OsteoCancerNet: An Efficient and Fast Bone Cancer Diagnostic Model Combining EfficientNet B4 and SVM with RBF Kernel for X-ray Image Analysis

Nashaat M. Hussain Hassan
nmh01@fayoum.edu.eg

Badr University in Cairo (BUC)

Ahmed S. Bayoumy
Badr University in Cairo (BUC)

Mohamed Hassan
Giza Institute for Higher Education and Technology

Article

Keywords: OsteoCancerNet, Computer-assisted diagnosis, Bone Cancer Diogenes, EfficientNet B4 Model, SVM Model, X-ray Image Analysis

Posted Date: August 20th, 2025

DOI: https://doi.org/10.21203/rs.3.rs-6530365/v1

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Scientific Reports on March 3rd, 2026. See the published version at https://doi.org/10.1038/s41598-026-38801-3.

OsteoCancerNet: An Efficient and Fast Bone Cancer Diagnostic Model Combining EfficientNet B4 and SVM with RBF Kernel for X-ray Image Analysis

*Nashaat M. Hussain Hassan[1, 2], Ahmed S. Bayoumy[1, 3], Mohamed Hassan M. Mahmoud[4]

[1]Faculty of Engineering and Technology, Badr University in Cairo (BUC), Badr 11899, Egypt
[2]Electronics and Communication Engineering Dept, Fayoum University, Fayoum 63514, Egypt
Corresponding author :nmh01@fayoum.edu.eg
[3]Physics and Engineering Mathematics Department, Faculty of Engineering, Kafrelsheikh University, Kafrelsheikh, 33516,
[4]Giza Institute for Higher Education and Technology, Giza, Egypt

Abstract

Bone cancer diagnosis is imperative for diagnosing and treating many forms of primary and metastatic bone cancers early. Traditional imaging techniques, such as CT, MRI, and X-ray scans, are effective but typically involve manual interpretation, which is laborious and prone to human mistake. With increased accuracy, dependability, and efficiency over conventional techniques, automated bone cancer diagnosis systems have been made possible by recent developments in machine learning (ML) and deep learning (DL). Even though a large portion of the recent literature has made significant strides in the diagnosis of bone cancer using deep learning techniques, many of these methods have numerous drawbacks, including computational complexity, overfitting in certain situations, and a lack of reliable databases. The objective of this research is to develop a method that can diagnose bone cancer as quickly, efficiently, and affordably as feasible. This method's main contribution is the combination of EfficientNetB4 and SVM algorithms, which improves accuracy, speed, and accessibility while making use of large datasets and reliable assessment measures. Combining EfficientNetB4 and SVM is crucial for diagnosing bone cancer because it harnesses the EfficientNetB4 technique's amazing capacity to extract useful features both quantitatively and qualitatively, as well as the SVM's significant advantage when it comes to binary separation. After a thorough analysis of numerous research, it was determined to combine these methods as they are distinguished by their great efficiency and simplicity in job implementation, especially in medical environments where accuracy and interpretability are of utmost importance The success of the suggested methods was shown by experiments on a large dataset (which contains 35244 X-ray images), which produced 98% precision, 97.47% recall, 98% accuracy, and 98% F1-score. The suggested method's better performance and computational economy are highlighted by comparison with machine learning, deep learning, and transfer learning techniques. Furthermore, the suggested system encounters a quick inference time of 41 ms, which qualifies it for clinical real-time applications. This study offers a viable strategy for early detection and better treatment outcomes by demonstrating the potential of integrating deep learning and traditional machine learning approaches for better bone cancer diagnosis.

Keywords: OsteoCancerNet; Computer-assisted diagnosis; Bone Cancer Diogenes; EfficientNet B4 Model; SVM Model; X-ray Image Analysis

1. Introduction

Though relatively uncommon cancer [1], bone cancer [2] presents substantial challenges to healthcare because of its profound impact on patients' survival rates and quality of life. It can be classified as primary bone cancer, which originates in the bone tissue directly, or metastatic bone cancer, which spreads from other tissue. The most effective treatment plan usually consists of chemotherapy, radiation therapy, surgery, or a combination of these. However, the subtle nature of early-stage bone cancer and the similar appearance of benign and malignant bone lesions on medical imaging can make it difficult to diagnose.

The conventional diagnostic methods, like computed tomography (CT) scans, magnetic resonance imaging (MRI), and X-rays, are the main tools used to evaluate bone abnormalities [3]. These methods provide detailed insights into the pathological and structural changes in bones, but despite their usefulness, the interpretation of these images is subjective and subject to variability because it relies heavily on radiologists' expertise. The manual analysis takes a lot of time and is prone to mistakes because of fatigue, limited experience with rare cases, or overlapping features between benign and malignant conditions, which emphasizes the need for better and more effective diagnostic methods.

Artificial intelligence (AI) transcendence [4,5], especially deep learning (DL) and, machine learning (ML) [6–9], has revolutionized medical image analysis recently. Deep learning models, like Convolutional Neural Networks (CNNs) [10–12], have been shown to be remarkably successful in a variety of diagnostic tasks, from disease detection to segmentation and classification. CNNs are particularly useful in addressing the complexity of bone cancer diagnosis because they are able to automatically learn hierarchical and complex features from raw medical images, surpassing traditional methods that rely on handcrafted feature extraction. Despite the fact that a lot of recently published work has advanced the recognition of bone cancer through methods utilizing deep learning, most of these approaches have a number of difficulties, such as high computational costs, overfitting in specific circumstances, and a lack of credible datasets. Additionally, a major portion of this work does not adequately verify the effectiveness of the methods they suggest since it does not compare their findings with those of several comparable techniques, including machine learning, deep learning, and transfer learning[13]. Also, the outcomes of the proposed methodologies for many of these literatures have not been reviewed using a number of assessment measures to assure the reliability of the results produced. Therefore, this study took into account that all these drawbacks are addressed by relying on a large database, then preparing the images well during pre-processing by resizing, improving the degree of image contrast, increasing the number of images and their display diversity through augmentation, and then normalization. Combining classical machine learning techniques with deep learning techniques by using the technique EfficientNet B4 to extract the distinctive features of X-ray images (this technique specifically because of its high efficiency in extracting the distinctive features of medical images) and then using the SVM technique to separate between normal image and abnormal ones. Also, to ensure the credibility of the results obtained, the results of the proposed technique (*Osteo*) were compared with the results of a number of the most famous machine learning, deep learning and transfer learning techniques[14]. The results were also reviewed using a large number of different evaluation mechanisms. The focus was also on reviewing the results of the computation times of the proposed technique (*Osteo*) and making the necessary treatments so that the time taken becomes appropriate for real clinical medicine experiments.

In the study presented in this research paper, the choice fell on using the SVM [15–17] technique as a classifier, as it showed great superiority in performance accuracy in binary classification tasks, in addition to its high processing speed and low implementation cost. It was also decided to use the EfficientNet B4 [18,19] technique to extract the features of X-ray images, as it is distinguished by its combination of its very high efficiency compared to its counterparts from other deep learning techniques in dealing with medical images, especially X-ray images (as performance accuracy for extracting features) and the short time it takes to perform the necessary calculations compared to other deep learning techniques. A hybrid technique that capitalizes on the advantages of both approaches is provided by combining SVM for classification with deep learning for feature extraction, resulting in a reliable and accurate diagnostic system.

This fusion's efficiency between machine and deep learning (which developed the suggested technique) allows it to outperform many traditional convolutional designs without requiring as much processing power. For the classification to become more structured with the sophisticated classification process, this model is utilized to extract picture features in depth. The next section is the Key Contributions of This Study: Utilizing the advantages of the EfficientNet B4 transfer learning framework, this study calculates and extracts a great volume of extensive, hierarchical features from X-ray bone images to ensure the entire input data is considered.

EfficientNet-B4 for feature extraction and SVM with RBF kernel for high-precision normalization of normal vs. abnormal bone image classification. It shows a viable and suitable deployment system that bespeaks high performance metrics; it reached an accuracy of 98%, a recall of 97.47%, and an inference time of 41 ms, making it competent for real-time diagnostics.

The proposed approach (*Osteo*) improves bone cancer diagnostics by using systematic feature extraction to analyze across disease pathologies, reducing human interpretation and errors. EfficientNet B4 + SVM focuses on scalability and resilience, even with limited training samples. The system's short inference time enables real-time decision-making, which can significantly improve patient well-being. Although the study's technological contributions are noteworthy, we believe its greatest promise lies in practical application. Early detection of bone cancer often leads to improved treatment outcomes and lower fatality rates. Integrating AI-powered diagnostic solutions, such as the proposed technique (*Osteo*), can improve accuracy and efficiency for healthcare providers and improve access to resource-poor areas with limited trained radiologists. The proposed hybrid model is applicable to various forms of medical imaging.

To summarize, this work constitutes a promising contribution towards automating bone cancer diagnosis. The proposed system (*Osteo*) focuses on transforming the methods of detecting bone cancer by leveraging the best practices feature extraction and strong classification methods to achieve better clinical output and enhance patient treatment.

This work is structured such that: Section 2 provides a review of earlier research on the identification and categorization of brain tumors. Section 3 outlines the recommended method for identifying and classifying apple defects. Following that, the test findings of the proposed method (*Osteo*) are carefully contrasted with those of similar approaches and assessed. At the end come the findings and suggestions for more research.

2. Literature Review

A range of diagnostic techniques such as X-rays, MRI, CT scans and biopsy [20–22] have been used to detect bone cancer, with an important contribution of medical image processing techniques [23,24]. Lately, machine learning (ML) and deep learning (DL) methods such as the use of Support Vector Machines (SVM), Convolutional Neural Networks (CNNs) and hybrid models [25,26] have demonstrated tremendous potential to automate the process of the bone cancer diagnosis. In this work, a review of the literature on bone cancer diagnosis via various approaches is provided.

Early diagnosis is key to finding the best treatment path, which usually includes surgery, chemotherapy and radiation therapy. X-rays, MRI, and CT scan imaging can aid both in the detection of bone tumors, as well as in evaluating their size, location, and extent, which are critical in helping to determine the appropriate treatment approach. Pulmonary imaging in the form of CT images also generates enormous amounts of raw data, manually dissected by experts who may take hours per exam, but are highly prone to human errors, thus facilitating the shifting towards AI-assisted systems and tools to create more reliable diagnosis options.

Traditionally, bone cancer diagnosis has relied on the use of imaging modalities such as X-ray, MRI, and CT 25. 3. X-ray Imaging − X-rays are generally used to screen bone lesions. Such X-rays could reveal morphological alterations (eg, bone destruction, abnormal bone configuration, osteolytic lesions) in bones that are potentially caused by a tumor or cancer. However, they have great difficulty in detecting smaller tumors or involvement of soft tissues. MRI: MRI is better for soft tissue than an X-ray and is commonly used for detailed studies of a bone tumor. It generates high-resolution pictures and is especially effective in delineating tumor margins, detecting bone marrow infiltration, and soft tissue extension.

