Application of Deep Neural Networks in Diagnosis of Osteoporosis and Osteopenia

Submitted in partial fulfilment of the requirement for the award of the degree of

BACHELOR OF TECHNOLOGY
IN
COMPUTER SCIENCE & ENGINEERING

2022-2023

FACULTY OF ENGINEERING AND TECHNOLOGY

SRM UNIVERSITY, DELHI-NCR, SONEPAT, HARYANA-131029

ACADEMIC YEAR: 2022-2023

SHREYA – 11021210078 ,
DHRUV CHOUDHURY – 11021210069 ,
RAGHAV SANT – 11021210054

ACKNOWLEDGMENT

I would like to express my sincere gratitude and thanks to my supervisor, Dr. Purushottam, Assistant Professor, Faculty of Engineering and Technology for his guidance, encouragement, and constructive feedback throughout this project. His expertise and insights have been invaluable in shaping and improving my work.
I would like to extend my gratitude to Dr. Puneet Goswami, Head of Department (CSE), for providing us wonderful opportunity and facilities for working on this project.
I am indebted to my colleagues and friends who have supported me throughout this journey.
Finally, we would like to express our deepest appreciation to our family for their love, patience, and encouragement. They have been our source of strength and motivation throughout this challenging process. I dedicate this work to them.

_____

Dr. Purushottam
B. Tech (CSE), IV semester
Department of Computer Science,
SRM University, Sonipat

CERTIFICATE

This is to certify, that the research paper entitled “APPLICATION OF DEEP NEURAL NETWORKS IN DIAGNOSIS OF OSTEOPOROSIS AND OSTEOPENIA” submitted by Dhruv Choudhury, Shreya, Raghav Sant is an outcome of our independent and original work. We have duly acknowledged all the sources from which the ideas and extracts have been taken. The project is free from any plagiarism and has not been submitted elsewhere.

The paper presents a novel approach to solve a problem in the field of study. The objectives of the paper are to state the research question, review the literature, propose a methodology, conduct experiments, analyze the results, and draw conclusions. The paper also discusses the implications and limitations of the findings, and suggests directions for future research.

The paper has been prepared according to the guidelines and format of the Dr. Purushottam Kumar.

_____

Dhruv Choudhury - 11021210069
Shreya - 11021210078
Raghav Sant - 11021210054

B. Tech (CSE), IV semester
Department of Computer Science
SRM University, Sonipat

TABLE OF CONTENTS

ABSTRACT

This project aims to explore the application of deep neural networks in diagnosing osteoporosis and osteopenia using medical imaging data. Deep neural networks, a subset of artificial intelligence, excel in recognizing patterns and abnormalities within these images, enabling them to identify potential diseases or conditions with high accuracy. Their ability to analyze vast amounts of data and extract complex patterns has paved the way for significant advancements in disease diagnosis and treatment.

Osteoporosis and osteopenia are bone disorders characterized by decreased bone density and strength, leading to an increased risk of fractures. Osteoporosis and osteopenia are prevalent bone diseases that pose significant health risks, particularly in aging populations. The diagnosis is done on the basis of evaluation of bone mineral density (BMD).

However, it is important to note that while deep neural networks have shown immense promise in medical disease diagnosis, they are not without limitations. Issues such as the lack of interpretability, potential biases in training data, and the need for extensive datasets for training are areas that researchers and developers are actively addressing.

The dataset has 3 classes: normal, osteopenia, osteopenia into which all the X-ray images are classified into and the model is applied and confusion matrix is plot for the same. Four deep neural network models have been trained on this dataset (VGG16, VGG19, DenseNet169,ConvNeXTiny). The Final Result suggests ConvNeXTiny as the best model for diagnosing with approximately 65% accuracy score.

The X-Rays are analyzed on the basis of difference in the opacity of the bones due to varying bone mineral density (BMD) as compared to normal healthy bones along with thinning of cortical bone in osteopenia and osteoporotic patients.

The “Final_Result” file attached along shows the accumulated predicted results presented by all the models applied for diagnosing and classifying into the different classes.

INTRODUCTION

Osteopenia and osteoporosis are significant health concerns worldwide, particularly in aging populations. These conditions affect millions of individuals, predominantly women, and can have severe consequences if left undiagnosed and untreated.

