Feature-Driven Deep Learning for Osteoporosis Detection: Leveraging Explainable AI

by

A.K.M. Salman Hosain
22166033

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
M.Sc. in Computer Science

Department of Computer Science and Engineering
Brac University
October 2024

Declaration

It is hereby declared that

1. The thesis submitted is my/our own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

A K M Salman Hosain
ID: 22166033

Approval

The thesis titled "Feature-Driven Deep Learning for Osteoporosis Detection: Leveraging Explainable AI" submitted by

1. A K M Salman Hosain (22166033)

Of Summer, 2024 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of M.Sc. in Computer Science on October 19, 2024.

Examining Committee:

External Examiner:
(Member)

Dr. Taskeed Jabid
Professor
Department of Computer Science and Engineering
East West University
Ph.D. in Computer Vision and Image Processing,
Kyung Hee University, South Korea

Internal Examiner:
(Member)

Dr. Md. Golam Rabiul Alam
Professor
Department of Computer Science and Engineering
BRAC University

Supervisor:
(Member)

Dr. Aniqua Nusrat Zereen
Assistant Professor
Department of Computer Science and Engineering
BRAC University

Program Coordinator:
(Member)

Dr. Md Sadek Ferdous
Associate Professor
Department of Computer Science and Engineering
BRAC University

Chairperson:

Sadia Hamid Kazi, Ph.D.
Associate Professor
Department of Computer Science and Engineering
BRAC University

Abstract

Osteoporosis, a widespread bone disorder affecting over 200 million people globally, is associated with significant morbidity, particularly due to increased fracture risk. While traditionally considered a condition afflicting the elderly, younger individuals are also susceptible. In this study, we present a deep learning framework utilizing Convolutional Neural Networks (CNNs) to detect osteoporosis from knee X-ray images. Two models, MobileNet V3 and EfficientNet B0, were employed and fine-tuned on a dataset of 774 knee X-ray images. We further improved model performance by curating a new dataset, emphasizing critical features using explainable AI (XAI). Our results show that while both models achieved an accuracy of 0.77 on the original dataset, EfficientNet B0 consistently outperformed MobileNet V3 on the new dataset with an accuracy of 0.9189 and an F1 score of 0.9315, compared to MobileNet V3's accuracy of 0.8243 and F1 score of 0.8169. These findings demonstrate the effectiveness of CNNs, particularly EfficientNet B0, in accurately diagnosing osteoporosis from medical images, and underscore the importance of both model selection and feature-focused data preprocessing in improving diagnostic performance.

Keywords: Osteoporosis, CNN, EffiecientNet B0, MobileNet V3, XAI

Acknowledgement

I, A K M Salman Hosain, a student from the Computer Science and Engineering (CSE) department at BRAC University, submit this paper as the culmination of two years of academic research, representing my final thesis.
I extend my heartfelt appreciation to my supervisor, Dr. Aniqua Nusrat Zereen, for her unwavering support, invaluable guidance, and insightful advice throughout this research process. Her dedication to academic excellence have been pivotal to the development of this study. I am deeply grateful to Dr. Zereen for her patience, encouragement, and the time she invested in mentoring me.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Osteoporosis: Impact and Complications

A consensus development conference statement defines osteoporosis as a condition marked by reduced bone mass and the deterioration of bone tissue's microarchitecture, resulting in increased bone fragility and a heightened risk of fractures. Bone mineral measurements are highly specific in assessing fracture risk and allow for the development of cutoff values to identify individuals at high risk before fractures occur. These measurements can also diagnose osteoporosis by establishing thresholds, such as 2.5 standard deviations (SD) below the young adult mean, which identifies approximately 30% of postmenopausal women as osteoporotic [1]. A lower threshold of 1 SD below the young adult mean could be used for prevention, as it includes many women at increased risk of fractures. Diagnostic categories for adult women based on bone mineral density (BMD) are as follows[1]:

- **Normal**: BMD within 1 standard deviation (SD) of the young adult mean.
- **Low Bone Mass (Osteopenia)**: BMD between 1 and 2.5 SD below the young adult mean, where prevention is crucial.
- **Osteoporosis**: BMD more than 2.5 SD below the young adult mean.
- **Severe Osteoporosis**: BMD more than 2.5 SD below the young adult mean with one or more fragility fractures.

Osteoporosis is a significant global health concern, characterized by a decrease in bone density and quality, which increases the susceptibility to fractures. The risk of osteoporosis is greater in women than men [1]. This condition predominantly affects older adults, particularly postmenopausal women, but it is not exclusive to this demographic; younger individuals can also suffer from osteoporosis, making it a pervasive and debilitating disorder across age groups [2]. The fundamental issue in osteoporosis is the disruption of bone homeostasis, a delicate balance maintained by a complex interplay of hormones, cytokines, and growth factors. These biological agents work in harmony to regulate bone formation and resorption, ensuring optimal bone mass and strength. When this balance is disturbed, bone mass decreases, bone microarchitecture deteriorates, and the risk of fractures significantly increases [6], [14], [16], [19], [27].

Research from the Global Longitudinal Study of Osteoporosis in Women (GLOW) highlights the multifaceted nature of osteoporosis and its intersection with other chronic conditions. The study identified that individuals with osteoporosis often have concurrent chronic diseases, including hypertension, heart disease, asthma, chronic obstructive pulmonary disease (COPD), arthritis, stroke, inflammatory bowel disease, Parkinson's disease, multiple sclerosis, and type I diabetes. These conditions are linked to an increased risk of fractures, compounding the health challenges faced by individuals with osteoporosis [4], [12]. A comprehensive study involving nearly 20,000 adults in Germany further supports these findings, revealing that 95% of osteoporosis patients had at least one other disease. Common comorbidities included arthrosis, arthritis, chronic low back pain, depression, and chronic heart failure, which were more prevalent in osteoporosis patients compared to those without the condition [10].

The global burden of osteoporosis is immense, affecting over 200 million people worldwide [12]. This condition poses a substantial public health challenge, with estimates suggesting that approximately one-third of women and one-fifth of men aged 50 and older will experience fragility fractures due to osteoporosis. These fractures can lead to significant morbidity, decreased quality of life, and increased mortality, emphasizing the urgent need for effective prevention and management strategies [12], [26]

1.2 Motivation

In recent years, the advancement of artificial intelligence (AI) has brought about revolutionary changes in various sectors, particularly in the field of medical diagnostics. Deep learning techniques, especially Convolutional Neural Networks (CNNs), have emerged as powerful tools in the analysis and interpretation of medical images. These AI-driven methods enable rapid and precise detection of a wide range of diseases, transforming disease diagnosis and management. CNNs, with their ability to learn and identify complex patterns in images, have been instrumental in diagnosing conditions in both humans and animals. This technology facilitates swift and accurate identification of disease symptoms, which is crucial for timely and effective treatment and prevention [22].

In the context of osteoporosis, traditional visual assessment of knee X-ray images by radiologists can provide an initial diagnosis. However, this process is time-consuming and subjective, often leading to variability in diagnostic accuracy due to human factors such as fatigue or lack of expertise. Misdiagnoses, whether false positives or false negatives, can have serious implications for patient care and outcomes. Implementing deep learning-based diagnostic systems offers a promising solution to these challenges. These systems can enhance diagnostic precision, reduce human error, and ensure consistent detection of medical conditions at early stages, ultimately improving patient care and outcomes [21].

To further improve the reliability and interpretability of AI models in medical diagnostics, Explainable AI (XAI) frameworks have been developed. These frameworks address the 'black box' problem of many machine learning models, where the decision-making process is not transparent. XAI frameworks provide detailed explanations of how AI models arrive at their predictions, making the models more understandable and trustworthy to healthcare professionals and patients alike. This

transparency is crucial for building confidence in AI-driven diagnostics and ensuring their effective integration into clinical practice [20], [25].
Despite the success of traditional machine learning in various applications, it faces challenges in real-world scenarios, particularly in obtaining sufficient labeled training data that matches the test data distribution. This process can be costly, time-consuming, and often impractical [13], [17], [23]. Transfer learning offers a solution by leveraging knowledge from previously trained models to enhance learning efficiency. In our work, we utilized models initially trained on the ImageNet dataset, a large-scale dataset with over 1.2 million labeled images across 1,000 classes, serving as a benchmark for image classification models [3]. Since these original models were not trained on osteoporosis-specific images, we fine-tuned and trained the models using our own dataset, keeping the initial layers frozen. This approach allowed us to measure the models' accuracy in detecting osteoporosis in knee X-ray images, demonstrating the potential of AI in enhancing osteoporosis diagnosis and management.
Osteoporosis is a prevalent bone disease affecting over 200 million people globally, characterized by reduced bone mass and deterioration in bone tissue microarchitecture, leading to increased fracture risk. Early detection of osteoporosis is crucial for effective management and prevention of fractures, especially in high-risk populations. While traditional diagnostic methods such as bone mineral density (BMD) measurements and visual assessments of X-rays by radiologists have been used, these methods can be time-consuming, subjective, and prone to human error. The advent of artificial intelligence (AI) and deep learning offers an opportunity to address these limitations by automating the detection process, enhancing accuracy, and reducing the reliance on human expertise. This research was conducted to explore the application of deep learning, particularly Convolutional Neural Networks (CNNs), in improving osteoporosis detection from medical images, leveraging the power of explainable AI (XAI) to ensure transparency in the decision-making process. The main motivation behind this work is to offer a more reliable, efficient, and consistent diagnostic tool for osteoporosis detection using AI.

1.3 Research Problem

While several studies have successfully applied machine learning and deep learning techniques to medical image classification, the field of osteoporosis detection using knee X-rays has been relatively underexplored. Prior research has primarily focused on transfer learning methods, utilizing models such as AlexNet, VGGNet, and ResNet for classifying osteoporosis based on hip X-rays, but knee X-rays, which are crucial in diagnosing osteoporosis, have received less attention. Moreover, most of these studies did not incorporate explainable AI methods, making the models' decision processes opaque, which limits their practical use in clinical settings. The lack of transparency (the "black box" issue) in deep learning models for medical diagnostics is a significant barrier to their adoption by healthcare professionals. Additionally, earlier works did not emphasize feature-focused data preprocessing techniques, which can significantly improve model performance by highlighting the critical areas in medical images. This research aims to bridge these gaps by employing explainable AI techniques such as LIME (Local Interpretable Model-agnostic Explanations) and Grad-CAM (Gradient-weighted Class Activation Mapping) alongside state-of-

the-art CNN architectures to not only improve osteoporosis detection accuracy but also to make the models more interpretable.