When examining local bone tumor staging, the MRI is deemed as accurate and arguably the best method. We turn to CT Scans because of their benefit in making 3D imagery of bone structures and in the search for metastatic cojns. The negative aspect of CT scans is that when compared to MRIs, they do have decreased levels of contrast in regards to soft tissues. However, the existing imaging techniques are enhanced considerably but this still directly correlates to the manual evaluations with respect to the images arising much subjectivity which in turn can affect the diagnostic results. ML algorithms have gained significant traction in the last few years in the field of targeting detection and classification of diseases, in bone cancer automation as well. The technology extracts unique attributes from a provided sample like images and accurately identifies the existence of a bone tumour. Support vector machines are now gaining considerable popularity as more and more applications in the field of image processing arise. The goal of an SVM is to construct the optimal hyperplane to separate the data into different classes. In the case of diagnosing bone cancer, SVM can be trained to determine whether an X-ray MRI or CT scan images show normal or abnormal (bone cancer). Ashish et al. (2021) [27] employed SVM for bone tumor detection from X-ray images. Using the HOG feature, their method distinguished between normal and pathological bone pictures with an accuracy of 92.5%; without the HOG feature, the accuracy was 87.5. The study relied on the use of a database consisting of only 105 X-ray bone images. Bhukya et al. (2020) [28] applied SVM based M3 filtered and Fuzzy C-Means (FCM) segmentation method to MRI images for classifying bone tumors, demonstrating high sensitivity and specificity, particularly in identifying osteosarcoma. Their approach achieved an accuracy of about 92%. This accuracy is achieved using a dataset consisting of only 109 X-ray bone images. Salvatore et al. (2024) [29] employed SVM for bone tumor detection on frontal view X-rays, MRI and CT. Their approach achieved an accuracy of about 82%. This accuracy is achieved using a dataset consisting ofonly1120 images.

Deep learning techniques, specifically CNNs, have revolutionized medical image analysis by automatically learning hierarchical features from raw image data without requiring manual feature extraction. CNNs are particularly effective in medical imaging tasks due to their ability to capture spatial relationships in image data. Liwen Song et al. (2024) [30] used the Primary Bone Tumor Classification Transformer Network (PBTC-TransNet) model to classify primary bone tumors from multimodal images in X-ray, CT, and MRI, achieving an impressive accuracy of 95%. This accuracy is achieved using a dataset consisting of only1305 images. Kanimozhi Sampath et al. (2024) [11] , this paper presents an analytical study to evaluate the accuracy of several deep learning techniques in diagnosing bone cancer. This study relied on the use of a database consisting of 1141 CT bone images. This paper presents an analytical study to evaluate the accuracy of several deep learning techniques (ten different deep learning techniques) in diagnosing bone cancer. The accuracy of the techniques used in this study ranged from 59% to 84%.

To enhance diagnostic performance, hybrid models that blend deep learning techniques (like CNNs) with conventional machine learning techniques (like SVM) have been investigated. These models leverage the strengths of both techniques, where deep learning is used for feature extraction and machine learning algorithms like SVM are used for classification. G. Suganeshwari et al. (2023) [31] proposed a hybrid approach combining CNN feature extraction with SVM classification for bone cancer detection in X-ray images. By uniting the decisive boundary development of SVMs with the strong feature learning of CNNs, this model enhanced accuracy and robustness. The research was based on a database of 100 X-ray bone images. Their model had an accuracy ratio of about 93%.

Performance diagnostics has gained a lot of attention due to the use of hybrid techniques and ensemble models which involve the combination of multiple algorithms. Bias and variance error computation suffers a decrease, while classification error rate occurs an increase as a result of ensemble learning. To further enhance bone cancer classification model accuracy, RF, AdaBoost, and Gradient Boosting are integrated into SVM and CNN. This system employed a CNN model to carry out the transfer learning algorithm and extract features from a pre-processed image. An SVM model was also incorporated to build the cancer detection model based on the extracted features. The database used for this study consisted of a measly 100 X-ray images. They obtained an accuracy of around 93%. Another great method when tackling data sufficiency is transfer learning, which utilizes neural networks ImageNet for residual model training and fine-tuning on bone cancer detection. Xia et al. used transfer learning on X- ray images of bone tumor diagnostics using a Res Net model and concluded that transfer learning with a small dataset is very efficient and accurate.

These advancements may sound fruitful, but skeleton cancer diagnosis automation is not without its challenges: Imbalanced dataset (problem with the image datasets due to for example, normal images being more present than the tumor images) can cause biased predictions. Oversampling, under-sampling, and cost-sensitive learning are widely used methods to tackle this problem. Lesions caused by bone cancer tend to differ slightly in shape, texture, and size.

Proper feature extraction or deep learning-based feature learning is essential for distinguishing between benign and malignant lesions. While deep learning models like CNNs offer high accuracy, they are often seen as "black-box" models, making it difficult for medical professionals to interpret how the model reached a particular diagnosis. Incorporating explainability into these models is an ongoing challenge.

While recent studies in bone cancer detection using deep learning and machine learning have shown promising results, several challenges and limitations still hinder their widespread implementation in clinical practice. Below are the key issues identified in current research: One of the most significant challenges in training deep learning models for medical image analysis, including bone cancer detection, is the limited availability of high-quality annotated datasets. Bone cancer, being relatively rare, often results in small and imbalanced datasets. This issue is particularly problematic for training machine learning models, as they tend to overfit to the majority class (normal images) and struggle to generalize well to the minority class (e.g., abnormal images), techniques like data augmentation and transfer learning are often employed, but they may not always fully compensate for the lack of sufficient data. Moreover, small datasets can also lead to issues such as inaccurate performance metrics, as models trained on limited data may fail to perform adequately on unseen, real-world data. Many studies lack standardization in the evaluation metrics used to judge model performance. Although metrics like accuracy, precision, recall, and F1-score are standard, they are not always the best choices for the problem at hand, and their use can dramatically impact how effective a model seems. Moreover, when models are applied to datasets that are public and well-known, their apparent performance is not always a good reflection of how they will do on the diverse, real-world datasets we see in medicine. This discrepancy can lead to overestimation of model capabilities. Overfitting is a common issue in deep learning, especially when the model is trained on a small and/or unbalanced dataset. When a model is overfitted, it performs beautifully on the training data but does not generalize well to new, unseen data. This is a huge problem in the medical domain when real-world data can look very different from the training data. Regularization techniques, cross-validation, and augmentation methods are commonly used to mitigate overfitting. These methods do not seem to eliminate the problem. Moreover, a model that performs well on a clean, controlled dataset may not be robust enough to handle the more diverse and noisier data we see in clinical medicine. Various studies have discussed the potential gains over traditional ML methods of combining EfficientNet-B4 with Support Vector Machines (SVMs) in the bone cancer classification task. EfficientNet, specifically its version B4, is a well-known deep learning model that excels as a feature extractor in the tokenization process, along with several other fields of computational science. However unlike older CNNs, which tend to be compute-hungry, EfficientNet is fine-tuned to successfully vary its performance, given its utilization of minimal computational resources. Such that it is beneficial for medical image classification where the extracted features quality is vital. EfficientNet-B4 can scale up any of these two factors that can better interpret the complex characteristics in X-ray and CT images, and can then be used for detecting subtle atypical signs of bone cancer that are hard to capture by the conspicuous-based methods. Table 1 outlines the benefits and drawbacks of the most significant studies published on this issue in the previous four years (2021-2024). Refer to Table 1 for more information.

Table 1. Benefits and drawbacks of the most significant research that has been published in the last the last four year

References	Publication Year	Technique	Benefits	drawbacks
[27]	2021	Bone Cancer Detection Using Feature Extraction Based Machine Learning Model	1- The study presented a proposal to separate cancerous bone from healthy bone. 2- The study relied on the use of SVM and RF algorithms. 3- The proposal presented in this study achieved an efficiency of 92% for SVM and 77% for RF.	1- The study relied on the use of a very small database that contained only 80 X-ray images. 2-The study did not indicate anything about the computation times of the proposed technique. 3- Accuracy can be improved 4- The study did not compare the proposal's outcomes to those of similar techniques (only SVM and RF algorithm results were compared).
[32]	2023	Bone Cancer Detection and Classification Using Owl Search Algorithm with Deep Learning on X-Ray Images	1- The model uses Inception v3 as a model that has been trained to extract features from X-ray images, eliminating the need for human segmentation. Additionally, the OSA serves as a hyperparameter optimizer to improve the effectiveness of the Inception v3 approach. Finally, the LSTM technique is employed to detect bone cancer. 2- The proposal presented by this	1- The study relied on the use of a very small database that contained only 200 X-ray images. 2- Although the author states that the study offers a proposal that reduces diagnosis time, the paper does not provide any data on inference time. 3- The authors compared their results to a few old ML and DL techniques, however it would have been more relevant to compare their results to current and

			study achieved a performance efficiency of 95%	heavyweight methodologies.
[33]	2024	Automated Bone Cancer Detection Using Deep Learning on X-Ray Images	1- This approach generates feature vectors using the SqueezeNet model and selects hyperparameters using the GSO algorithm. The extracted features can be classified using improved cuckoo search and a long short-term memory model. 2- The proposal presented by this study achieved a test performance efficiency of 94%	1- The authors relied on a very small database of only 200 X-ray images, which reduces the reliability of the results obtained. 2- The authors did not mention what measures they took to reduce the possibility of overfitting of the results obtained. 3- The authors did not mention any data of computation times, to consider the possibility of benefiting from the proposal in real experiments or not. 4- The authors did not compare their results with the results of modern techniques of the related techniques.
[31]	2023	DTBV: A Deep Transfer-Based Bone Cancer Diagnosis System Using VGG16 Feature Extraction	1- This study presents a technique for detecting and classifying bone cancer. This technique is based on the combination of VGG16 feature extraction techniques and SVM for bone classification. 2- This technique achieved a performance efficiency of 93.9%.	1- This study was used to train and then test the efficiency of the presented technique on a database of only 100 X-ray bone images. 2- The study provided no information on the computational times of the proposed technique. 3. The accuracy can be increased 4-The authors compared their findings to a few ancient ML and DL techniques, although it would have been more appropriate to compare them to current and heavyweight procedures.

