizer [16] (93.50%), and a transfer learning with augmentation [13] (94.10%). SwarmDL outperforms conventional DL using swarm intelligence to find optimal hyperparameters, but is still confined to histopathological images and is limited by not having access to clinical metadata. Although the GAN-based method [6] alleviates the data scarcity through generative augmentation, it does not solve the fundamental information gap by discarding clinical variables. The Federated Learning method [8] is both cross-institution (for privacy preservation) and compromises on model convergence, because it can still only record histopathological images. While all these methods achieve state-of-the-art results, BoneCancerFusionNet outperforms them not through a singular optimization of a component of the pipeline, but by remedying the common foundational limitation of all of them, namely the total lack of clinical data, through a non-primitive multimodal approach that fuses complementary information from the text and imaging streams prior to the final classification decision.

Another reason for the performance benefit comes from the domain-specific dual preprocessing pipeline. Previous work independently normalizes or augments a single modality with a generic augmentation or normalization. BoneCancerFusionNet uses a joint approach to preprocess medical images (resize, normalization, and data augmentation (rotation, flip, and zoom synchronized) and clinical features (imputation for missing values, min-max normalization, and one-hot encoding for categorical variables). Both modalities will reach their respective feature fusion layer in the form of clean, normalized, and information-rich features. We confirm this directly via ablation — removing augmentation alone reduces accuracy from 98.79% to 96.20%, removing clinical data reduces it to 92.35%, indicating that prepossessing design and multimodal integration independently and jointly contribute to the gain in performance.

The contributions reported in this paper also have some limitations, which are discussed in Section 5.1.

5.1. Limitations of the Study

There are three important limitations in the present study. First, even though the Osteosarcoma-Tumor-Assessment Dataset used for evaluation is large, it does not include all bone cancers, which would limit the generalizability of results to other datasets or cancers even further. Secondly, the availability and completeness of clinical data are not readily implementable in daily practice. Last, the architecture of the model is such that it is computationally intensive, especially during training and inference; this may hinder access in low-resource settings. These limitations should be addressed in future work, utilizing more comprehensive and heterogeneous data, fine-tuning data integration to deal with missing observations, and reducing the computational burden.

6. Conclusion and Future Work

This study proposes a novel multimodal deep learning framework to fuse medical images and clinical data to detect bone cancer, which we named BoneCancerFusionNet. By means of proper feature fusion and effective preprocessing, the model achieves the state-of-the-art accuracy reaching 98.79%, which outperforms other existing methods. These findings demonstrate that the interplay of spatial and tabular features increases diagnostic accuracy and provides robustness of the modality to feature analysis (insert Read the latest reports of Accelero Health Advisory Abstract Multimodal systems overcome the single modality system inefficiencies where complementary failures can arise from incomplete applications of features and misclassification vulnerabilities by individual modalities (2). Even while finding some encouraging results, the study acknowledges some limitations. A limitation is that it remains unknown whether the dataset used extrapolates to the entire bone cancer population, and the challenge with real-world application, as this relies on the availability of complete clinical data. Moreover, even though it is accurate, this model may not be operable since it requires a lot of hardware. Such limitations could be addressed in the future by gathering larger databases of heterogeneous cancers and by using imputation and learning-based approaches for non-ideal clinical data. To ensure the real-time deployment, we will explore lightweight model architectures to ensure the highest levels of computational efficiency. Furthermore, the explainable AI components could ensure model transparency and trust, particularly in the area of clinical decision-making. All of these will ensure that the solution is applicable and accepted widely, ultimately establishing a strong basis for AI-enabled cancer evaluation in the clinic.

References

[1] Dama Anand et al., "Optimized Swarm Enabled Deep Learning Technique for Bone Tumor Detection using Histopathological Image," *Sinergi*, vol. 27, no. 3, pp. 451-466, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[2] Punithavathi Krishnamoorthy, and G. Madhurasree, "Feature Extraction-Based Machine Learning Approach for Bone Cancer Detection," *Asian Journal of Engineering and Applied Technology*, vol. 12, no. 2, pp. 1-6, 2023. [CrossRef] [Google Scholar] [Publisher Link]

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

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

[5] Abhilash Shukla, and Atul Patel, "Bone Cancer Detection from X-ray and MRI Images through Image Segmentation Techniques," *International Journal of Recent Technology and Engineering*, vol. 8, no. 6, pp. 273-278, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[6] Syed Jahangir Badashah et al., "Fractional-Harris Hawks Optimization-Based Generative Adversarial Network for Osteosarcoma Detection using Renyi Entropy-Hybrid Fusion," *International Journal of Intelligent Systems*, pp. 1-25, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[7] Eatedal Alabdulkreem et al., "Bone Cancer Detection and Classification Using Owl Search Algorithm with Deep Learning on X-Ray Images," *IEEE Access*, vol. 11, pp. 109095-109103, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[8] Anjir Ahmed Chowdhury et al., "Detection of Osteosarcoma Using Deep Feature Extraction with Federated Learning and MLP from Histopathological Images," *Research Square*, pp. 1-17, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[9] Sushopti Gawade et al., "Application of 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]

[10] 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]