1.4 Research Objective

The objectives of our research is to:

- **Proposed Deep Learning Framework**: Introduce a deep learning pipeline utilizing transfer learning method for detecting osteoporosis from knee X-ray images.
- **Explainable AI (XAI) Implementation**: Address the 'black box' issue of machine learning models by incorporating explainable AI
- **Feature-focused Dataset Creation**: Conduct feature extraction method and propose a dataset for improved quantitative performance parameters
- **Comparative Analysis of Models**: Conduct a comparative analysis of models with original and proposed dataset to showcase performance improvement across various quantitative parameters

Overall, the study advances the field by integrating cutting-edge AI technologies with medical imaging, focusing on both accuracy and interpretability.

1.5 Thesis Organization

The thesis is organized as follows:

- **Chapter 1: Introduction** – This chapter introduces the problem of osteoporosis, its global impact, and the potential for AI-driven diagnostics to improve early detection and management. It outlines the motivation for this research, the objectives, and the structure of the thesis.
- **Chapter 2: Related Work** – This chapter reviews the existing literature on osteoporosis detection, transfer learning methods, and the application of deep learning models to medical image classification. It also highlights the gaps in current research, particularly the limited focus on knee X-rays and the lack of explainable AI in previous studies.
- **Chapter 3: Methodology** – This chapter details the dataset used for training and testing, the preprocessing steps, and the architecture of the deep learning models. It also explains the XAI techniques employed in this research, including LIME and Grad-CAM, and describes the steps taken to improve model performance by creating a feature-focused dataset.
- **Chapter 4: Results** – This chapter presents the performance metrics of the models, including accuracy, precision, recall, and F1 score. It compares the performance of EfficientNetB0 and MobileNet V3 on both the original and feature-extracted datasets, demonstrating the improvements achieved through the proposed methodology.

- **Chapter 5: Conclusion** – This chapter summarizes the key findings of the research, discusses the implications of the results, and suggests directions for future work in applying AI and explainable AI to medical diagnostics.

This organization ensures a logical flow from identifying the problem and gaps in current research to proposing and evaluating a solution, followed by a discussion of its implications and future prospects.

Chapter 2

Related Work

Osteoporosis is a condition characterized by weakened bones and an increased risk of fractures, commonly affecting the elderly. It often progresses silently until a fracture occurs, making early detection crucial. Transfer learning, a technique in machine learning where a model developed for one task is reused as the starting point for a model on a second task, proves to be highly effective in medical image classification due to its ability to leverage pre-trained networks and adapt them to specific datasets. Recognizing the importance of accurate diagnosis, Wani et al. employ transfer learning techniques utilizing AlexNet, VGGNet-16, ResNet, and VGGNet-19 to classify knee joint X-Ray images into three categories: normal, osteopenia, and osteoporosis [32]. Their dataset comprises 381 knee X-Ray images. Among the models used, the pretrained AlexNet achieves the highest accuracy, reaching 91.1%. Abubakar et al. also adopt transfer learning techniques with VGG 16 to detect osteoporosis from the Mendeley dataset, which consists of 323 radiographs of knees affected by osteoporosis and 323 normal knee radiographs [18]. Given that osteoporosis leads to decreased bone density and quality, their model's high accuracy of 88% with fine-tuned parameters is particularly significant for early intervention and treatment.

Ensemble learning, which combines multiple models to improve overall performance, shows promising result in enhancing the accuracy of osteoporosis detection through medical imaging. Klontzas et al. use ImageNet-trained CNN architectures, including VGG-16, InceptionResNetV2, and InceptionV3, to differentiate between 210 cases of avascular necrosis and 210 cases of transient osteoporosis in hip Magnetic Resonance Images (MRI) [24]. By adopting an ensemble decision-making process, they improve classification accuracy, highlighting the potential of ensemble learning in medical diagnostics. In their work, Inception-ResNet-V2 attained the highest AUC score of 97.62%. Kumar et al. also leverage ensemble learning by employing a fuzzy rank-based unification of classifiers with convolutional neural network (CNN) architectures such as Inception v3, Xception, and ResNet 18 [30]. Their approach achieves an impressive accuracy of 93.5% with a loss of 0.082. Their dataset includes X-Rays from 240 subjects, with 37 having normal bone density, 154 having osteopenia, and 49 having osteoporotic bone density. This study underscores the effectiveness of ensemble methods in accurately detecting varying degrees of bone density. Rasool et al. propose a weighted ensemble learning model using EfficientNetB0 and DenseNet121 to detect knee osteoporosis from a dataset of 372 knee images. By assigning different weights to these models, they identify the optimal combination to

achieve the highest accuracy. This innovative approach further demonstrates the potential of ensemble learning to improve the detection and diagnosis of osteoporosis, ultimately aiding in better patient outcomes [33].

One prominent technique in Explainable AI (XAI) is Local Interpretable Model-agnostic Explanations (LIME). LIME is designed to elucidate the predictions of complex, black-box models by approximating them with locally interpretable models. This technique can be applied to any machine learning model, irrespective of its type (e.g., neural networks, decision trees, Support Vector Machine), as it treats the model as a black box and requires no knowledge of its internal mechanics. LIME is primarily focused on explaining individual predictions rather than the overall model behavior. It generates explanations valid for the specific instance being analyzed, thereby providing insights into why the model made a particular prediction. The interpretable model is trained on perturbations of the instance being explained, with these perturbations weighted according to their proximity to the original instance. Osteoporosis is a condition characterized by decreased bone mineral density and mass, leading to weak and brittle bones. Often undiagnosed until fractures occur, it requires early detection. Utilizing an open-source dataset of 1,493 patients, multiple heterogeneous machine learning frameworks were designed by Khanna et al. to predict osteoporosis risk [29]. The best-performing model, combining Forward Feature Selection and a custom multi-level ensemble learning stack, achieved 89% accuracy. Explainable AI tools like SHAP, LIME, ELI5, and Qlattice were used to interpret model predictions, aiding physicians in diagnostic decision-making. Osteoporosis requires early detection for effective management, yet traditional screening methods and machine learning models often struggle with low accuracy and lack individualized explanations. A study by Suh et al. develops an interpretable deep-learning (DL) model for osteoporosis risk screening, using XAI techniques to provide individual feature contributions and insights [31]. The model, trained on NHANES and KNHANES datasets, achieves high AUC values of 0.851 and 0.922 for femoral neck and total femur bone mineral density, respectively. By employing methods like LIME, the study offers detailed, individualized explanations of risk factors, enhancing both the accuracy and interpretability of osteoporosis screening. Similar techniques like Grad-CAM highlights the regions of an image that are critical for a particular class prediction by producing a coarse localization map that can be overlaid on the original image to indicate the areas the model is focusing on. Grad-CAM employs the gradients of the target class flowing into the final convolutional layer to generate this localization map. These gradients reveal the importance of each neuron in the convolutional layer for the target class prediction [9]. Kassem et al. performs feature extraction using GoogleNet, ResNet50, and AlexNet networks, trained on the ImageNet dataset, to classify pelvis images into fractions and normal categories [28]. The networks are fine-tuned with new layers while retaining the pre-trained ones to speed up training and reduce data requirements. To enhance interpretability, the study employs Gradient-weighted Class Activation Mapping (Grad-CAM), which uses gradients to highlight the important regions in images that influence classification decisions, validating that the network focuses on relevant pelvic areas for accurate predictions. Jang et al. developed a deep learning model to predict osteoporosis from simple hip radiography, using 1,001 images and VGG16 with a non-local neural network [15]. The model achieves an accuracy of 81.2% and an AUC of 0.867. Grad-CAM is used to visualize and validate that the model cor-

rectly focuses on relevant features like the proximal femur cortex and trabecular patterns in the radiographs, confirming its potential as a practical screening tool for osteoporosis.

Table 2.1: Summary of Literature Review

Paper	**Method**	**Dataset**	**Accuracy**	**XAI**	**Disease Detected**
Wani et al. [32]	AlexNet, VGGNet-16, ResNet, VGGNet-19 (Transfer Learning)	381 knee X-rays	91.1% (AlexNet)	No	Osteoporosis, Osteopenia
Abubakar et al. [18]	VGG 16 (Transfer Learning)	323 osteoporosis knee X-rays, 323 normal knee X-rays	88%	No	Osteoporosis
Klontzas et al. [24]	VGG-16, InceptionResNetV2, InceptionV3 (Ensemble Learning)	210 avascular necrosis, 210 transient osteoporosis hip MRIs	97.62% (Inception-ResNet-V2)	No	Transient Osteoporosis
Kumar et al. [30]	Fuzzy rank-based unification of CNN classifiers (Ensemble Learning)	X-Rays from 240 subjects (37 normal, 154 osteopenia, 49 osteoporosis)	93.5%	No	Osteoporosis, Osteopenia
Khanna et al. [29]	Forward Feature Selection, Multilevel ensemble learning stack	1,493 patient records	89%	Yes	Osteoporosis
Suh et al. [31]	Interpretable deep-learning (DL) with XAI (LIME, SHAP)	NHANES, KNHANES datasets	AUC: 0.851 (femoral neck), 0.922 (total femur)	Yes	Osteoporosis (Femoral Neck, Total Femur)
Jang et al. [15]	VGG16 with non-local neural network, Grad-CAM (XAI)	1,001 hip X-rays	81.2% (Accuracy), AUC 0.867	Yes	Osteoporosis

Chapter 3

Methodology

In this work, we propose a proposed pipeline to detect osteoporosis from knee X-Ray images employing transfer learning method with EffecientNetB0 [11], and MobileNet [8]. We use XAI architecture, LIME [7] to differentiate the features that our models focus primarily on the classification task. Subsequently, we implement a feature extraction method to isolate the prominent features from the original images, creating a separate dataset, and train our models on this extracted feature dataset to enhance classification accuracy. We utilize various numerical parameters to demonstrate the improvement of model accuracy by adopting our proposed framework. The complete framework is depicted in Fig. 3.1.