[34]	2023	A fusion of VGG-16 and ViT models for improving bone tumor classification in computed tomography	1- The study presents a technique based on the combination of VGG-16 and ViT models to classify bone tumors into three categories (benign tumor - cancer - no tumor). 2- This technique achieved a high efficiency of 97.6%.	1- This technique achieved a high efficiency of 97.6% This technique relied on a small data base of only 786 CT images. 2- The authors of this work did not provide any data related to the inference time, so that we can know the suitability of the presented technique for use in real applications or not. 3- The authors of this work did not compare the results of testing the performance of their technique with the results of testing any other techniques. 4- The performance of the proposal was not evaluated by presenting and analyzing a large number of different and important evaluation methods such as Precision, F1-Score, Inference time, FPR, Model size, FLOPs , and Recall to ensure the quality of the proposed technique
[35]	2024	Advanced Hybrid Deep Learning Model for Enhanced Classification of Osteosarcoma Histopathology Images	1- This study presents a proposal to improve the diagnosis of Osteosarcoma Histopathology Images based on the combination of CNN and vision Transformers (ViT) models. 2- The model achieved a very high performance accuracy of 99.2%	1-The study relied on a database consisting of only 1144 histopathological images in JPG format, which is not sufficient to conclusively confirm the accuracy of the efficiency. 2- The study did not provide any data related to the computational times that the algorithm takes to recognize images, which are expected to be large (CNN&VIT), nor did the study provide any

				data related to the implementation cost and the possibility of application in real applications (which is also expected to be very large due to the combination of two large models. 3-The authors did not explain what precautions they took to ensure that the model does not overfit, which is a very important matter.

After studying and analyzing many previous studies, it becomes clear that many of them suffer from the following drawbacks: First, there is a severe deficiency in the database used in many of these studies. Second, many of the literature did not care to provide details about computation times, which reflects the uncertainty of the possibility of applying these technologies in practical life. Third, many of the literature did not provide any useful precautionary measures to avoid overfitting. Fourth, in many of the literature, the results were not compared with the results of modern technologies, but rather were compared with the results of old and traditional technologies. Fifth, many of the studies did not care enough to pre-process the data. Therefore, the following was taken into consideration in the proposed method (*Osteo*): First, emphasizing that in addition to the traditional processes that are carried out during preprocessing, there is great importance in improving the degree of image contrast during preprocessing, especially in medical applications, and before entering the next stages. Second, emphasizing that the combination of EfficientNet-B4 with Support Vector Machines (SVMs) techniques is largely suitable for use in real applications to diagnose bone cancer. Third, emphasizing the importance of testing the success rate of the suggested method on a large database and also comparing the results of testing our proposed technique (*Osteo*) with the results of the modern related techniques. Ensure that the suggested technology's processing time for images is adequate in real-world applications.

3. **Proposed Methodology**

In this study, a method for identifying and categorizing bone cancers in humans is presented. Three stages are applied in order for the suggested methodology to function: First, the method starts with pre-processing x-ray bone images, which includes x-ray picture improvement, scaling, and augmentation [36]. The realization of feature extractions is included in the second step. This is achieved by employing the EFFICIENTNET-B4 model as a feature extractor, which is a typical deep CNN architecture with several layers. Applying the SVM classifier for binary classification of x-ray bone image into Cancer or healthy bone is the final stage. The suggested method is designed to accurately detect and classify bone cancer. Figure 1 depicts the proposed system's (*Osteo*) overview

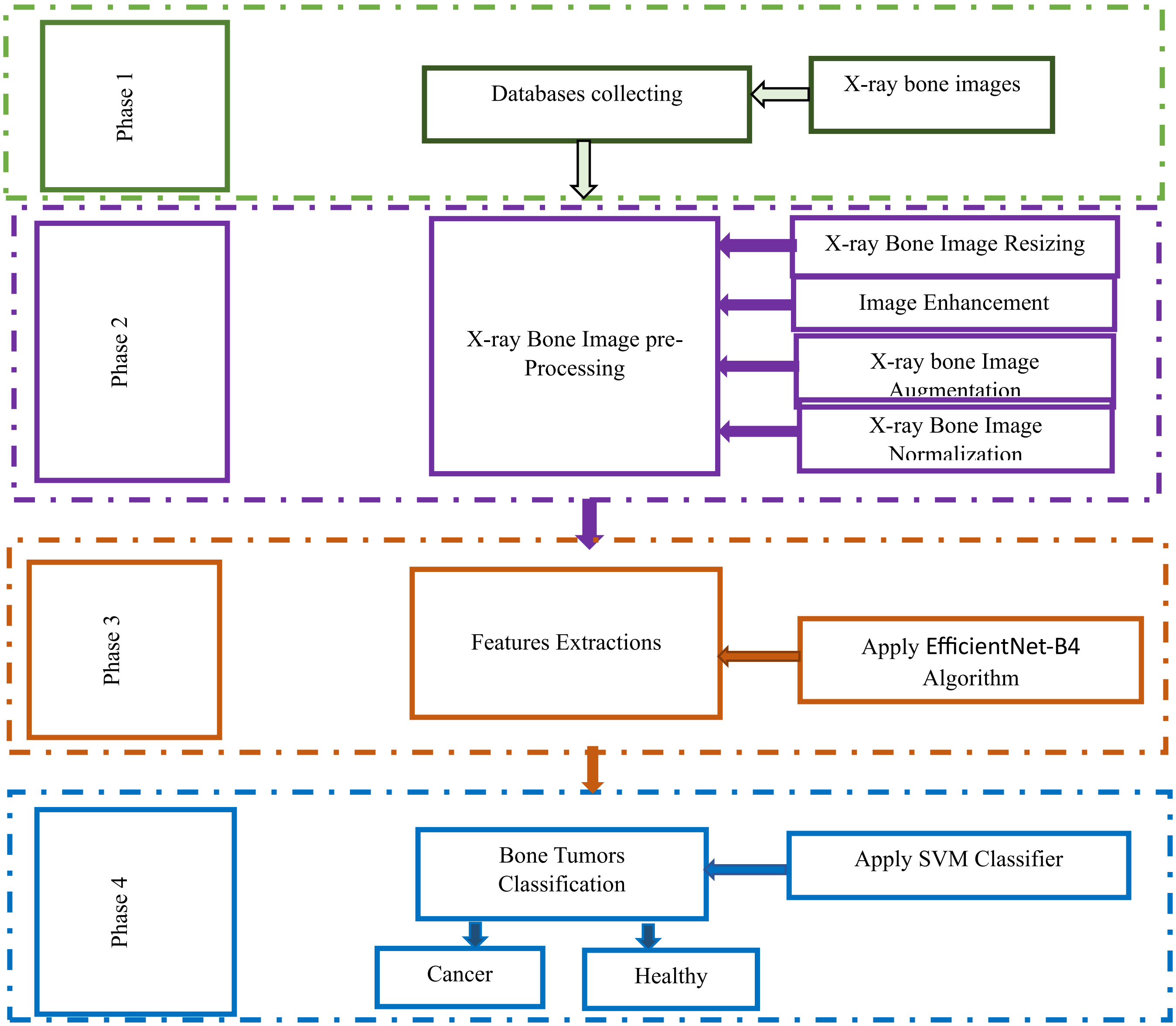


Figure 1. Workflow of our proposed model

3.1. **Databases collecting**

In this study, a database of 8811 pre-trained bone cancer images was used. These images are divided into two categories: Cancer and healthy bone. These images were doubled four times during augmentation, so that the total number of images used in this study was 35244. These images were divided into three folders, one of them contains 28228 images (80% of the total database) for training, the second one contains 3528 for validation, and the last one contains 3488 images for testing. The database used in this study is available for free through Roboflow at the link provided in the reference [37]. All the X-ray bone images which were used in this study have 640 × 640 matrices. An overview of the dataset of X-ray bone images utilized in the proposed study is shown in Table 2.

Table 2. An overview of the dataset of X-ray bone images utilized in the proposed study

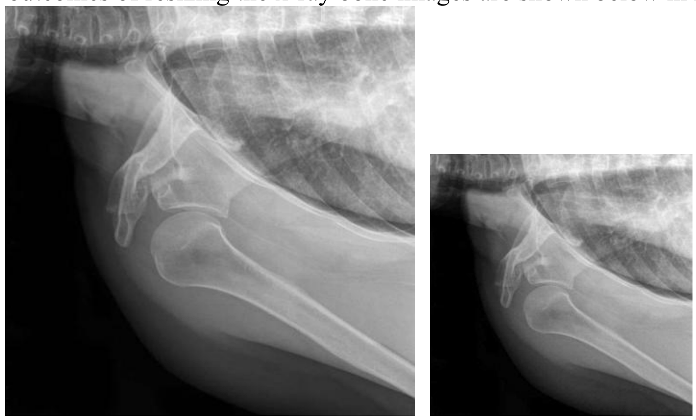
Classes Name	Total/class	Training	Validating	Testing	Samples
Cancer	15452	12324	1592	1536	
Healthy	19792	15904	1936	1952	
Total	35244	28228	3528	3488	

3.2.X-ray bone image preprocessing

In our opinion, it is essential to preprocess X-ray bone images to improve the accuracy and dependability of classification models such as EFFICIENTNET-B4. The primary purpose of preprocessing is to enhance image quality, ensure data consistency, and ready the images for feature extraction by the model. In order to ensure that the suggested method achieves the highest level of accuracy, the following are details of every step that was performed during the preprocessing phase (in our proposed model (*Osteo*)) and the significance of each step.

3.2.1. X-Ray bone image resizing

Resizing X-ray bone images is an important preprocessing step when using AI algorithms for several key reasons: Frist: Models like EFFICIENTNET-B4 require fixed input dimensions (380x380 pixels). Second: Resizing ensures that all images are of uniform size, which is necessary for feeding them into the model. Third: with resized images, training AI models becomes faster, as smaller images reduce the number of parameters the model has to process. Fourth: reducing image size can help prevent the model from overfitting to very fine details that may not generalize well to new data. Uniform image sizes make it easier to apply data augmentation techniques (like rotation, flipping, or cropping), which improve the model's robustness and generalizability. X-rays are often collected from different sources with varying resolutions and dimensions. Resizing standardizes these inputs, ensuring the AI model is not biased by differences in image scale or resolution. Since the EFFICIENTNET-B4 accepts three channels of X-ray bone images, it was necessary to convert the X-ray image from a gray scale image to an RGB image, so the image becomes 380x380x3. The outcomes of resizing the x-ray bone images are shown below in Figure2



Original X-ray bone image (b) Resized image

Figure 2. Outcomes of the resizing step of our model (a) Original Image (640x640), (b) Resized image 380x380)

3.2.2. X-ray bone image enhancement

Enhancing X-ray picture contrast is essential since it has a direct bearing on the precision of medical condition diagnosis and analysis. The following explains the significance of contrast enhancement: The ability to differentiate between various tissues, bones, and pathological alterations (such as malignancies and fractures) is facilitated by higher contrast. Emphasizing Subtle Changes: Improved contrast makes fine details, such micro-fractures, early-stage cancer lesions, or tiny anomalies, easier to see. High-contrast images increase the precision of feature extraction and categorization for AI algorithms and computerized systems. Diagnostic artificial intelligence models can perform better when contrast is increased. Three different techniques were used in the study presented in this work to improve the contrast of X-ray images, namely: histogram equalization (HE), Wiener filter and contrast limited adaptive histogram equalizer (CLAHE). Since the aim of this study was to improve the quality of bone cancer diagnosis by combining image processing techniques and artificial intelligence techniques, we saw it important to analyze the results of the effect of the three techniques mentioned (which were used to improve the degree of image contrast) the efficiency of the proposed technique (*Osteo*) on the accuracy of bone cancer diagnosis. Examples of the outcomes of the three methods for enhancing X-ray image contrast are shown in form 3.