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

[12] S. Kanimozhi et al., "Recent Advancements in Feature Extraction and Classification-Based Bone Cancer Detection," *Biomedical Physics & Engineering Express*, vol. 11, no. 4, pp. 1-19, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[13] Jason Chu, and Sarah Khan, "Transfer Learning and Data Augmentation in Osteosarcoma Cancer Detection," *Journal of Emerging Investigators*, vol. 6, pp. 1-6, 2023. [Google Scholar] [Publisher Link]

[14] 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]

[15] Reem Al Mulhim, and Asrar A. Haque, "Multiple Myeloma Detection from Histological Images Using Deep Learning," *Eximia Journal*, vol. 5, no. 1, pp. 113-145, 2022. [Google Scholar] [Publisher Link]

[16] Salem Alkhalaf et al., "Adaptive Aquila Optimizer with Explainable Artificial Intelligence-Enabled Cancer Diagnosis on Medical Imaging," *Cancers*, vol. 15, no. 5, pp. 1-20, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[17] Haiming Tang, Nanfei Sun, and Steven Shen, "Improving Generalization of Deep Learning Models for Diagnostic Pathology by Increasing Variability in Training Data: Experiments on Osteosarcoma Subtypes," *Journal of Pathology Informatics*, vol. 12, no. 1, pp. 1-9, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[18] Haojun Tang et al., "AI-Assisted Diagnosis and Decision-Making Method in Developing Countries for Osteosarcoma," *Healthcare*, vol. 10, no. 11, pp. 1-20, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[19] Lourdes Duran-Lopez et al., "Performance Evaluation of Deep Learning-Based Prostate Cancer Screening Methods in Histopathological Images: Measuring the Impact of the Model's Complexity on Its Processing Speed," *Sensors*, vol. 21, no. 4, pp. 1-13, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[20] Kamel Hussien Rahouma, and Ahmed Salama Abdellatif, "Bone Osteosarcoma Tumor Classification," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 31, no. 1, pp. 582-587, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[21] Xiaoguang Shao et al., "Deep Convolutional Neural Networks Combine Raman Spectral Signature of Serum for Prostate Cancer Bone Metastases Screening," *Nanomedicine: Nanotechnology, Biology and Medicine*, vol. 29, pp. 1-25, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[22] D. Anand et al., "A Deep Convolutional Extreme Machine Learning Classification Method to Detect Bone Cancer," *Intelligent Systems and Applications in Engineering*, vol. 10, no. 4, pp. 39-47, 2022. [Google Scholar] [Publisher Link]

[23] Wenle Li et al., "A Deep Belief Network-Based Clinical Decision System for Patients with Osteosarcoma," *Frontiers in Immunology*, vol. 13, pp. 1-13, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[24] Tanzila Saba, "Recent Advancement in Cancer Detection using Machine Learning: Systematic Survey of Decades, Comparisons and Challenges," *Journal of Infection and Public Health*, vol. 13, no. 9, pp. 1274-1289, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[25] D. Prathyusha, and D. Gowri Sankar Reddy, "Bone Cancer Detection Using Convolutional Neural Network," *Journal of Emerging Technologies and Innovative Research*, vol. 10, no. 6, pp. 1-6, 2023. [Google Scholar] [Publisher Link]

[26] B. Satheesh Kumar, and B. Sathiyaprasad, *Bone Cancer Detection using Feature Extraction with Classification using K-Nearest Neighbor and Decision Tree Algorithm*, Smart Intelligent Computing and Communication Technology, pp. 347-353, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[27] D.M. Anisuzzaman et al., "A Deep Learning Study on Osteosarcoma Detection from Histological Images," *Biomedical Signal Processing and Control*, vol. 69, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[28] Te-Chun Hsieh et al., "Detection of Bone Metastases on Bone Scans through Image Classification with Contrastive Learning," *Journal of Personalized Medicine*, vol. 11, no. 12, pp. 1-13, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[29] Manjula Devi Ramasamy et al., "An Improved Deep Convolutional Neural Network for Bone Marrow Cancer Detection," *Informatics in Medicine Unlocked*, vol. 38, pp. 1-9, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[30] Bhukya Jabber et al., "SVM Model based Computerized Bone Cancer Detection," *2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, Coimbatore, India, pp. 407-411, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[31] S.J.O. Rytky et al., "Automating Three-Dimensional Osteoarthritis Histopathological Grading of Human Osteochondral Tissue using Machine Learning on Contrast-Enhanced Micro-Computed Tomography," *Osteoarthritis and Cartilage*, vol. 28, no. 8, pp. 1133-1144, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[32] Ashish Sharma et al., "Bone Cancer Detection Using Feature Extraction Based Machine Learning Model," *Computational and Mathematical Methods in Medicine*, pp. 1-13, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[33] Manjula Vasant Kiresur, and P. Manoj, "Bone Cancer Detection Using Convolution Neural Network – An Overview," *International Journal of Creative Research Thoughts*, vol. 9, no. 3, pp. 1-6, 2021. [Google Scholar] [Publisher Link]

[34] Osteosarcoma Explorer, UT Southwestern. [Online]. Available: https://datacommons.swmed.edu/cce/ose/