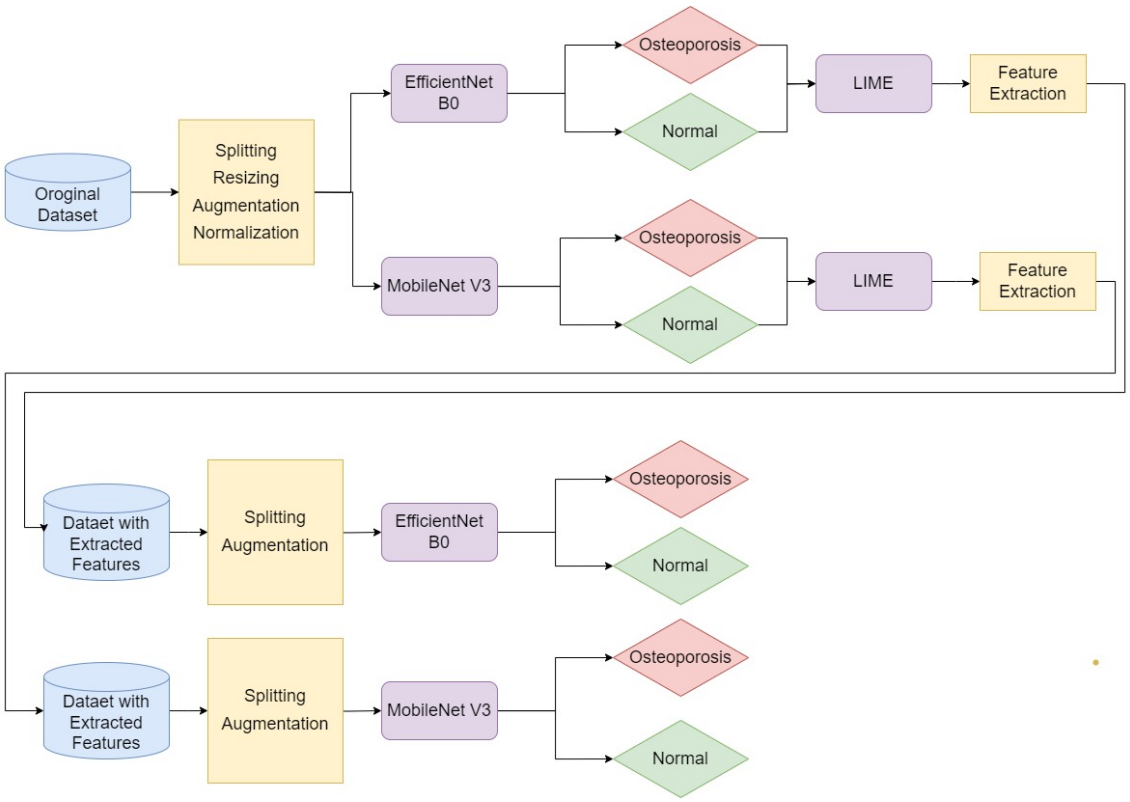


Figure 3.1: Proposed framework to detect osteoporosis from Knee X-Ray images.

3.1 Dataset Description

We utilize X-Ray images from available dataset in Kaggle [34]. The dataset contains images of knee radio-graphs of two categories: normal, and osteoporosis. A total of 774 images, distributed equally across two categories .

(a) Normal Knee X-Ray

(b) Osteoporosis Knee X-Ray

Figure 3.2: Sample images from dataset.

3.2 Dataset Pre-processing

We split the dataset into three sets: training, validation, and test by the ration 7:2:1. The training set contains a total of 520 images, the validation set 150 images and the test 74 images. The distribution of images in dataset is depicted in Fig. 3.3.

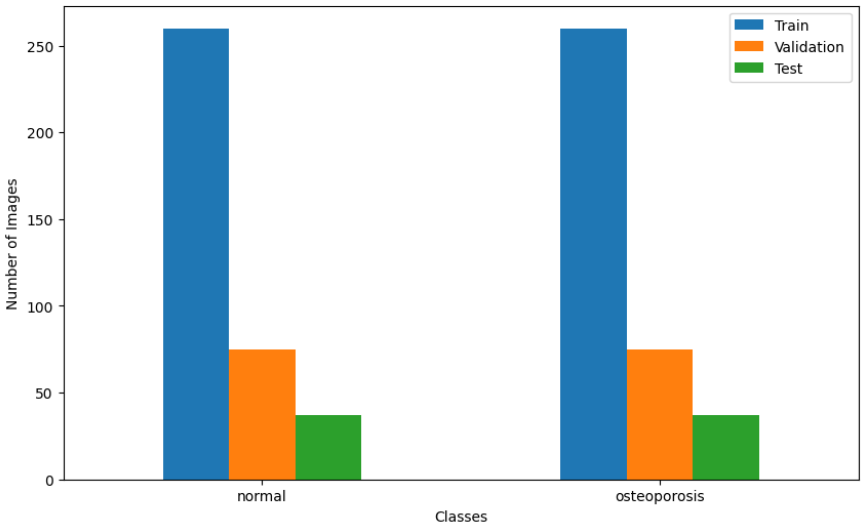


Figure 3.3: Image distribution of training, validation, and test sets.

The images are of various sizes, as depicted in Fig. 3.4. Other statistic parameters of image sizes are presented in Table 3.1. The table provides a summary of the image sizes in the training set, detailing key statistical measures for both width and height in pixels. The mean width of the images is 629.29 pixels, and the mean height is 956.12 pixels, indicating the average dimensions of the images in the dataset. The

standard deviation for the width is 756.87 pixels, and for the height, it is 909.49 pixels. The minimum width of the images is 128 pixels, and the minimum height is 256 pixels, which are the smallest dimensions in the dataset. The first quartile (25%) for both width and height is 128 pixels and 256 pixels respectively, meaning 25% of the images have dimensions equal to or smaller than these values. The median (50%) values are also 128 pixels for width and 256 pixels for height, indicating that half of the images have dimensions equal to or smaller than these values. The third quartile (75%) for width is 1068 pixels and for height is 1994 pixels, meaning that 75% of the images have dimensions equal to or smaller than these values. The maximum width recorded in the dataset is 2430 pixels, and the maximum height is 2660 pixels, representing the largest image dimensions.

To ensure uniformity and increase model memory and computational efficiency, we resize all the images to 224×224 pixels. We further augment the images to increase the number of training images. The augmentation parameters we adopt include randomly rotating images by up to 20 degrees, randomly shifting images horizontally by up to 20% of the width, randomly shifting images vertically by up to 20% of the height, randomly applying shear transformations by up to 20%, randomly zooming in/out on images by up to 20%, and randomly flipping images horizontally.

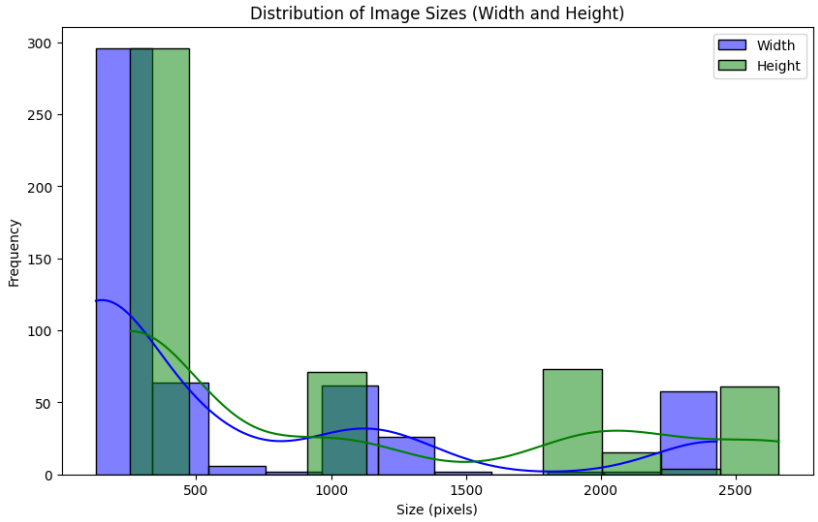


Figure 3.4: Image size distribution.

Table 3.1: Summary of Image Sizes in the Training Set

Statistic	Width	Height	Units
Mean	629.29	956.12	pixels
Standard Deviation	756.87	909.49	pixels
Min	128	256	pixels
25%	128	256	pixels
50%	128	256	pixels
75%	1068	1994	pixels
Max	2430	2660	pixels

3.3 Model Description

The proposed framework utilizes a pipeline to detect osteoporosis from knee X-Ray images using transfer learning method with EffecientNetB0, and MobileNet.

3.3.1 EfficientNetB0

EfficientNet introduces a Convolutional Neural Network (CNN) architectures by leveraging compound scaling across width, depth, and resolution. This method ensures that the resulting models are efficient in terms of both computational resources and performance, making them highly suitable for deployment in real-world applications where efficiency and accuracy are critical factors [11].
The proposed CNN architecture utilizes the EfficientNetB0 model as a foundational feature extractor for image classification. This model is initialized with weights learned from extensive training on a large-scale image classification dataset, ImageNet [3], providing a robust starting point for feature extraction in our detection task.
In the feature extraction phase, the top layer of the EfficientNetB0 model is excluded, transforming it into a feature extractor. This modification enables the efficient extraction of high-level features from the input images, which are crucial for capturing the intricate patterns necessary for accurate classification.
Following feature extraction, custom classification layers are added to refine and classify the extracted features. The first layer is a Global Average Pooling layer, which serves to reduce the spatial dimensions while preserving the essential features of the input. This is followed by a Dense layer with 1024 units and a Rectified Linear Unit (ReLU) activation function. The ReLU function, defined in Equation 3.1.

$$\text{ReLU}(x) = \max(0, x) \tag{3.1}$$

where $\max(0, x)$ returns x if x is positive and 0 otherwise. This activation function is chosen for its effectiveness in introducing non-linearity into the model, thereby enabling it to learn complex patterns within the data.
The final layer is a Dense layer with a softmax activation function, designed for multi-class classification. The softmax function, defined in Equation 3.2.

$$\text{Softmax}(z)_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \quad \text{for } i = 1, \ldots, K \tag{3.2}$$

where z is a vector of raw class scores, K is the number of classes, and e is the base of the natural logarithm. This function produces a probability distribution over the classes, ensuring that the output probabilities for each class sum to one. This facilitates the model in making clear, probabilistic predictions for each class in the dataset.
The model compilation phase involves setting up the model for training by defining the loss function and the optimizer. Categorical crossentropy is selected as the loss function, which is well-suited for multi-class classification problems. It is defined in Equation 3.3.

$$\text{Categorical Crossentropy} = -\frac{1}{N} \sum_{i=1}^{N} \sum_{c=1}^{C} y_{ic} \log(p_{ic}) \tag{3.3}$$

where y_{ic} represents the true label and p_{ic} represents the predicted probability for class c for the i-th sample. The Adam optimizer is chosen for its adaptive learning rate capabilities, which help in efficiently navigating the optimization process during training. Through this meticulous setup, the model is well-prepared for robust training and accurate prediction tasks.
The proposed model is depicted in Fig. 3.5.