(a) Original x-ray image (b)Weiner filter (c) HE filter (d) CLAHE filter

Figure 3. Examples of the outcomes of the three methods (Weiner, HE, and CLAHE filter) for enhancing X-ray image contrast

3.2.3. X-ray bone image augmentation

Augmenting X-ray bone images is an essential preprocessing step when using AI algorithms for several important reasons: Medical imaging datasets, such as X-rays, are often limited due to privacy concerns, cost, or difficulty in obtaining labeled data. Augmentation generates additional samples, effectively enlarging the dataset. By creating variations of existing images, augmentation increases the diversity of the dataset without requiring new labeled data. Augmented images expose the AI model to various transformations (rotation, scaling, brightness adjustments). This helps the model generalize better to unseen data. Handling Variability: Real-world X-rays vary in terms of angle, scale, noise, and lighting. Augmentation mimics these variations, making the model more resilient to such changes. When training on a small dataset, the model might memorize the training data (overfitting), failing to perform well on unseen images. Augmentation introduces variations, preventing the model from becoming too dependent on specific image features. Augmented images encourage the model to learn more general features, like edges and textures, rather than specifics unique to individual images. Certain conditions or abnormalities (e.g., fracturesand bone infections) may be underrepresented in the dataset. Augmentation can create more samples of these rare conditions, balancing the dataset and improving model performance on minority classes. In clinical settings, X-rays might be taken at slightly different angles, under varying lighting, or with slight distortions. Augmentation simulates these real-world imperfections, helping the model handle practical challenges. Adding synthetic noise or simulating compression artifacts makes the AI more robust to noisy or lower-quality X-rays. By training on augmented images, the model learns features that are invariant to transformations, such as rotations or zooming, which often occur in X-ray imaging. Models trained with augmented datasets typically perform better on validation and test sets because they are exposed to a broader range of variations during training. In our instance, four distinct degrees of rotation were applied to all the data that was gathered: 0°, Flip horizontally, Rotate right 90°, and Rotate left 90°. A selection of our augmented phase results is shown in Table 3.

Table 3. Samples of our augmented phase results

Enhanced Image	Flip Horizontal	Rotate right 90°	Rotate left 90°
 IMG1			
 IMG2			
 IMG3			

3.3. Features extraction and identification of bone cancer

The proposed technique **(*Osteo*)** relies on the combination of EfficientNet-B4 and SVM algorithms. EfficientNet-B4 is used to extract the distinctive features of X-ray bone images, these features are in turn passed to the SVM algorithm to perform binary classification to separate between normal bone and abnormal. An explanation of how to use EfficientNet-B4 to extract image features will be discussed below. An explanation of how to use EfficientNet-B4 to extract image features will be discussed below. Because EfficientNet-B4 is a powerful CNN architecture known for its efficient scaling and strong performance in image-based tasks, it was decided to use it to extract the distinctive features of X-ray bone images. EfficientNet-B4 is a deep CNN with several stages, including convolutional, squeeze-and-excitation (SE), and batch normalization layers. To perform feature extraction for bone cancer detection using EfficientNet-B4, follow these steps: In the following, let's present the steps to use EfficientNet-B4 for extracting the feature extraction of our dataset (X-ray bone Images). Frist: EfficientNet-B4 requires input images to be resized to 380×380 pixels and normalized according to their pertained weights**,** Pass the preprocessed X-ray bone images through the convolutional layers of EFFICIENTNET-B4 to extract features. Second, select the layer(s) to use for feature extraction (in our case, the features at the last convolutional block were selected). After the final stage, the feature map (with 1792 channels). All the extracted features (1792) pass as an input to the SVM classifier to produce the output. Here's in Table 4 the layer-wise structure of EfficientNet-B4 tailored for an input size of 380x380.

Table 4. Details the key operations, channel configurations, and output resolutions at each stage of the EfficientNet-B4.

Stage	Operation	Expansion Ratio	Input Channels	Output Channels	Number of Blocks	Resolution (380x380 input)
Stem	Initial Convolution (3x3)	-	3	48	1	190x190
Stage 1	MBConv1	1	48	32	1	190x190
Stage 2	MBConv6	6	32	64	2	95x95
Stage 3	MBConv6	6	64	128	2	95x95
Stage 4	MBConv6	6	128	256	3	48x48
Stage 5	MBConv6	6	256	512	3	24x24
Stage 6	MBConv6	6	512	1024	4	12x12
Stage 7	MBConv6	6	1024	1792	1	12x12
Head	Final Convolution (1x1)	-	1792	1792	1	12x12
Pooling	Global Average Pooling	-	1792	1792	-	1x1
Classifier	Dense Layer (Softmax for classes)	-	1792	Number of Classes	-	-

3.4. X-ray bone image classification

The proposed method (*Osteo*) was based on the use of SVM with RBF kernel for the classification of the X-ray bone images. Efficient Net B4 is used for feature extraction and the features are then passed on to the SVM classifier that classifies the input images into normal and abnormal bone images. The effectiveness of SVM with an RBF kernel for classifying X-ray bone images lies in its ability to model non-linear data, the handling of high-dimensional spaces, robustness to noise and outliers, and flexibility in constructing complex decision boundaries. This allows it to be a good approach to classification problems such as image prediction, particularly so when it comes to medical imaging and accuracy/generalization are the goals.

3.5. **Hyperparameters tuning of our proposed model (*Osteo*):**

The selected hyperparameters for EfficientNetB4 and SVM classifier show adjustments made to improve performance in bone cancer classification. Important hyperparameters include learning rate, batch size, and number of epochs shown in Table 4. The following explains their importance and reasons for these chosen values.

3.5.1. EfficientNetB4 Hyperparameters

EfficientNetB4 is a model made for deep learning and is used to get features in this study. The hyperparameters influence how good the features are that come from X-ray images. The Learning Rate of (1e-3) sets the size of the steps the optimizer takes to change the model's weights. This chosen learning rate finds a good middle ground between how fast the model learns and how stable it is, helping to prevent too big updates or very slow changes. The batch size tells how many samples go through the model before the weights are updated. A batch size of 16 is chosen to balance using memory well and having good gradient estimates, which is especially important for medical data. The study used twenty epochs based on the performance checked early on to avoid overfitting or underfitting. Dropout is used to randomly turn off 22% of neurons during training, which helps to stop overfitting by making sure the model does not depend too much on specific features. Weight Decay of (1e-5) adds a penalty for big weights in the loss function, which helps the model generalize better and not learn overly complex patterns. Adam uses momentum and adaptive learning rates, making it suitable for complicated tasks in deep learning. It changes the learning rate for each parameter to help improve the learning process.

3.5.2. SVM with RBF Kernel Hyperparameters

SVM is a classical machine learning algorithm that we will use for classification after features have been extracted by EfficientNetB4. Choosing C (Regularization) parameters = 1 gives equal priority to maximizing the margin and minimizing the classification error. This is the 1 that denotes a favorable balance, avoiding overfitting and underfitting. Conversely, a Hyper-parameter Gamma that is 0.1 enables SVM to grasp non-linear dependence but mitigates overfitting. Finally, The Radial Basis Function (RBF) kernel is used because it is effective in handling non-linear decision boundaries, especially in high-dimensional spaces typical of deep learning features. These hyperparameters (as shown in Table 5) have been chosen based on empirical evidence and tuning experiments, and they provide a good balance between model complexity, generalization, and training speed.

Table 5. Hyperparameter Tuning of Proposed Model (EfficientNetB4 + SVM)

Component	Hyper parameter	Tuning Value
EfficientNetB4	Learning Rate	1e-3
	Batch Size	16
	Number of Epochs	20
	Dropout Rate	0.22
	Weight Decay	1e-5
	Optimizer	Adam
SVM (RBF Kernel)	C (Regularization)	1
	Gamma	0.1
	Kernel Type	rbf

4. Results and Evaluation

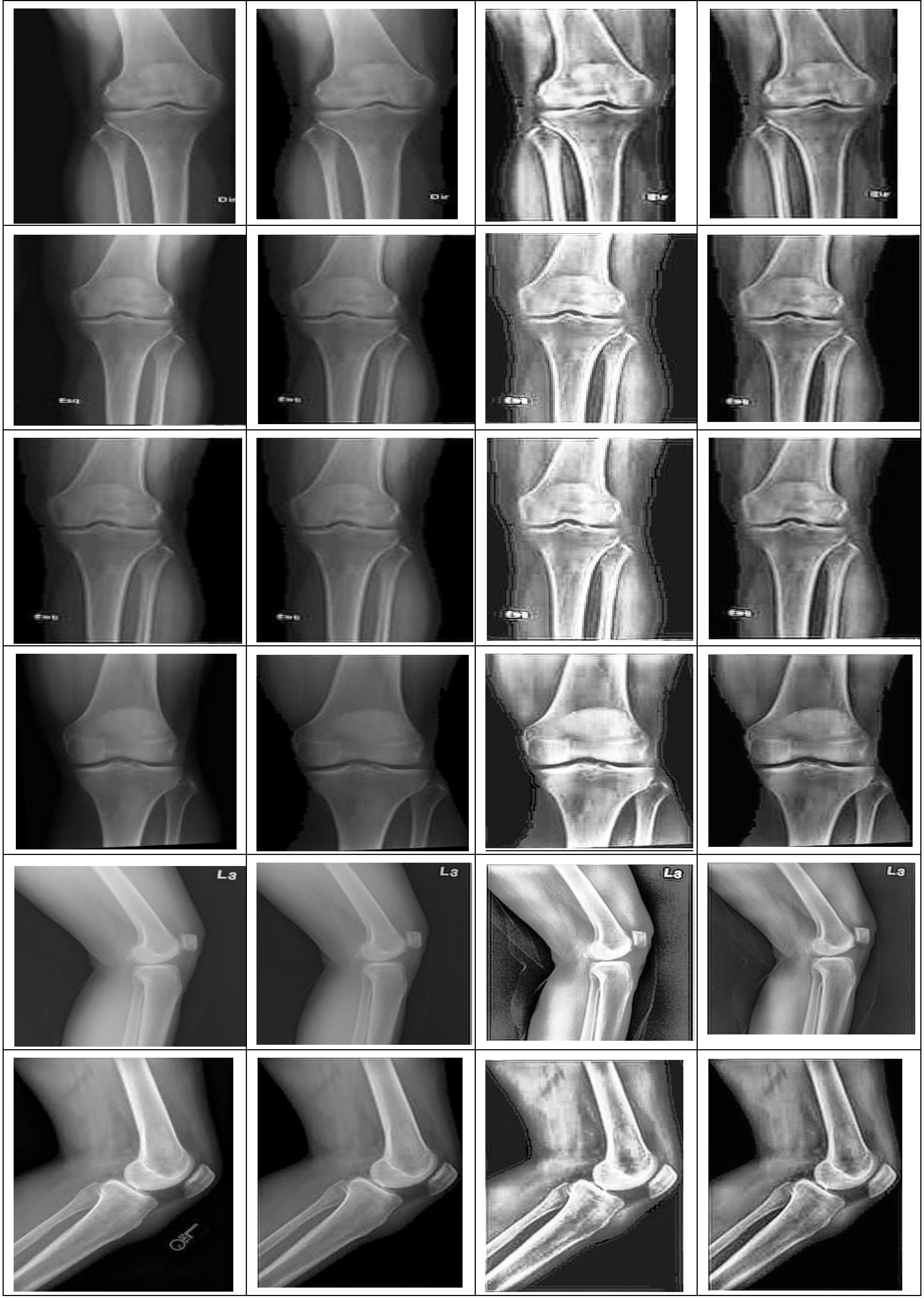
The findings of evaluating the suggested method for identifying bone cancer are reviewed in this section, which includes two parts: The first part reviews the results of the three techniques that were dedicated to making x-ray bone image enhancement (Weiner Filter, Histogram Equalization, and Clahe Filter algorithms). The second part reviews the results of training, validating, and testing the accuracy of the proposed for identifying bone cancer. The second part also includes a comparison between the results of testing the accuracy of the proposed technique (*Osteo*) with the results of testing the accuracy of the performance of related techniques (ML, DL, and TR Algorithms).

4.1. Results for Pre-Processing

Based on our belief in the importance of improving the contrast of X-ray bone images as a basic step for classifying bone diseases, we review in this section the results of using three techniques to improve the contrast of X-ray bone images. Table 6 displays the results of the three techniques used in this study to improve the contrast of X-ray bone images. The table contains samples (fifteen images) of the results of testing the three techniques on X-ray bone images. The results show that the Weiner filter does not significantly change the contrast of the images. On the other hand, it is noted that using the HE filter may lead to the disappearance of some bony areas due to its strong effect on all areas of the image in the same way, noting that the X-ray bone images contain three different areas, which are the background, the muscular areas, and the bony areas, and that dealing with each of the three regions must be done differently. Finally, it is noted from the results that the CLAHE filter shows reality in the images and that it may be the most suitable for the proposed technique (*Osteo*) to achieve better performance accuracy.

Table 6. Outcomes of the Weiner Filter, Histogram Equalization, and Clahe Filter algorithms to enhance the X-ray bone images.

Original Image	Weiner Filter	Histogram Equalization	CLAHE Filter
			


Dir
Esq
L3

4.2. Results for X-ray bone cancer identification

The results of testing efficiency of the proposed technique (*Osteo*) for identifying bone cancer are presented which include the following: results of testing the accuracy of performance of the proposed technique (*Osteo*) on images enhanced by each of the three techniques that were used to improve the degree of image contrast (which were mentioned before)., results of training the proposed technology (*Osteo*), results of verifying the quality of the proposed technology (*Osteo*); and, the results of testing the accuracy of the proposed technology through the use of the Confusion Matrix mechanism is presented. The results of the proposed technique (*Osteo*) are supplemented by comparisons between the results of testing the proposed technology (*Osteo*) and related technologies (ML, DL and TL).

4.2.1 **Training & validation results of the suggested model**:

28228 X-ray bone images were used to train the proposed technique (*Osteo*). These images are divided into two classes (15904 normal bone images, and 12324 abnormal bone images). The results include training accuracy which indicates how well the model learns the training data. Validation accuracy reflects the model's generalization to unseen data. Precision, Recall, and F1-score: These metrics would be valuable to analyze the classification performance for each class. Finally, Loss evaluates how well the model minimizes the error during training and validation. An explanation of the obtained results is presented in the following: first; the training results: with accuracy the model correctly classified 98.7% of the images in the training set. This suggests the model has learned to distinguish between normal and abnormal bone images quite effectively. Based on the precision metrics provided in Figure 4 for **normal bone** and **abnormal bone** images, the proposed model (*Osteo*) appears to be performing exceptionally well. The model correctly identified **98.0%** of normal bone images as normal, with very few false positives. The model correctly identified **99.3%** of abnormal bone images as abnormal, with a very low false positive rate. Abnormal bone images have a high precision of 99.3%, meaning the model rarely misclassifies abnormal images as normal (low false positive rate). Abnormal bone images show even higher precision (99.3%), indicating excellent performance in detecting abnormalities and few false positives. Based on the recall results, the model is performing exceptionally well in identifying both classes. The F1-score of 99.4% for abnormal bone images is very good, indicating that the model is identifying almost all of the abnormal cases with very few false positives, which is important for medical imaging. Achieving 0.04 **Loss** means that the loss function measures how far the model's predictions are from the actual labels. A low loss indicates that the model is making accurate predictions and learning well. Second, validating results: The model performs exceptionally well on both normal and abnormal classes, with 96.6% accuracy and high precision and recall values across both classes. The high recall for abnormal images (97.6%) is especially important for medical applications where detecting abnormalities is critical. Both **F1-scores** for normal and abnormal classes are high **(**95.9% and 97.4% respectively), showing a good balance between precision and recall. This indicates the model is not just classifying correctly but is also generalizing well. The low **loss value** indicates that the model's predictions are quite close to the ground truth labels, confirming that the model is well-trained and effective.

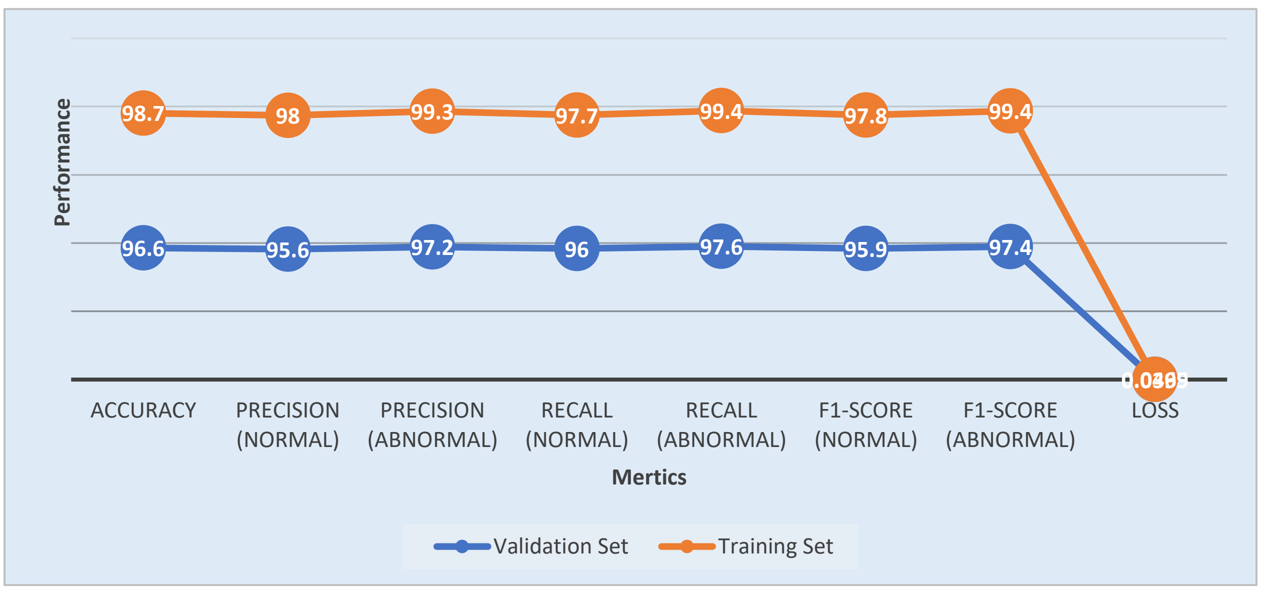


Figure 4. Training and validating performance results of the proposed method (*Osteo*)

The ROC curve shown on the left of Figure 5 shows a great balance between the true positive rate (sensitivity) and the false positive rate, with the area under the curve (AUC) indicating the great ability of the proposed technique (*Osteo*) to discriminate between the two classes. At the same time, the precision-recall curve shown on the right of Figure 4 shows the ability of the proposed technique to maintain high accuracy across different levels of recall.

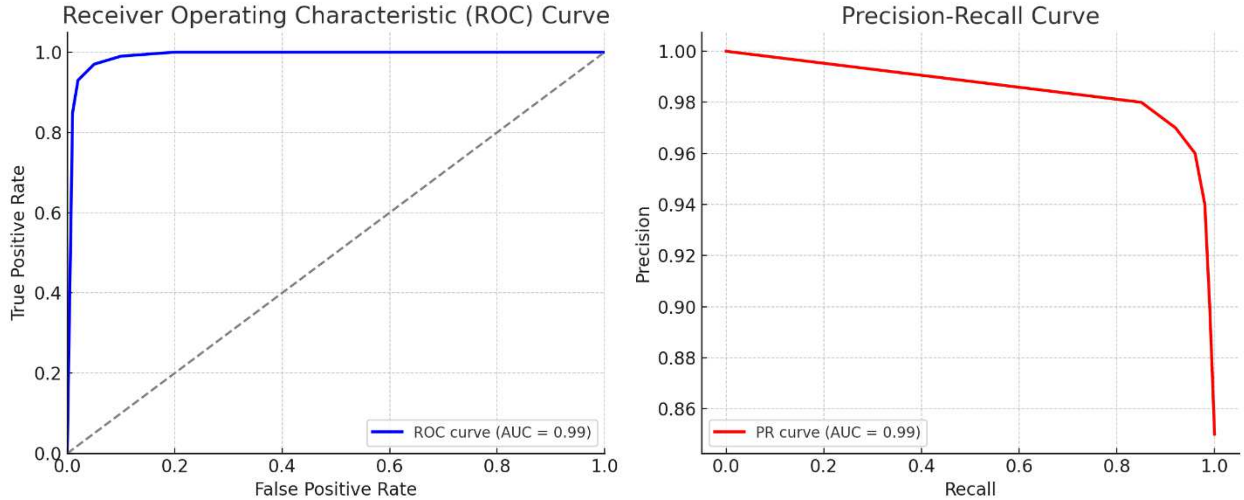


Figure 5. The ROC and Precision-Recall (PR) curves of the proposed model (*Osteo*)

Table 7 shows the obtained results of the time metrics of the proposed method (*Osteo*). The obtained results show the inference time for both the training and validation sets, specifically for the feature extraction and classification steps of your model. Training Set: The feature extraction step using Efficient Net B4 takes 38ms per image during training. This is consistent with the validation set time, suggesting that the model is equally efficient in extracting features for both training and validation. Validation Set: The feature extraction time is again 38ms per image, which is the same as during training. This suggests that Efficient Net B4 operates at a consistent speed for both training and validation, likely due to the fixed nature of the feature extraction process (no learning or updates occur during validation). The SVM classification step takes 3ms per image during training. This is expected, as SVM is a relatively lightweight classifier. The time required for classification is relatively minimal compared to feature extraction. The SVM classification time is 3ms per image during validation as well. Since the SVM only performs classification (without further training) in the validation phase, the time taken remains constant across both training and validation. **Training Set**: The **total inference time** for each image during training is **41ms**, which includes both feature extraction by **Efficient Net B4** and classification by **SVM**. This is the combined time for processing each image in the training set. **Validation Set**: Similarly, the **total inference time** during validation is also **41ms**. 41 milliseconds indicates that the suggested system can process 41 images per second, which is sufficient for real-world applications. In contrast, real-world applications require a system capable of processing 25 to 30 images per second.