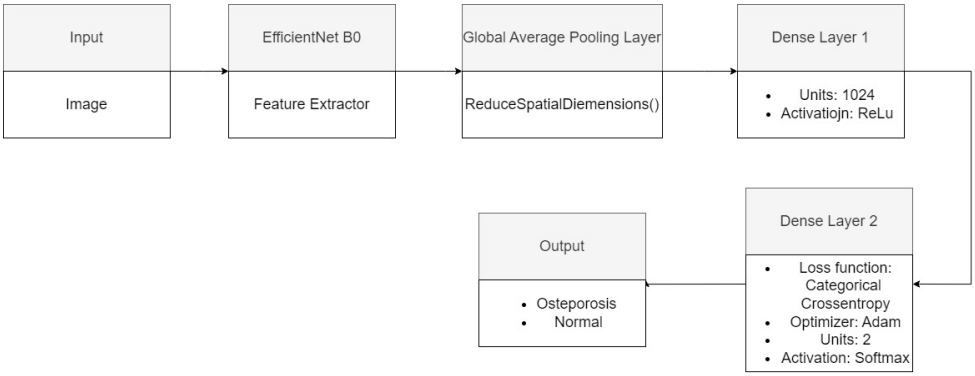


Figure 3.5: Proposed model architecture based on EfficientNetB0 as encoder.

3.3.2 MobileNet V3

MobileNet is built on a streamlined architecture that employs separable convolutions to construct lightweight deep neural networks. This architecture splits the standard convolution operation into two distinct layers: a depthwise convolution and a pointwise convolution. This separation drastically reduces the number of parameters and computations compared to traditional convolutions, making MobileNet models both efficient and effective. These models are designed to strike a balance between accuracy and computational cost, performing well on benchmark datasets while maintaining low resource requirements [8].
The architecture begins with an Input Layer that accepts input images followed by a Convolutional Layer, which consists of an initial standard convolution layer with a large number of filters, accompanied by Batch Normalization and ReLU activation. The ReLU activation function, is defined in Equation 3.1. Which is applied to introduce non-linearity into the model. The core of MobileNet comprises Depthwise Separable Convolution Blocks. Each block contains a depthwise convolution followed by a pointwise convolution. These blocks significantly reduce the computational load and parameter count, forming the essential building units of MobileNet.
After the convolutional blocks, the architecture includes a Global Average Pooling layer. This layer reduces each feature map to a single value by averaging, thus decreasing the spatial dimensions and further reducing the computational complexity. Finally, a Fully Connected Layer is employed for classification tasks. This dense layer has a number of neurons equal to the number of output classes, completing the architecture and enabling the model to perform its classification functions efficiently.
The proposed model using MobileNet V3 as feature extractor is depicted in Fig. 3.6

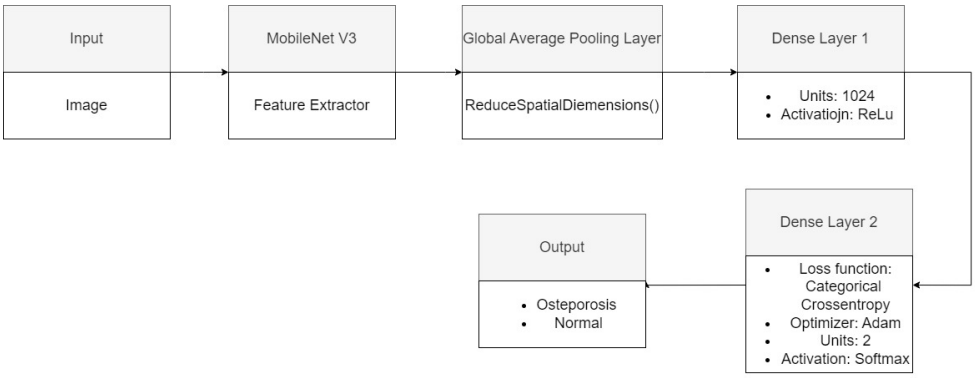


Figure 3.6: Proposed model architecture based on MobileNet V3 as encoder.

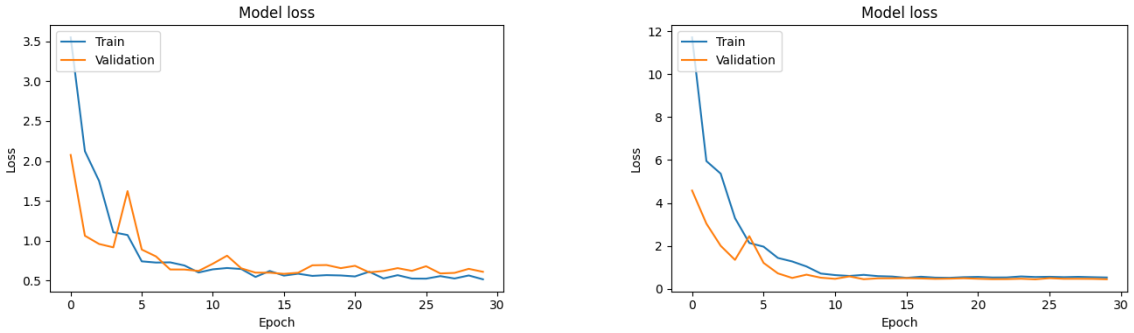
3.4 Model Training

The models are initially trained on the original dataset, which consist of 520 images, for 30 epochs. Following the training, EfficientNetB0 achieves an accuracy of 87.88%, while MobileNet V3 attains an accuracy of 83.08%. For both models, the 'Adam' optimizer [5] and the 'Categorical Cross Entropy' loss function are utilized. The categorical cross-entropy loss is defined as:

$$L = -\sum_{i=1}^{N}\sum_{c=1}^{C} y_{i,c}\log(p_{i,c}) \tag{3.4}$$

where:

- L is the categorical cross-entropy loss,
- N is the number of samples,
- C is the number of classes,
- $y_{i,c}$ is a binary indicator (0 or 1) if class label c is the correct classification for sample i,
- $p_{i,c}$ is the predicted probability of sample i being in class c.



(a) EfficientNetB0 training loss

(b) MobileNet V3 training loss

Figure 3.7: Training loss of models on original dataset

Fig. 3.7 shows a figure depicting two subplots that compare the training loss and validation loss of two different neural network models over a number of training epochs.
Fig. 3.7a displays the training and validation loss curves for the EfficientNetB0 model. The x-axis represents the number of training epochs, while the y-axis shows the loss values ranging from 0.5 to 3.5. Two distinct curves are plotted. The blue curve represents the loss on the training dataset, which decreases rapidly during the initial epochs and stabilizes as training progresses. The orange curve represents the loss on the validation dataset, which follows a similar trajectory, initially decreasing but with some fluctuation. Toward the end of the training process, the validation loss becomes almost flat, signifying that the model's performance has plateaued.
Fig. 3.7b presents the training and validation loss curves for the MobileNet V3 model. The x-axis represents the number of training epochs, while the y-axis shows the loss values, which range from 0 to 5. The blue curve corresponds to the training loss, which rapidly declines in the early epochs, indicating that the model is learning quickly. Similar to EfficientNetB0, the loss gradually stabilizes after the initial few epochs. The orange curve represents the validation loss, which also decreases rapidly during the first few epochs but shows some fluctuations in the early stages. However, like the training loss, it levels off as training progresses.

3.5 Explainable AI

Further to improve the models, we adopt two XAI techniques: LIME [7], and Grad-Cam [9]. After classifying the images into two classes with the models, we utilize these XAI architectures to address the black box issue of the deep learning models. It illustrates the regions of the images on which the models primarily focus for classification.

3.5.1 Local Interpretable Model-agnostic Explanations (LIME)

We employ LIME to illustrate the features of the images that the models predominantly utilized for classification. An example of such an image is presented in Fig. 3.10. This figure shows the region where the model with EfficientNetB0 feature extractor primarily focuses to classify the image.

- **Original Image (Fig. 3.8a):** The first sub-image presents the original X-ray of a knee joint, which serves as the input to the EfficientNetB0 model. The X-ray displays key structural elements of the knee, such as the bones and the joint area. This is the primary image used by the model to assess the presence of osteoporosis, a condition that affects bone density and can be diagnosed through changes visible in such medical images.

- **Highlighted Focus Area (Fig. 3.8b):** The second sub-image shows the regions of the original X-ray that the model considers most relevant for its prediction, delineated by a yellow contour. These highlighted areas represent the regions of the knee joint that the model focuses on when making a classification decision. In this case, the model primarily concentrates on the joint

and surrounding bone regions, which are critical indicators in diagnosing osteoporosis. The yellow boundaries offer a clear visual representation of the model's selective attention within the entire X-ray image.

- **Focus Area of the Original Image (Fig. 3.8c):** The third sub-image isolates the regions of the X-ray that were highlighted by the model in the previous sub-image. This focused area is extracted from the original image, allowing for a more direct view of the regions the model deems important. By concentrating on this region, the model minimizes distractions from irrelevant parts of the image and hones in on specific areas where osteoporosis-related changes are more likely to occur. This visualization emphasizes how the deep learning model selectively filters the input image to make more accurate predictions.