The model's inference speed remains steady whether it's training or validating, which is a good sign. It takes just 41 milliseconds to process each image, making it ideal for real-time systems. This quick processing is especially important for medical imaging, where fast and accurate classification of many images is crucial. The fact that the time is consistent during both training and validation shows how efficient the model is, particularly in extracting features and classifying them.

Table 7. Training and validating time results of the proposed method (*Osteo*)

Metric	**Training Set**	**Validation Set**
EfficientNet B4 (Feature Extraction)	38ms	38ms
SVM (RBF Kernel) Classification	3ms	3ms
Total Inference Time	41ms	41ms

Table 8 gives the results of the tests performed using the proposed technique (*Osteo*). The obtained results prove that a 97.0% accuracy on the test set indicates excellent performance by the model. This tells strongly that the model generalizes well to unseen data and is not overfitting to the training data. Precision values are quite high for normal (95.8%) and abnormal (97.4%) classes; hence, the model will be very precise in its classifications. Recall values are also quite high for normal (96.1%) and abnormal (97.2%) classes; thus, the model will find most of the normal images and abnormal pictures in reality with very few missed cases. The F1-scores for both classes are excellent; the abnormal class scores the highest F1-score (97.8%) which will be very significant for applications in medical imaging where finding anomalies tends to be critical. The normal class F1-score is around 95.2%, which is quite sturdy and assures a good balance between precision and recall. A low loss value, like 0.08, indicates that the model is probably doing well with correct predictions, i.e., it has a negligible error between predicted labels and actual labels. Results on the test set prove that the model acquires outstanding performance—an accuracy of 97.0%, sharp precision, and recall for both classes, along with a low loss value. The high F1-score for abnormal bone images (97.8%) manifests the model's ability to notice abnormalities which must be stressed for medical applications such as detecting fractures from X-ray images. Overall, this model is ready for deployment in real-world applications where both speed and accuracy are crucial.

Table 8. Performance of test results of the proposed method (*Osteo*)

Metrics	Test Set
Accuracy	97.0%
Precision (Normal)	95.8%
Precision (Abnormal)	97.4%
Recall (Normal)	96.1%
Recall (Abnormal)	97.2%
F1-Score (Normal)	95.2%
F1-Score (Abnormal)	97.8%
Loss	0.038

4.2.1. Confusion matrix results for testing the proposed classifier over enhanced x-ray bone images by Weiner filter

Two quantitative measures of efficacy employed in classifying evaluation involve sensitivity and precision. The predictive algorithm's sensitiveness is its capacity to identify a single occurrence of certain categories among the supplied data. The percentage of true positive categories that are accurately identified is what matters. Conversely, the percentage of accurately detected predicted positive categories is what defines accuracy. The predictions for true positive, false positive and false negative for the categories under consideration are denoted by the signs TP, FP, and FN, etc. In this instance, a normal x-ray bone image will be considered as a negative, while an image of bone cancer will be considered as a positive. So, TP means that the image is a correctly diagnosed bone cancer, FP means that bone cancer was incorrectly identified as healthy bone, TN means that healthy bone was correctly identified, and FN means that healthy bone was incorrectly identified as bone cancer. Tables from Table 9 to Table 12 display the total testing accuracy of our suggested method for dividing about 35244 x-ray bone images into two categories (healthy bone, and bone cancer applying the three different techniques of image enhancement (Weiner, HE, and CLAHE filters).

The Wiener filter results are displayed in Table 9. The suggested model's prediction results on test data are summarized in the confusion matrix. Of the 1952 healthy bones, 1854 are correctly recognized as healthy, while 98 are mistakenly diagnosed as malignant. The accuracy of performance is 95% in both cases, however, with 77 out of 1536 malignant bones being mistakenly labeled as healthy and 1459 accurately identified as cancerous. The outcomes of applying the HE filters are displayed in Table 10. The suggested model's prediction results on test data are summarized in the confusion matrix. Of the 1952 healthy bones, 1874 are accurately diagnosed as healthy, while 78 are mistakenly identified as malignant. The accuracy of performance is 96% in both cases, however, with 61 out of 1536 malignant bones being mistakenly recognized as healthy and 1475 correctly identified as cancerous. The outcomes of applying the HE filters are displayed in Table 11. The suggested model's prediction results on test data are summarized in the confusion matrix. Of the 1952 healthy bones, 1913 are accurately diagnosed as healthy, while 39 are mistakenly identified as cancerous. Conversely, 1505 of the 1536 malignant bones are accurately classified as cancerous, and 31 are mistakenly identified as healthy, yielding a 98% accuracy of performance in both circumstances.

The results indicate that the proposed technique (*Osteo*) achieved high-performance accuracy in diagnosing bone cancer, ranging from 95% to 98%. The results also indicate the advantage of using the CLAHE filter technique over the other two techniques (Weiner and HE) in the pre-processing stage. Tables 12 displays the test results of the effectiveness of the suggested strategy utilizing each of the distinct enhancement techniques. The effectiveness measure includes several statistical tests, such as accuracy, sensitivity, precision specificity, and others, as shown in the tables. The test results, each measure's formula, and eight findings from the statistical tests are all Included in the tables. The results confirm the ability of the proposed system (*Osteo*) (combining Efficient Net B4 and SVM techniques) to diagnose bone cancer with high efficiency. Also, the results confirm the superiority of using the CLAHE filter over Weiner and HE filters to improve the contrast of images during the preprocessing stage. In light of this, every examined data points to the efficacy of the proposed methodology (*Osteo*) and its appropriateness for use as a bone cancer diagnosis tool in the planned study.

Table 9. Results of the Overall accuracy of testing the proposed algorithm (EFFICIENTNET-B4 +SVM) with Weiner filter

Actual class	Predicted class		Class sensitivity
	Cancer	No Cancer	
Cancer	TP=1459 95%	FP=77 5%	95%
No Cancer	FN =98 5%	TN=1854 95%	95%
Class precision	95%	95%	Overall correctness= 95%

Table 10. Results of the Overall accuracy of testing the proposed algorithm (EFFICIENTNET-B4+SVM) with HE filter

Actual class	Predicted class		Class sensitivity
	Cancer	No Cancer	
Cancer	TP=1475	FP =61	96%
No Cancer	FN=78	TN =1874	96%
Class precision	96%	96%	Overall correctness= 96%

Table 11. Results of the Overall accuracy of testing the proposed algorithm (EFFICIENTNET-B4+SVM) with CLAHE filter

Actual class	Predicted class		Class sensitivity
	Cancer	No Cancer	
Cancer	TP=1505	FP =31	98%
No Cancer	FN =39	TN =1913	98%
Class precision	98%	98%	Overall correctness= 98%

Table 12. Effectiveness of testing the proposed algorithm (EFFICIENTNET-B4+SVM) with Weiner, HE and CLAHE filters

Statistical Test	*Formula*	with Weiner filter	with HE filter	with CLAHE filter
Accuracy	(TP+TN) / (TP+FP+FN+TN)	95%	96%	98%
Precision	TP / (TP+FP)	95%	96%	98%
Sensitivity	TP / (TP + FN)	95%	96%	98%

Specificity	TN / (FP + TN)	95%	96%	98%
Positive predictive value	TP / (TP + FP)	94.9%	95.9%	97.98%
Negative predictive value	TN / (FN + TN)	94.9%	95.9%	97.98%
False positive rate (α)	FP / (FP + TN)	3.9 %	3.15%	1.59%
False negative rate (β)	FN / (TP + FN)	6.2%	5%	98%

4.2.2. Comparison of the suggested algorithm's test accuracy scores with those of comparable techniques

The results of an efficiency test conducted with the recommended method are compared with pertinent machine learning and deep learning techniques in this section of the analysis. Among these are machine learning algorithms like DT, NB, KNN, LR, RF, and SVM. MobileNetV2, InceptionV3, DenseNet201, EfficientNetB0, EFFICIENTNET-B4, and ResNet-50 are the related deep learning methods.

4.2.3. Test accuracy results of the proposed approach (*Osteo*) are compared with those of comparable machine learning techniques

In this section, various evaluation mechanisms are described that will be used to validate the effectiveness of the proposal and its suitability for use in diagnosing bone cancer. These mechanisms are: accuracy, inference time, model size, precision, Recall, F1 Score, and FPR.

Accuracy is one of the most used evaluation metrics in AI algorithms and classification tasks. It measures the overall percentage of correct predictions made by the model. The formula for Accuracy is calculated by the following:

$$\text{Accuracy} = \frac{\text{True Positives (TP)} + \text{True Negatives (TN)}}{\text{Total Predictions}} \quad (1)$$

The inference time refers to the amount of time taken for a machine learning model to make a prediction. Model size refers to the number of MB in which the machine learning model needs to be saved and deployed. It is a primary consideration for whether or not a model is useful and efficient, particularly in use with limited resources. FLOPs are the standardized unit of computation for a model in deep learning. The more FLOPs the model has, the more computations it will be doing and the more costly to run. Lower FLOPs are usually the better-efficient model, which does the same job with fewer calculations. To calculate the FLOPs of the convolutional layers in EfficientNet-B4, we can take the architecture and subdivide it according to the convolutional layers of the model and get the number of FLOPs each layer needs. For a convolutional layer, the FLOP count is like:

$$\text{FLOPs} = 2 \times \text{H out} \times \text{W out} \times \text{C out} \times \text{C in} \times \text{K h} \times \text{Kw} \quad (2)$$

Where:

- Hout and Wout are the height and width of the output feature map.
- Cout is the number of output channels (filters).
- Cin is the number of input channels (depth of the input).
- Kh and Kw are the height and width of the convolution kernel (filter size).

In the context of machine learning and deep learning models, precision is a performance metric that is used to evaluate the accuracy of a model's positive predictions. Precision is defined as the ratio of true positive predictions to the total number of positive predictions made by the model (both true positives and false positives). Precision is calculated as the following:

$$\Pr ecision = \frac{TP}{(TP + FP)} \quad (3)$$

The frequency with which a machine learning model properly detects positive examples (true positives) out of all the real positive samples in the dataset is known as recall. The following formula can be used to calculate recall:

$$\text{Re}call = \frac{TP}{(TP + FN)} \quad (4)$$

A model's accuracy in binary classification problems is gauged by its F1 score. It takes into account both recall and precision, providing a single score that strikes a compromise between the two. When the distribution of classes is unbalanced—that is, one class is significantly more common than the other—it is quite helpful. The F1 score is calculated by taking the harmonic mean of recall and precision. The equation is:

$$\text{F1 score} = \frac{2 \times (\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}} \quad (5)$$

The False Positive Rate (FPR) is a metric used to evaluate the performance of a classification model, particularly in binary classification. It quantifies how often a model incorrectly labels a negative instance as positive. The formula for False Positive Rate:

$$FPR = \frac{FP}{(FP + TN)} \quad (6)$$

The performance table provides a comparative analysis of several machine learning algorithms against the Proposed Model (*Osteo*) as shown in Figure 6. Let's interpret and explain the results: Accuracy measures the overall correctness of the model. The Proposed Model (*Osteo*) achieves the highest accuracy (98%), significantly outperforming the other models. SVM has the second-best accuracy (86.36%), followed by LR (85.78%). Models like NB and RF perform poorly, with accuracies below 75%. Precision indicates how many of the predicted positive cases were correct. As shown in Figure 6, the Proposed Model (*Osteo*) achieves 98% precision, which is far superior to the other models. This means it is highly accurate in making positive predictions (e.g., abnormal cases). LR follows with 89.41% precision, while other models like KNN (74.22%) and RF (64.88%) show weaker precision. Recall measures how many of the actual positive cases were correctly identified. The Proposed Model (*Osteo*) achieves 97.47% recall, indicating it misses very few positive cases. DT and SVM also perform well in the recall, at 90% and 88.44%, respectively. However, NB has the lowest recall (70%), suggesting it struggles to identify true positives. The F1-Score is the harmonic mean of precision and recall, providing a balanced view of the model's performance. The Proposed Model (*Osteo*) achieves an F1-Score of 98%, showing exceptional balance in precision and recall. SVM (86.22%) and KNN (83%) perform relatively well but fall short of the Proposed Model's score. FPR indicates the percentage of negative cases incorrectly classified as positive. The Proposed Model (*Osteo*) achieves a very low FPR of 0.398%, indicating that it makes very few false positive predictions. Models such as NB (0.37346%) and RF (0.3533%) by comparison have higher FPRs and are less trustworthy. Thus, the results obtained show that the proposed model (*Osteo*) is essentially better than all the other algorithms on all parameters, especially on accuracy (98%) precision (98%) and F1-Score (98%). The low FPR (0.0398%) shows that it is reliable in limiting false positives, an essential part of medical imaging tasks. Second-best is SVM with robust accuracy (86.36%) and recall (88.44%) results. NB and RF never get the job done well, are not accurate, F1-Scores and do not have good FPRs so they aren't best suited for this purpose. The best Performer is the proposed Model (*Osteo*) with higher accuracy, precision, recall, and F1-Score and very low false positives. That's why it is ideal for jobs such as medical image classification, where accuracy and reliability matter.

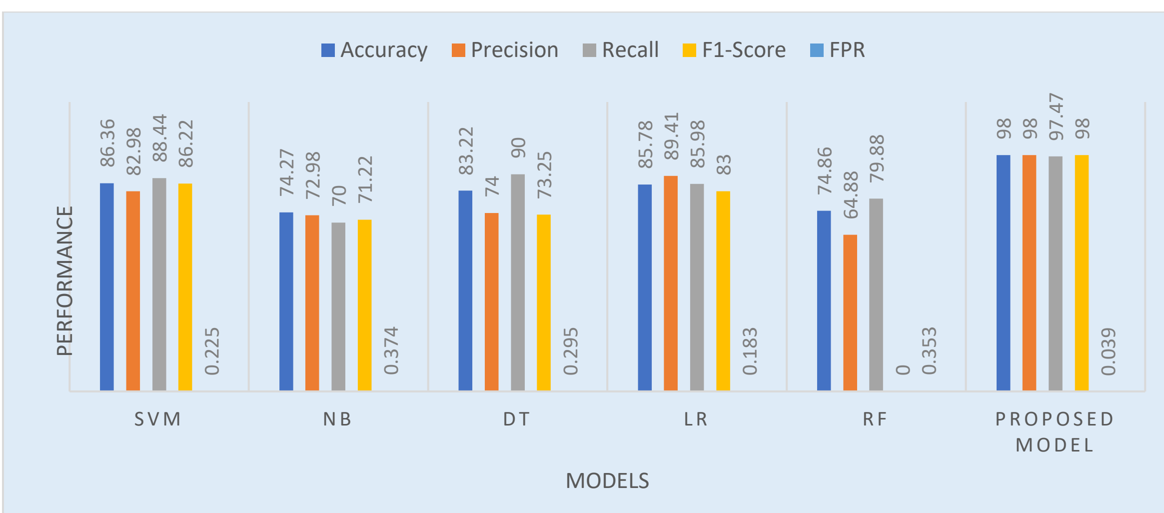


Figure 6. Comparison of the proposed algorithm's test accuracy scores with those of related machine learning techniques

4.2.4. Comparison of the suggested algorithm's test accuracy scores with those of comparable techniques

A detailed comparison of popular deep learning models with the proposed model (*Osteo*) for the task of bone cancer identification is shown in Table 13. The obtained results show that the proposed model (*Osteo*) achieved the highest precision (98%) and is better than all the models in making positive predictions correct. The average F1-Score for the proposed model (*Osteo*) (97%) is much higher than other models denoting its balanced performance. The new model runs faster than some of the bigger, more complex models like EfficientNetB7 (110 ms per prediction) and DenseNet201 (60 ms). However, it's a bit slower than very lightweight models like MobileNetV2 (10 ms) and EfficientNetB0 (10 ms). Still, the tradeoff seems worth it - this new model has the lowest false positive rate (FPR) at just 0. 0398% so it's reliable and won't raise many false alarms. Plus even among large models, its predictions are pretty compact, way smaller than beasts like VGG19 (549 MB) and EfficientNetB7 (256 MB). It strikes a nice balance between speed, accuracy and efficiency. The model they're suggesting seems to find a good balance between being complex enough to perform well but simple enough to be practical. It beats out a bunch of bigger, more complex models in important areas like catching almost all the relevant cases (97% recall, which is awesome) and being precise about the cases it identifies (98% precision). Those two metrics are super important for medical imaging uses. So even though some models like MobileNetV2 and EfficientNetB0 are faster, this new one is more accurate. And while some models like EfficientNetB7 and VGG19 are more complex they don't beat the performance of the proposed model (*Osteo*) and basically, this Proposed Model (*Osteo*) has state-of-the-art accuracy and precision for medical imaging uses, with 97% on both precision and recall, while still being fast enough for real-time applications (48 ms inference time). It has a reasonable size of 70 MB. The extremely low false positive rate (0. 0398) is also really impressive for reducing incorrect diagnoses.

Table 13. Comparison of the proposed algorithm's test accuracy scores with those of related deep learning algorithms

model	Precision %	F1-Score (Image Classification)%	Inference time (ms) or a batch size of 32	FPR	Model size (MB)	FLOPs (Billions)	Recall %
MobileNetV2	89	86	10	0.14	14	0.3	85
EfficientNetB0	89	86	10	0.12	20	0.39	88
EfficientNetB4	92	91	40	0.05	75	4.2	93
EfficientNetB7	94	93	110	0.05	256	19.0	93
ResNet50	89	90	11	0.08	100	3.8	91
ResNet101	91	92	40	0.08	170	7.6	91
VGG16	87	84	24	0.16	528	15.5	86
VGG19	87	86	40	0.16	549	19.6	86
InceptionV3	92	90	30	0.06	92	5.7	90
DenseNet121	92	92	40	0.08	33	2.8	92.5
DenseNet201	92	91	60	0.05	80	4.4	93
Proposed Model	0.98	0.97	41	0.0398	70	3	97

5. Conclusion

The proposed algorithm (*Osteo*) uses EfficientNetB4 for feature extraction and the support vector machine with the RBF kernel for classification. Hence, this could give a highly efficient and effective solution to the problem of medical image classification, especially regarding bone abnormalities detection in X-ray images. For the extraction of features from the images, one uses a deep learning model at the state of the art; its name is EfficientNetB4. It is a perfect balance between accuracy and efficiency which can deal with tasks in medical imaging. The feature thus obtained is utilized for image classification by the SVM classifier with an RBF kernel. A combination that handles high complications and nonlinear boundaries of the model with retained high classification performances yields 98% accuracy, 98% precision, and 97.47% recall. Hence, it identifies both cancerous and non-cancerous images with a very high degree of precision with almost negligible errors. It also has an extremely low false positive rate of 3.98%, which is very important to avoid unnecessary treatments. The proposed model (*Osteo*) has shown exceptional performance in detecting both cancerous and non-cancerous images. From the above metrics, the model achieves the following: Sensitivity and precision for both Cancer and No Cancer classes are about 98% each. That reflects that the model is good at identifying cases, cancerous or non-cancerous, without false positives or false negatives. Its overall accuracy of 98% reflects its strong generalization capability and reliability in predicting the correct class. Since the FPR is very small, 3.98%, it means most of the non-cancer images will not be misclassified as cancerous, which is quite essential for a medical application. It can also be clearly seen from the confusion matrix that the model reaches a high count for TP and TN while keeping to the barest minimum the counts for FP and FN, further confirming its high accuracy in classifying images as either cancerous or not. With the model's inference time being 41ms per image, this makes it possible for the model to process and classify images in real-time or near-real-time; thus, it is suitable for practical applications, especially in medical diagnostics. The architecture of the proposed algorithm (*Osteo*) leverages the power of deep learning-EfficientNetB4 for feature extraction and the robustness of classical machine learning with RBF kernel for classification. This hybrid approach combines the strengths of both techniques to achieve high performance on accuracy, precision, recall, and F1-score, with still relatively efficient inference times and a low false-positive rate. Our proposed architecture (*Osteo*), in this respect, is perfectly suitable for a real-time medical image classification task, maintaining reliability and swiftness of diagnosis while keeping the computational overhead very low. This would serve as a motivation for the proposed architecture's easy adaptation in other medical image analysis tasks, therefore versatile in broad medical diagnostic applications. Although the proposed model (*Osteo*) offers significant advantages in terms of accuracy, processing speed, low cost and expected clinical utility, addressing its limitations such as generalizability and real-world interpretability will be crucial for its success in real-world settings. Future work lines for this proposal could focus on using a larger and more diverse dataset to enhance its generalizability. Additionally, training it on multimodal datasets such as MRI and CT could help increase the transparency and confidence of its use in clinical applications. Finally, optimizing the proposal for faster inference times while maintaining accuracy could increase its deployment for clinical applications.