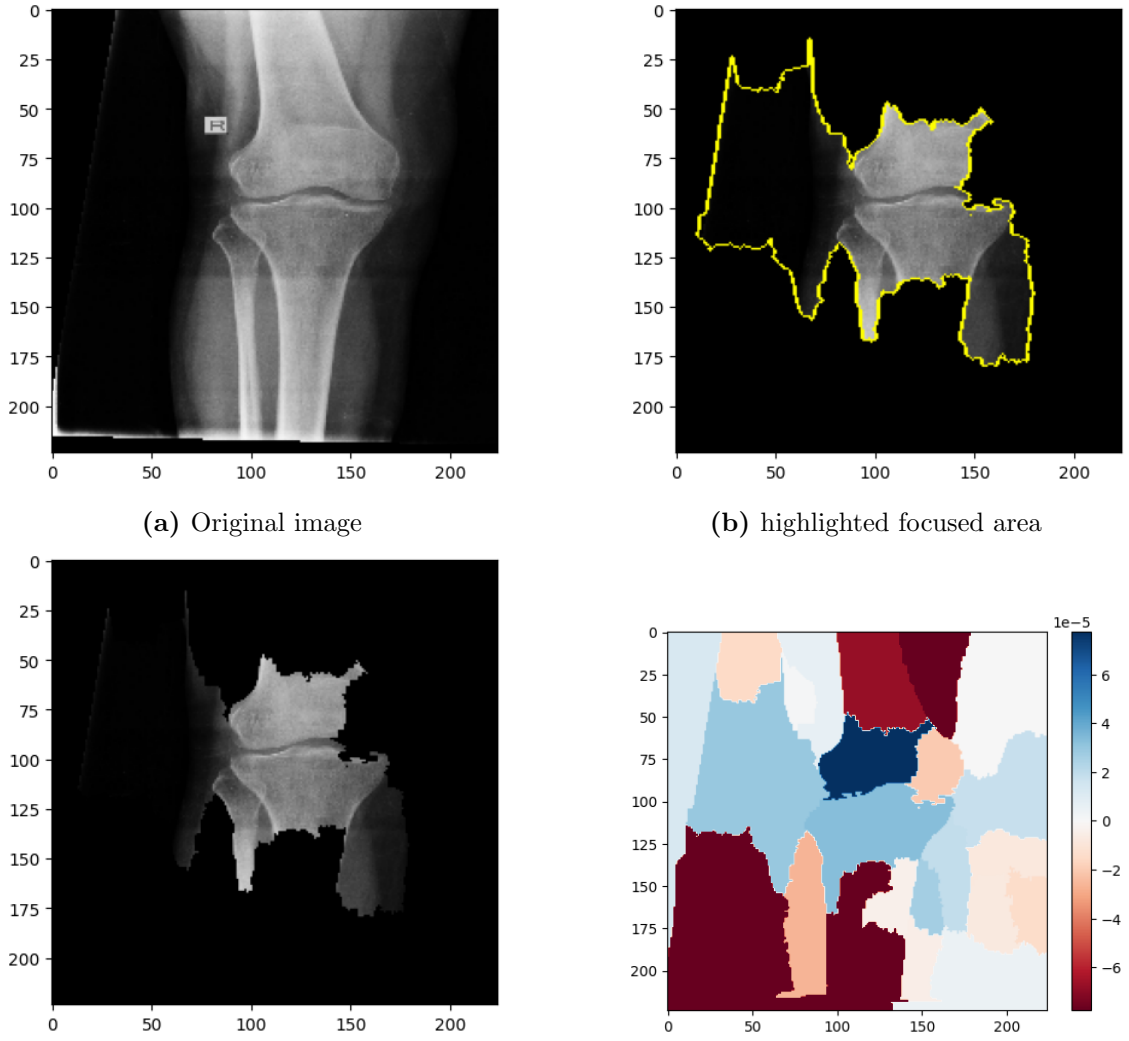
- **Heatmap (Fig. 3.8d):** The final sub-image provides a heatmap that visually represents the distribution of the model's attention across different regions of the X-ray image. In this heatmap, the color gradient ranges from deep blue to red. In this case, the blue areas indicate regions of higher focus by the model, meaning that these areas are considered most relevant in the prediction of osteoporosis. Conversely, the red areas represent regions of lesser importance, where the model allocates minimal attention. This color-coded representation provides a clear understanding of how the model emphasizes certain regions, particularly those in blue, for making its classification decisions. The heatmap thus highlights the spatial regions that play a critical role in the model's prediction process.

Same process is followed for MobileNet V3 for an example image in fig. 3.9.

3.5.2 Gradient-weighted Class Activation Mapping (Grad-Cam)

Fig. 3.10 depicts the focused regions of the knee joint as highlighted by the models EfficientNetB0 and MobileNetV3 using Grad-CAM (Gradient-weighted Class Activation Mapping). Grad-CAM is a popular interpretability technique that provides visual explanations for deep learning models, allowing us to observe which areas of an image are most influential in the model's decision-making process. The sub-images (a), (b), (c), and (d) demonstrate how both models concentrate on certain areas of the X-ray for the classification task, which in this context is likely related to osteoporosis detection. These visualizations are crucial for validating and understanding the models' behavior in medical imaging.

- **Knee Region Focused by EfficientNetB0:** The first sub-image shows the region of the knee X-ray that the EfficientNetB0 model focuses on. The heatmap overlay uses a color gradient, where warmer colors such as red and orange signify areas of greater focus or importance for the model's decision. In this case, the model's attention is primarily directed to the knee joint, with emphasis on the surrounding bone structures. These areas are crucial for detecting changes in bone density and structure, which are characteristic of osteoporosis. The use of Grad-CAM allows us to visualize how EfficientNetB0



(a) Original image

(b) highlighted focused area

(c) Focus area of the original image

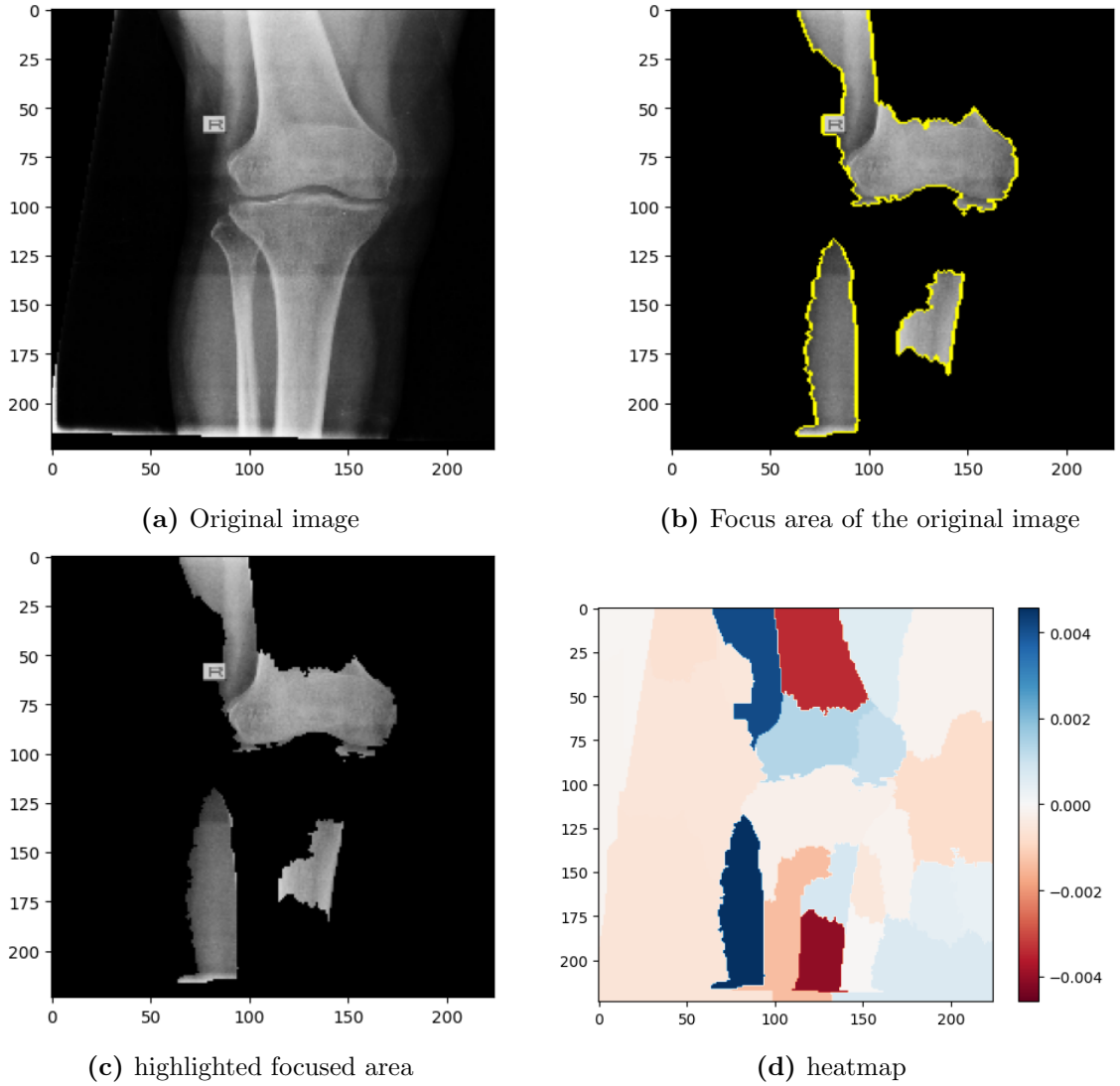
(d) Heatmap

Figure 3.8: Focused area of an image with EfficientNetB0.

assigns higher relevance to the central joint region in its decision-making process.

- **Knee Region of 2nd Knee X-Ray Image Focused by EfficientNetB0:** The second sub-image provides focus areas identified by EfficientNetB0 of another knee x-ray image, further validating its consistency in targeting key anatomical structures of the knee. Similar to the previous sub-image, the concentration of warm colors in the joint and surrounding bone regions suggests that these areas play a significant role in the model's classification output. The focus on these regions reaffirms the model's capability to hone in on clinically significant areas when analyzing the X-ray for potential signs of osteoporosis.

- **Knee Region Focused by MobileNetV3:** The third sub-image presents the Grad-CAM visualization for the MobileNetV3 model applied to the knee X-ray. MobileNetV3 also concentrates its attention on the knee joint, similar to EfficientNetB0, but with slight variations in the exact regions of focus. The warmer colors, particularly around the lower section of the knee joint,



(a) Original image

(b) Focus area of the original image

(c) highlighted focused area

(d) heatmap

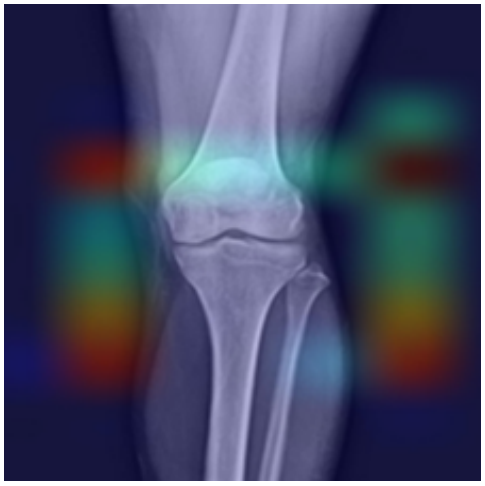
Figure 3.9: Focused area of an image with MobileNet V3 on the same image as EfficienNet B0.

indicate the areas the model finds most relevant. These differences in focus can provide insights into how different models approach the same task, offering complementary perspectives in detecting osteoporosis or other knee-related abnormalities.

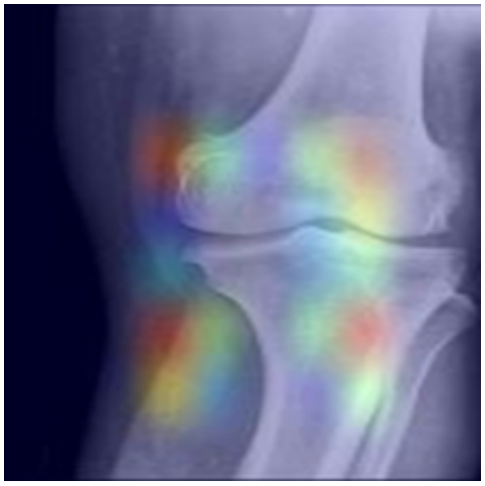
- **Knee Region of 2nd Knee X-Ray Image Focused by MobileNetV3):** This sub-image on the focus areas of 2nd knee x-ray image identified by MobileNetV3. The heatmap again highlights the central areas of the knee joint, though the focus appears more distributed along the joint axis compared to the EfficientNetB0 model. MobileNetV3, like EfficientNetB0, prioritizes the joint and adjacent bone structures, reinforcing the significance of these areas in medical imaging analysis for osteoporosis detection.

Fig. 3.10 highlights the importance of using Grad-CAM to interpret deep learning models like EfficientNetB0 and MobileNetV3, particularly in the context of medical imaging. Both models focus on critical anatomical regions such as the knee joint and surrounding bone areas, which are key indicators in diagnosing conditions like

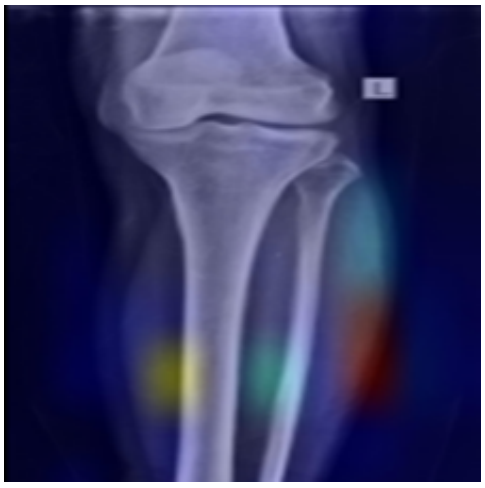
osteoporosis. The visualizations provided by Grad-CAM enable us to understand the inner workings of these models, providing transparency in the decision-making process. This is especially valuable in medical applications, where trust and interpretability are essential for ensuring that models focus on clinically relevant regions of interest.



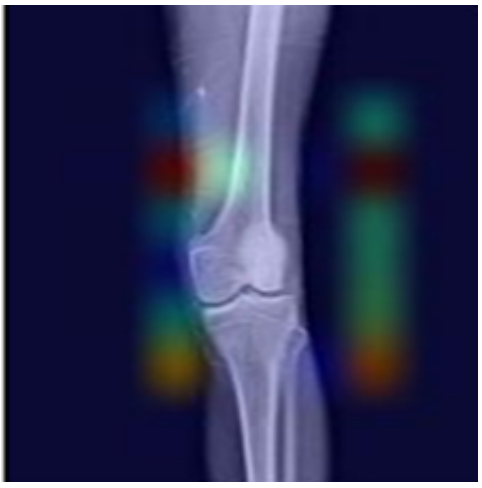
(a) Knee region focused by EfficientNet B0



(b) Knee region focused by EfficientNet B0



(c) Knee region focused by Mobilenet V3

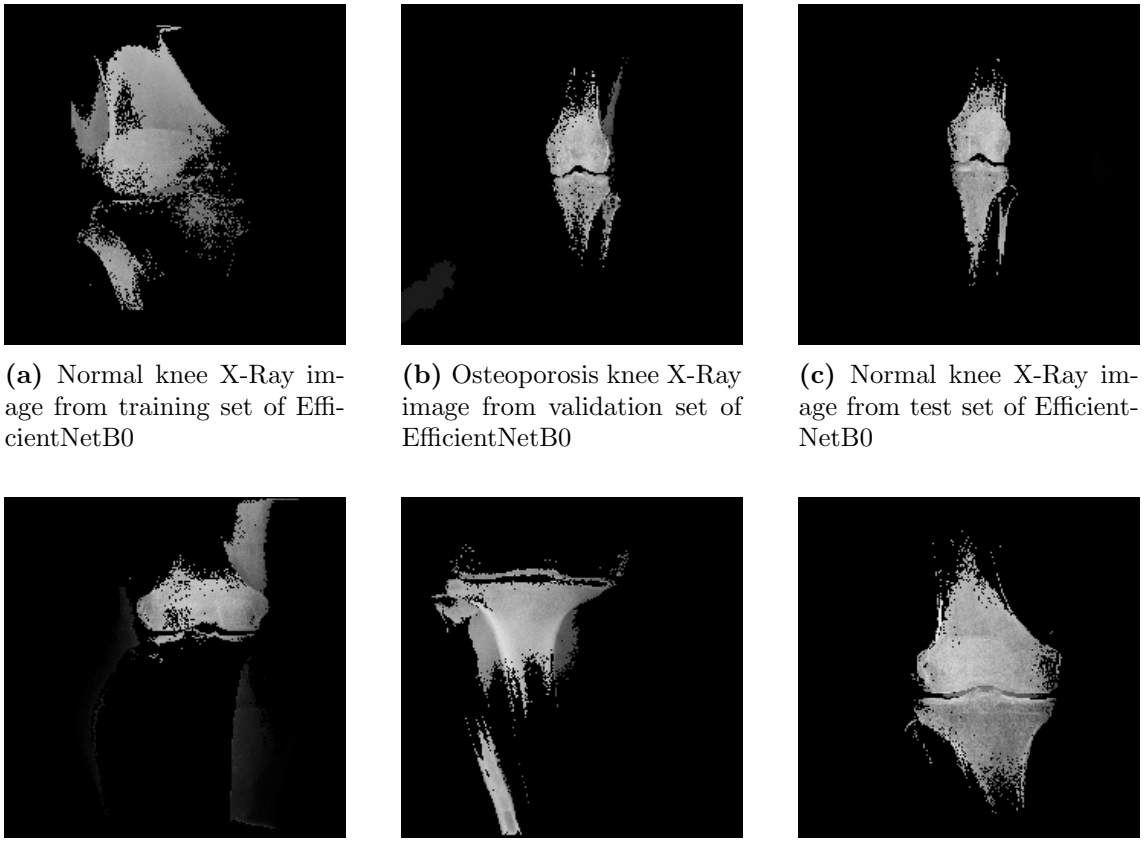


(d) Knee region focused by Mobilenet V3

Figure 3.10: Focused knee region by the models depicted through GradCam.

3.6 Dataset with Extracted Feature

We create two separate datasets by extracting key areas from the images for both models. We follow the same process to resize, split and change statistic parameters for the images of the proposed dataset as the previous dataset.
The training images of these datasets are augmented following the same procedure described in 3.2. Sample training, validation, and test images from this dataset are depicted in Fig. 3.11.



(a) Normal knee X-Ray image from training set of EfficientNetB0

(b) Osteoporosis knee X-Ray image from validation set of EfficientNetB0

(c) Normal knee X-Ray image from test set of EfficientNetB0

(d) Osteoporosis knee X-Ray image from training set of MobileNet V3

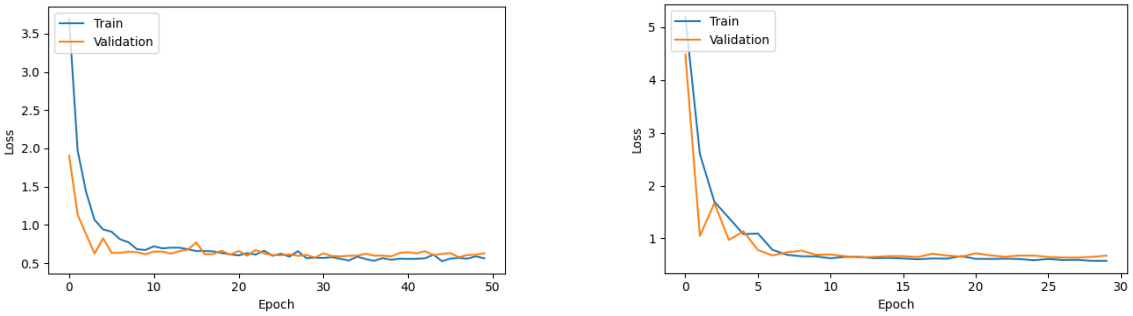
(e) Normal knee X-Ray image from validation set of MobileNet V3

(f) Osteoporosis knee X-Ray image from test set of MobileNet V3

Figure 3.11: Sample feature extracted images from proposed dataset.