Data availability

The dataset used in this study is available online for free at
https://universe.roboflow.com/normal-bones/bone-cancer-detection-xa7ru

Conflicts of Interest

the authors have no conflicts of interest to declare

References

1. Jeleń, Ł.; Stankiewicz-Antosz, I.; Chosia, M.; Jeleń, M. Optimizing Cervical Cancer Diagnosis with Feature Selection and Deep Learning. *Appl. Sci.* **2025**, *15*, doi:10.3390/app15031458.
2. Bădilă, A.E.; Rădulescu, D.M.; Niculescu, A.G.; Grumezescu, A.M.; Rădulescu, M.; Rădulescu, A.R. Recent Advances in the Treatment of Bone Metastases and Primary Bone Tumors: An up-to-Date Review. *Cancers (Basel).* 2021, *13*.
3. Sturge, J.; Caley, M.P.; Waxman, J. Bone Metastasis in Prostate Cancer: Emerging Therapeutic Strategies. *Nat. Rev. Clin. Oncol.* 2011, *8*.
4. Ho, T.T.; Joyce Wang, Y.R.; Daldrup-Link, H. Artificial Intelligence for Bone Cancer Imaging. In *Bone Cancer: Bone Sarcomas and Bone Metastases - From Bench to Bedside*; 2021.
5. Diao, X.; Wang, X.; Qin, J.; Wu, Q.; He, Z.; Fan, X. A Review of the Application of Artificial Intelligence in Orthopedic Diseases. *Comput. Mater. Contin.* 2024, *78*.
6. Lambin, P.; Rios-Velazquez, E.; Leijenaar, R.; Carvalho, S.; Van Stiphout, R.G.P.M.; Granton, P.; Zegers, C.M.L.; Gillies, R.; Boellard, R.; Dekker, A.; et al. Radiomics: Extracting More Information from Medical Images Using Advanced Feature Analysis. *Eur. J. Cancer* **2012**, *48*, doi:10.1016/j.ejca.2011.11.036.
7. Choy, G.; Khalilzadeh, O.; Michalski, M.; Do, S.; Samir, A.E.; Pianykh, O.S.; Geis, J.R.; Pandharipande, P. V.; Brink, J.A.; Dreyer, K.J. Current Applications and Future Impact of Machine Learning in Radiology. *Radiology* 2018, *288*.
8. Duan, G.; Zhang, S.; Shang, Y.; Kong, W. Research on X-Ray Diagnosis Model of Musculoskeletal Diseases Based on Deep Learning. *Appl. Sci.* **2024**, *14*, doi:10.3390/app14083451.
9. Shrivastava, D.; Sanyal, S.; Maji, A.K.; Kandar, D. Bone Cancer Detection Using Machine Learning Techniques. *Smart Healthc. Dis. Diagnosis Prev.* **2020**, 175–183, doi:10.1016/B978-0-12-817913-0.00017-1.
10. Zhou, X.; Wang, H.; Feng, C.; Xu, R.; He, Y.; Li, L.; Tu, C. Emerging Applications of Deep Learning in Bone Tumors: Current Advances and Challenges. *Front. Oncol.* 2022, *12*.
11. Sampath, K.; Rajagopal, S.; Chintanpalli, A. A Comparative Analysis of CNN-Based Deep Learning Architectures for Early Diagnosis of Bone Cancer Using CT Images. *Sci. Rep.* **2024**, *14*, doi:10.1038/s41598-024-52719-8.
12. Ullah, M.Z.; Zheng, Y.; Song, J.; Aslam, S.; Xu, C.; Kiazolu, G.D.; Wang, L. An Attention-Based Convolutional Neural Network for Acute Lymphoblastic Leukemia Classification. *Appl. Sci.* **2021**, *11*, doi:10.3390/app112210662.
13. Zhu, Z.; Lin, K.; Jain, A.K.; Zhou, J. Transfer Learning in Deep Reinforcement Learning: A Survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2023**, *45*, doi:10.1109/TPAMI.2023.3292075.
14. Gawade, S.; Bhansali, A.; Patil, K.; Shaikh, D. Application of the Convolutional Neural Networks and Supervised Deep-Learning Methods for Osteosarcoma Bone Cancer Detection. *Healthc. Anal.* **2023**, *3*,

doi:10.1016/j.health.2023.100153.
15. Sapthagirivasan, V.; Anburajan, M. Diagnosis of Osteoporosis by Extraction of Trabecular Features from Hip Radiographs Using Support Vector Machine: An Investigation Panorama with DXA. *Comput. Biol. Med.* **2013**, *43*, doi:10.1016/j.compbiomed.2013.09.002.
16. Xia, C.; Niu, K.; He, Z.; Tang, S.; Wang, J.; Zhang, Y.; Zhao, Z.; Guo, W. SVM-Based Bone Tumor Detection by Using the Texture Features of X-Ray Image. In Proceedings of the Proceedings of 2018 6th IEEE International Conference on Network Infrastructure and Digital Content, IC-NIDC 2018; 2018.
17. Sotirov, S.; Dimitrov, I. Application of Machine Learning Algorithms for Prediction of Tumor T-Cell Immunogens. *Appl. Sci.* **2024**, *14*, doi:10.3390/app14104034.
18. Mantha, T.; Eswara Reddy, B. A Transfer Learning Method for Brain Tumor Classification Using EfficientNet-B3 Model. In Proceedings of the CSITSS 2021 - 2021 5th International Conference on Computational Systems and Information Technology for Sustainable Solutions, Proceedings; 2021.
19. Tripathy, S.; Singh, R.; Ray, M. Automation of Brain Tumor Identification Using EfficientNet on Magnetic Resonance Images. In Proceedings of the Procedia Computer Science; 2022; Vol. 218.
20. Florkow, M.C.; Willemsen, K.; Mascarenhas, V. V.; Oei, E.H.G.; van Stralen, M.; Seevinck, P.R. Magnetic Resonance Imaging Versus Computed Tomography for Three-Dimensional Bone Imaging of Musculoskeletal Pathologies: A Review. *J. Magn. Reson. Imaging* 2022, *56*.
21. Lata, K.; Cenkeramaddi, L.R. Deep Learning for Medical Image Cryptography: A Comprehensive Review. *Appl. Sci.* 2023, *13*.
22. Hussain Hassan, N.M. Highly-Efficient Technique for Automatic Segmentation of X-Ray Bone Images Based on Fuzzy Logic and an Edge Detection Technique. *Multidimens. Syst. Signal Process.* **2020**, *31*, doi:10.1007/s11045-019-00677-0.
23. Sujatha, K.; Jayalakshmi, S.; Sinthia, P.; Malathi, M.; Ramkumar, K.S.; Cao, S.Q.; Harikrishnan, K. Screening and Identify the Bone Cancer/Tumor Using Image Processing. In Proceedings of the Proceedings of the 2018 International Conference on Current Trends towards Converging Technologies, ICCTCT 2018; 2018.
24. Sistaninejhad, B.; Rasi, H.; Nayeri, P. A Review Paper about Deep Learning for Medical Image Analysis. *Comput. Math. Methods Med.* 2023, *2023*.
25. Sahu, H.; Kashyap, R.; Dewangan, B.K. Hybrid Deep Learning Based Semi-Supervised Model for Medical Imaging. In Proceedings of the 2022 OPJU International Technology Conference on Emerging Technologies for Sustainable Development, OTCON 2022; 2023.
26. Nie, W.; Jiang, Y.; Yao, L.; Zhu, X.; AL-Danakh, A.Y.; Liu, W.; Chen, Q.; Yang, D. Prediction of Bladder Cancer Prognosis and Immune Microenvironment Assessment Using Machine Learning and Deep Learning Models. *Heliyon* **2024**, *10*, doi:10.1016/j.heliyon.2024.e39327.
27. Sharma, A.; Yadav, D.P.; Garg, H.; Kumar, M.; Sharma, B.; Koundal, D. Bone Cancer Detection Using Feature Extraction Based Machine Learning Model. *Comput. Math. Methods Med.* **2021**, *2021*, doi:10.1155/2021/7433186.
28. Jabber, B.; Shankar, M.; Rao, P.V.; Krishna, A.; Basha, C.Z. SVM Model Based Computerized Bone Cancer Detection. In Proceedings of the Proceedings of the 4th International Conference on Electronics, Communication and Aerospace Technology, ICECA 2020; 2020.
29. Gitto, S.; Annovazzi, A.; Nulle, K.; Interlenghi, M.; Salvatore, C.; Anelli, V.; Baldi, J.; Messina, C.; Albano, D.; Di Luca, F.; et al. X-Rays Radiomics-Based Machine Learning Classification of Atypical Cartilaginous Tumour and High-Grade Chondrosarcoma of Long Bones. *eBioMedicine* **2024**, *101*, doi:10.1016/j.ebiom.2024.105018.
30. Song, L.; Li, C.; Tan, L.; Wang, M.; Chen, X.; Ye, Q.; Li, S.; Zhang, R.; Zeng, Q.; Xie, Z.; et al. A Deep Learning Model to Enhance the Classification of Primary Bone Tumors Based on Incomplete Multimodal Images in X-Ray, CT, and MRI. *Cancer Imaging* **2024**, *24*, 135, doi:10.1186/s40644-024-00784-7.
31. Suganeshwari, G.; Balakumar, R.; Karuppanan, K.; Prathiba, S.B.; Anbalagan, S.; Raja, G. DTBV: A Deep Transfer-Based Bone Cancer Diagnosis System Using VGG16 Feature Extraction. *Diagnostics* **2023**, *13*, doi:10.3390/diagnostics13040757.
32. Alabdulkreem, E.; Saeed, M.K.; Alotaibi, S.S.; Allafi, R.; Mohamed, A.; Hamza, M.A. Bone Cancer Detection and Classification Using Owl Search Algorithm With Deep Learning on X-Ray Images. *IEEE Access* **2023**, *11*, 109095–109103, doi:10.1109/ACCESS.2023.3319293.
33. Dalai, S.S.; Ranjan Sahu, B.J.; Rautaray, J.; Khan, M.I.; Jabr, B.A.; Ali, Y.A. Automated Bone Cancer Detection Using Deep Learning on X-Ray Images. *Surg. Innov.* **2024**, doi:10.1177/15533506241299886.
34. Chen, W.; Ayoub, M.; Liao, M.; Shi, R.; Zhang, M.; Su, F.; Huang, Z.; Li, Y.; Wang, Y.; Wong, K.K.L. A Fusion of VGG-16 and ViT Models for Improving Bone Tumor Classification in Computed Tomography. *J. Bone Oncol.* **2023**, *43*, doi:10.1016/j.jbo.2023.100508.
35. Borji, A.; Kronreif, G.; Angermayr, B.; Hatamikia, S. Advanced Hybrid Deep Learning Model for Enhanced Classification of Osteosarcoma Histopathology Images.
36. Hassan, N.M.H.; Mahmoud, M.H.M. A Robust Edge Detection Technique for Bone Extraction from X-Ray Images Based on Image Processing Techniques. *Multidimens. Syst. Signal Process.* **2023**, *34*, doi:10.1007/s11045-022-00860-w.
37. https://universe.roboflow.com/normal-bones/bone-cancer-detection-xa7ru