3.7 Model Training Improvement

To train the model on our proposed dataset, we follow the same procedure as described in 3.4. In this case, the training set of the proposed dataset are utilized instead of the original one. The training loss of the models on this proposed training set is described in Fig. 3.12.
On the proposed dataset, the training loss curve of EfficientNet B0 drops more smoothly and reaches below 0.5, which is significantly lower than in Fig. 3.7. This indicates a much more efficient learning process. The validation loss also shows improvement, stabilizing closer to the training loss with fewer fluctuations. The validation loss stabilizes at around 0.75, a much closer convergence compared to the old dataset. This suggests that the model generalizes better on the proposed dataset with reduced overfitting. The EfficientNet B0 model performs significantly better on the proposed dataset (Fig. 3.12). The overall loss values are lower, and the gap between training and validation losses has reduced, indicating better generalization and less overfitting. The smoother curves suggest the model converged more efficiently, likely due to better data quality or a more appropriate dataset for training.
In Fig. 3.11f, MobileNet V3 shows significant improvement. The training loss decreases more consistently, reaching below 1.0 early on, and stabilizes near 0.5 by the end of the 30 epochs. The validation loss also improves significantly compared to Figure 3.7. It initially fluctuates, but stabilizes much faster around 0.75, which is much closer to the training loss than in the previous figure. The reduction in fluctuation and the smaller gap between training and validation losses in Fig. 3.11f indicates improved generalization on the proposed dataset. This means the model learned more effectively and managed to reduce overfitting. The MobileNet V3 model also shows substantial performance improvement on the proposed dataset. Both training and validation losses are significantly lower, and the validation loss curve stabilizes more quickly with fewer fluctuations. The close convergence between the two curves in Fig. 3.11f indicates that MobileNet V3 has better generalization ability on the proposed dataset, whereas it struggled more with the old dataset where the larger gap and higher validation loss suggested potential overfitting.



(a) EfficientNetB0 training loss

(b) MobileNet V3 training loss

Figure 3.12: Training loss of models on proposed dataset

Chapter 4

Results

We conduct a comparative analysis of our models using various quantitative evaluation metrics, including accuracy, precision, recall, and F1 score. Initially, we assess the models' performance using the original dataset.
Subsequently, after generating a second dataset consisting of the focused regions, we conduct a second round of quantitative analysis. The comparative performance of our models between the original and proposed datasets is showcased in Table 4.1. In both cases, the test sets are used to measure performance.
Accuracy: Accuracy is a metric that quantifies the proportion of correctly predicted instances (both true positives and true negatives) relative to the total number of instances. It provides an overall measure of the model's effectiveness shown in Equation 4.1.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{4.1}$$

Precision: Precision measures the proportion of true positive predictions among all positive predictions made by the model. It reflects the accuracy of the positive predictions shown in Equation 4.2.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{4.2}$$

Recall: Recall quantifies the proportion of actual positive instances that are correctly identified by the model. It indicates the model's ability to capture all positive instances shown in Equation 4.3.

$$\text{Recall} = \frac{TP}{TP + FN} \tag{4.3}$$

F1 Score: The F1 score is the harmonic mean of precision and recall, providing a single metric that balances both aspects shown in Equation 4.4. It is particularly useful when dealing with imbalanced datasets, as it gives equal weight to precision and recall.

$$\text{F1 Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4.4}$$

Here, TP = True positive, TN = True negative, FP = False positive, FN = False negative.

Table 4.1: Performance metrics for MobileNet v3 and EfficientNetB0 on the original and proposed datasets

Dataset	Model	Accuracy	Precision	Recall	F1 Score
Original	MobileNet v3	0.77	0.7083	0.9189	0.8000
	EfficientNetB0	0.77	0.7632	0.7838	0.7733
Proposed	MobileNet v3	0.8243	0.8529	0.7838	0.8169
	EfficientNetB0	0.9189	0.9444	0.9189	0.9315

The comparative analysis of MobileNet v3 and EfficientNetB0 using various quantitative evaluation metrics on both the original and proposed datasets are shown in Table 4.1.

For the original dataset, both MobileNet v3 and EfficientNetB0 achieve an accuracy of 0.77. However, their performance varies across other metrics. EfficientNetB0 exhibits a precision of 0.7632, which is approximately 7.7% higher than MobileNet v3's precision of 0.7083. Conversely, MobileNet v3 demonstrates a superior recall of 0.9189, which is about 17.2% higher than EfficientNetB0's recall of 0.7838. In terms of the F1 score, MobileNet v3 has a slight edge with an F1 score of 0.8000, representing a 3.4% improvement over EfficientNetB0's F1 score of 0.7733. These results suggest that while EfficientNetB0 has a higher precision, MobileNet v3 is more effective in correctly identifying positive cases, as indicated by its higher recall and F1 score.

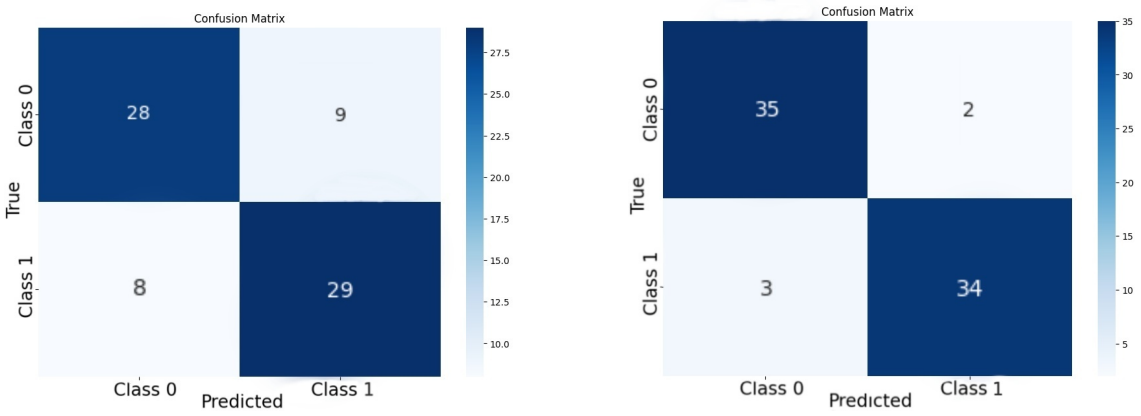
On the proposed dataset, both models show significant improvement across all metrics. MobileNet v3 achieves an accuracy of 0.8243, a precision of 0.8529, a recall of 0.7838, and an F1 score of 0.8169. EfficientNetB0, however, outperforms MobileNet v3 with an accuracy of 0.9189, a precision of 0.9444, a recall of 0.9189, and an F1 score of 0.9315. The performance improvement of EfficientNetB0 is notable, with its accuracy being approximately 11.5% higher than that of MobileNet v3. Additionally, EfficientNetB0's precision is 10.7% higher, its recall is 17.2% higher, and its F1 score is 14% higher than MobileNet v3's. These results indicate that EfficientNetB0 not only maintains a high level of precision but also significantly improves its recall, leading to a substantial increase in the F1 score.

However, there was a slight decrease of MobileNet's recall on proposed dataset. The decrease in MobileNet's recall on the proposed dataset can be attributed to several factors. First, the proposed dataset was generated by extracting key areas of the images for classification, and while this improves the model's precision by reducing false positives, it may have caused MobileNet to miss some positive cases that require a broader view of the image. This narrow focus might have led the model to overlook subtle or dispersed patterns indicative of osteoporosis. Additionally, MobileNet, being a lightweight architecture, may have overfitted to these extracted regions, reducing its ability to generalize and identify true positive cases that involve more complex or less obvious patterns outside the focused areas. Finally, compared to EfficientNetB0, MobileNet has fewer parameters and a lower capacity to learn intricate patterns from the feature-focused dataset, which could have contributed to its reduced recall performance. This suggests that while MobileNet is efficient, it might struggle to fully capture the complexities of the proposed dataset, particularly when fine-tuning its attention to critical features.

The comparative analysis reveals that EfficientNetB0 consistently outperforms MobileNet v3, especially on the proposed dataset where feature extraction focuses on the critical regions of the images. The enhanced performance of EfficientNetB0 can be attributed to its architecture, which is designed to capture intricate patterns in the data more effectively than MobileNet v3. The pre-processing step of refining the dataset to focus on regions of interest likely contributes to the improved performance of both models, allowing them to better utilize the most relevant features for classification.

The confusion matrix shown in Fig. 4.1a represents the EfficientNetB0 for the original dataset. In the confusion matrix, class 0 represents normal knee X-Ray images, and class 1 represents osteoporosis knee X-Ray images. The EfficientNetB0 correctly classifies 28 normal knee X-Ray images while misclassifying nine normal knee X-Ray images as osteoporosis images. Additionally, the model correctly classifies 29 osteoporosis knee X-Ray images and misclassifies eight osteoporosis knee X-Ray images as normal knee X-Ray images.

Fig. 4.1b. represents the confusion matrix for the proposed dataset. Here, EfficientNetB0 correctly classifies 35 normal knee X-Ray images (compared to 28 in the original test set) and misdiagnoses only two normal knee X-Ray images as osteoporosis (compared to nine in the original test set). Additionally, the model correctly classifies 34 osteoporosis knee X-Ray images and misclassifies three osteoporosis knee X-Ray images as normal knee X-Ray images (compared to 29 and eight, respectively, in the original test set).



(a) Confusion matrix of EfficientNetB0 on original dataset

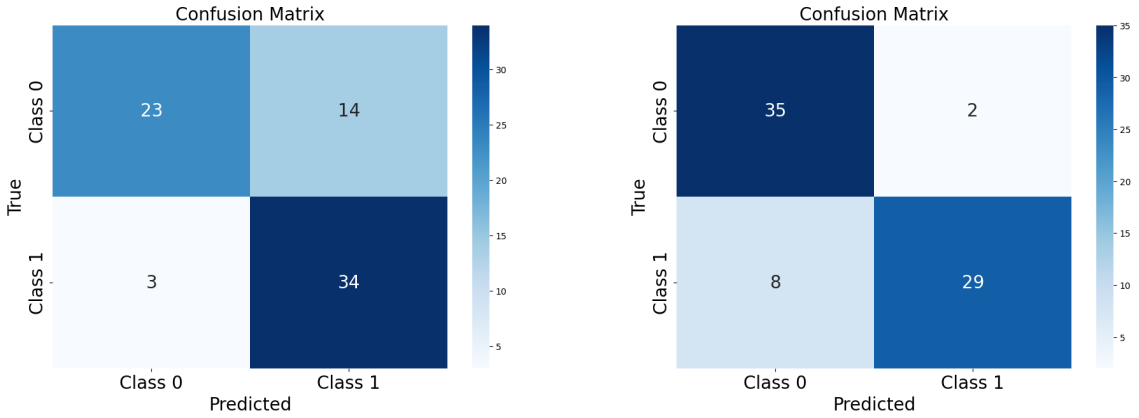
(b) Confusion matrix of EfficientNetB0 on proposed dataset

Figure 4.1: Confusion matrix of EfficientNetB0.

Fig. 4.2a. shows the confusion matrix for MobileNet V3 applied to the original dataset. In this matrix, class 0 corresponds to normal knee X-Ray images, while class 1 corresponds to osteoporosis knee X-Ray images. The MobileNet V3 correctly identifies 23 normal knee X-Ray images but misclassifies 14 normal images as osteoporosis. Furthermore, the model accurately classifies 34 osteoporosis knee X-Ray images, with only three being incorrectly identified as normal.

Fig. 4.2b. illustrates the confusion matrix for the proposed dataset. For the proposed dataset, MobileNet V3 correctly classifies 32 normal knee X-Ray images (up from 23 in the original test set) and misclassifies only five normal knee X-Ray images as osteoporosis (compared to 14 in the original test set). Additionally, the model

correctly identifies 29 osteoporosis knee X-Ray images and misclassifies eight osteoporosis images as normal (compared to 34 correct and three incorrect classifications in the original test set).



(a) Confusion matrix of MobileNet V3 on original dataset

(b) Confusion matrix of MobileNet V3 on proposed dataset

Figure 4.2: Confusion matrix of MobileNet V3.

In conclusion, while both MobileNet v3 and EfficientNetB0 exhibited competent performance on the original dataset, the refined dataset allowed both models to achieve better results. EfficientNetB0 demonstrated superior performance across all metrics, making it the preferable choice for this classification task. These findings underscore the importance of both model selection and data preprocessing in enhancing the performance of deep learning models in medical imaging tasks.

Chapter 5

Consclusion

This study presented a comparative analysis of two deep learning models—MobileNet V3 and EfficientNet B0—for detecting osteoporosis from knee X-ray images. While both models exhibited satisfactory performance on the original dataset, the results significantly improved when evaluated on a refined dataset highlighting key features using XAI. EfficientNet B0 outperformed MobileNet V3 across all metrics, particularly excelling in precision, recall, and F1 score. The enhanced performance of EfficientNet B0 underscores its ability to capture intricate patterns within the dataset, making it a superior choice for this classification task. These findings highlight the importance of advanced deep learning architectures and feature-focused data preprocessing in improving the accuracy of medical image classification. The implementation of such AI-driven diagnostic tools can offer more reliable, efficient, and consistent osteoporosis detection, ultimately enhancing patient care by reducing human error and improving early diagnosis.

Bibliography

[1] W. Peck, "Consensus development conference: Diagnosis, prophylaxis, and treatment of osteoporosis," *American Journal of Medicine*, vol. 94, no. 6, pp. 646–650, 1993.

[2] J. Compston, "Use of alendronate and risk of incident atrial fibrillation in women," *The Journal of clinical endocrinology & metabolism*, vol. 91, no. 9, pp. 3084–3090, 2006.

[3] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*, Ieee, 2009, pp. 248–255.

[4] E. M. Dennison, J. E. Compston, J. Flahive, *et al.*, "Effect of co-morbidities on fracture risk: Findings from the global longitudinal study of osteoporosis in women (glow)," *Bone*, vol. 50, no. 6, pp. 1288–1293, 2012.

[5] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[6] C. J. Rosen, "The epidemiology and pathogenesis of osteoporosis," 2015.

[7] M. T. Ribeiro, S. Singh, and C. Guestrin, """ why should i trust you?" explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 2016, pp. 1135–1144.

[8] A. G. Howard, M. Zhu, B. Chen, *et al.*, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, 2017.

[9] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 618–626.

[10] M.-T. Puth, M. Klaschik, M. Schmid, K. Weckbecker, and E. Münster, "Prevalence and comorbidity of osteoporosis–a cross-sectional analysis on 10,660 adults aged 50 years and older in germany," *BMC musculoskeletal disorders*, vol. 19, pp. 1–8, 2018.

[11] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *International conference on machine learning*, PMLR, 2019, pp. 6105–6114.

[12] M. A. Clynes, N. C. Harvey, E. M. Curtis, N. R. Fuggle, E. M. Dennison, and C. Cooper, "The epidemiology of osteoporosis," *British medical bulletin*, vol. 133, no. 1, pp. 105–117, 2020.

[13] F. Zhuang, Z. Qi, K. Duan, *et al.*, "A comprehensive survey on transfer learning," *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43–76, 2020.

[14] Y. Gao, S. Patil, and J. Jia, "The development of molecular biology of osteoporosis," *International journal of molecular sciences*, vol. 22, no. 15, p. 8182, 2021.

[15] R. Jang, J. H. Choi, N. Kim, J. S. Chang, P. W. Yoon, and C.-H. Kim, "Prediction of osteoporosis from simple hip radiography using deep learning algorithm," *Scientific reports*, vol. 11, no. 1, p. 19 997, 2021.

[16] H. Li, Z. Xiao, L. D. Quarles, and W. Li, "Osteoporosis: Mechanism, molecular target and current status on drug development," *Current medicinal chemistry*, vol. 28, no. 8, pp. 1489–1507, 2021.

[17] A. N. Zereen, S. Corraya, M. N. Dailey, and M. Ekpanyapong, "Two-stage facial mask detection model for indoor environments," in *Proceedings of International Conference on Trends in Computational and Cognitive Engineering: Proceedings of TCCE 2020*, Springer, 2021, pp. 591–601.

[18] U. B. Abubakar, M. M. Boukar, and S. Adeshina, "Evaluation of parameter fine-tuning with transfer learning for osteoporosis classification in knee radiograph," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 8, 2022.

[19] A. Aibar-Almazán, A. Voltes-Martínez, Y. Castellote-Caballero, D. F. Afanador-Restrepo, M. d. C. Carcelén-Fraile, and E. López-Ruiz, "Current status of the diagnosis and management of osteoporosis," *International journal of molecular sciences*, vol. 23, no. 16, p. 9465, 2022.

[20] A. S. Hosain and M. G. R. Alam, "Demystifying hypothyroidism detection with extreme gradient boosting and explainable ai," in *2022 25th International Conference on Computer and Information Technology (ICCIT)*, IEEE, 2022, pp. 260–265.

[21] A. S. Hosain, M. Islam, M. H. K. Mehedi, I. E. Kabir, and Z. T. Khan, "Gastrointestinal disorder detection with a transformer based approach," in *2022 IEEE 13th annual information technology, electronics and mobile communication conference (IEMCON)*, IEEE, 2022, pp. 0280–0285.

[22] A. S. Hosain, M. H. K. Mehedi, T. J. Jerin, *et al.*, "Rice leaf disease detection with transfer learning approach," in *2022 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET)*, IEEE, 2022, pp. 1–6.

[23] A. S. Hosain, M. H. K. Mehedi, and I. E. Kabir, "Pconet: A convolutional neural network architecture to detect polycystic ovary syndrome (pcos) from ovarian ultrasound images," in *2022 International Conference on Engineering and Emerging Technologies (ICEET)*, IEEE, 2022, pp. 1–6.

[24] M. E. Klontzas, I. Stathis, K. Spanakis, A. H. Zibis, K. Marias, and A. H. Karantanas, "Deep learning for the differential diagnosis between transient osteoporosis and avascular necrosis of the hip," *Diagnostics*, vol. 12, no. 8, p. 1870, 2022.

[25] M. H. K. Mehedi, A. S. Hosain, S. Ahmed, *et al.*, "Plant leaf disease detection using transfer learning and explainable ai," in *2022 IEEE 13th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, IEEE, 2022, pp. 0166–0170.

[26] Y. Shen, X. Huang, J. Wu, *et al.*, "The global burden of osteoporosis, low bone mass, and its related fracture in 204 countries and territories, 1990-2019," *Frontiers in Endocrinology*, vol. 13, p. 882 241, 2022.

[27] B. Adejuyigbe, J. Kallini, D. Chiou, and J. R. Kallini, "Osteoporosis: Molecular pathology, diagnostics, and therapeutics," *International journal of molecular sciences*, vol. 24, no. 19, p. 14 583, 2023.

[28] M. A. Kassem, S. M. Naguib, H. M. Hamza, M. M. Fouda, M. K. Saleh, and K. M. Hosny, "Explainable transfer learning-based deep learning model for pelvis fracture detection," *International Journal of Intelligent Systems*, vol. 2023, no. 1, p. 3 281 998, 2023.

[29] V. V. Khanna, K. Chadaga, N. Sampathila, *et al.*, "A decision support system for osteoporosis risk prediction using machine learning and explainable artificial intelligence," *Heliyon*, vol. 9, no. 12, 2023.

[30] S. Kumar, P. Goswami, and S. Batra, "Fuzzy rank-based ensemble model for accurate diagnosis of osteoporosis in knee radiographs," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 4, 2023.

[31] B. Suh, H. Yu, H. Kim, *et al.*, "Interpretable deep-learning approaches for osteoporosis risk screening and individualized feature analysis using large population-based data: Model development and performance evaluation," *Journal of medical Internet research*, vol. 25, e40179, 2023.

[32] I. M. Wani and S. Arora, "Osteoporosis diagnosis in knee x-rays by transfer learning based on convolution neural network," *Multimedia Tools and Applications*, vol. 82, no. 9, pp. 14 193–14 217, 2023.

[33] M. A. Rasool, S. Ahmed, U. Sabina, and T. K. Whangbo, "Konet: Towards a weighted ensemble learning model for knee osteoporosis classification," *IEEE Access*, 2024.

[34] *Osteoporosis — kaggle.com*, https://www.kaggle.com/datasets/mrmann007/osteoporosis, [Accessed 15-07-2024].