According to the International Osteoporosis Foundation (IOF), osteoporosis affects an estimated 200 million people globally. This number is expected to rise with the aging population and changing lifestyle factors. Osteopenia, which represents a precursor to osteoporosis, is even more prevalent and affects a substantial portion of the population.

By utilizing neural networks, healthcare professionals can benefit from more precise and efficient diagnostic tools. These networks can process a wide range of input data and provide objective assessments, assisting doctors in making informed decisions. They can also help overcome limitations in human interpretation by incorporating complex interactions among different variables.

However, it is important to note that the implementation of deep neural networks in the diagnosis of osteoporosis and osteopenia is still an evolving field. Challenges include the need for large and diverse datasets for training, addressing potential biases in the data, and ensuring the interpretability of the models' decision-making process.

This research paper presents an extension of the pioneering work conducted by the Lublin University of Technology, aiming to explore the application of deep neural networks, including VGG16, VGG19, DenseNet169, and ConvNeXTiny, for diagnosing osteoporosis and osteopenia. The performance of these models will be evaluated based on accuracy scores and F1-scores, providing valuable insights into identifying the most effective network architecture for this purpose.

The evaluation will primarily focus on two key metrics: accuracy score and F1-score. The accuracy score represents the overall correctness of the model's predictions, while the F1-score considers both precision and recall to provide a comprehensive assessment of the model's diagnostic performance. By comparing these metrics across the four deep neural networks, we aim to identify the architecture that exhibits the highest level of accuracy and diagnostic reliability.

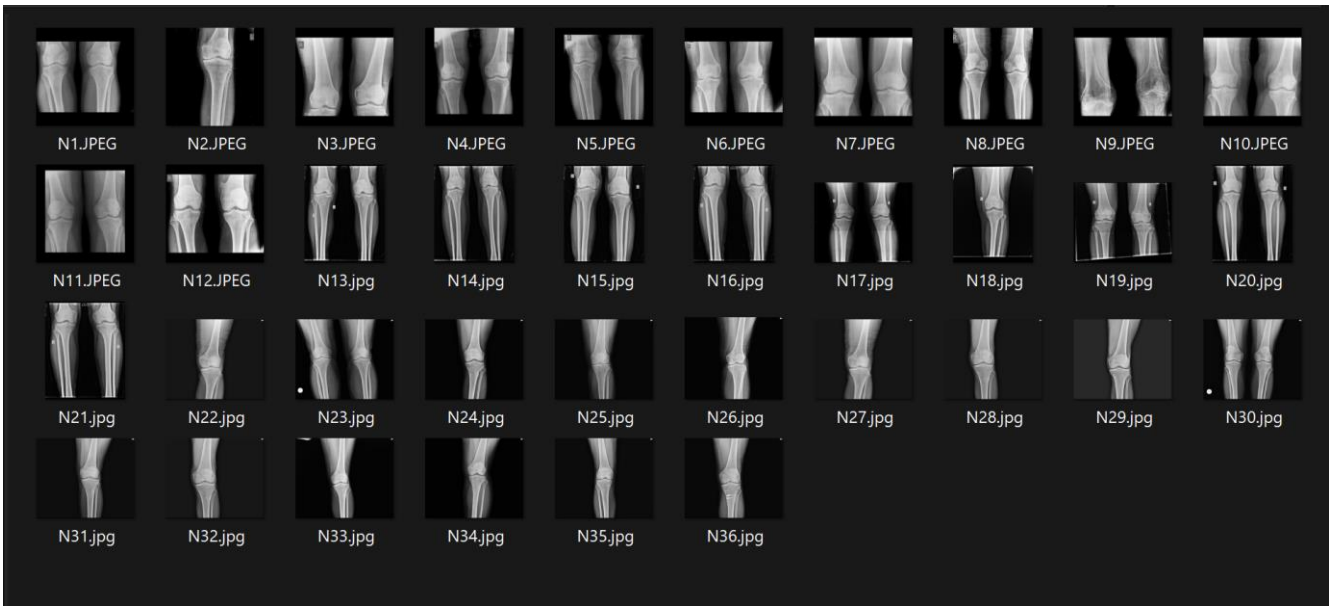
Overall, this project presents a promising avenue for leveraging deep neural networks to enhance the diagnosis of osteoporosis and osteopenia. By harnessing the power of artificial intelligence and medical imaging, we strive to contribute to the advancement of bone disease management and improve the quality of care for individuals at risk of these conditions.

DATASET DESCRIPTION

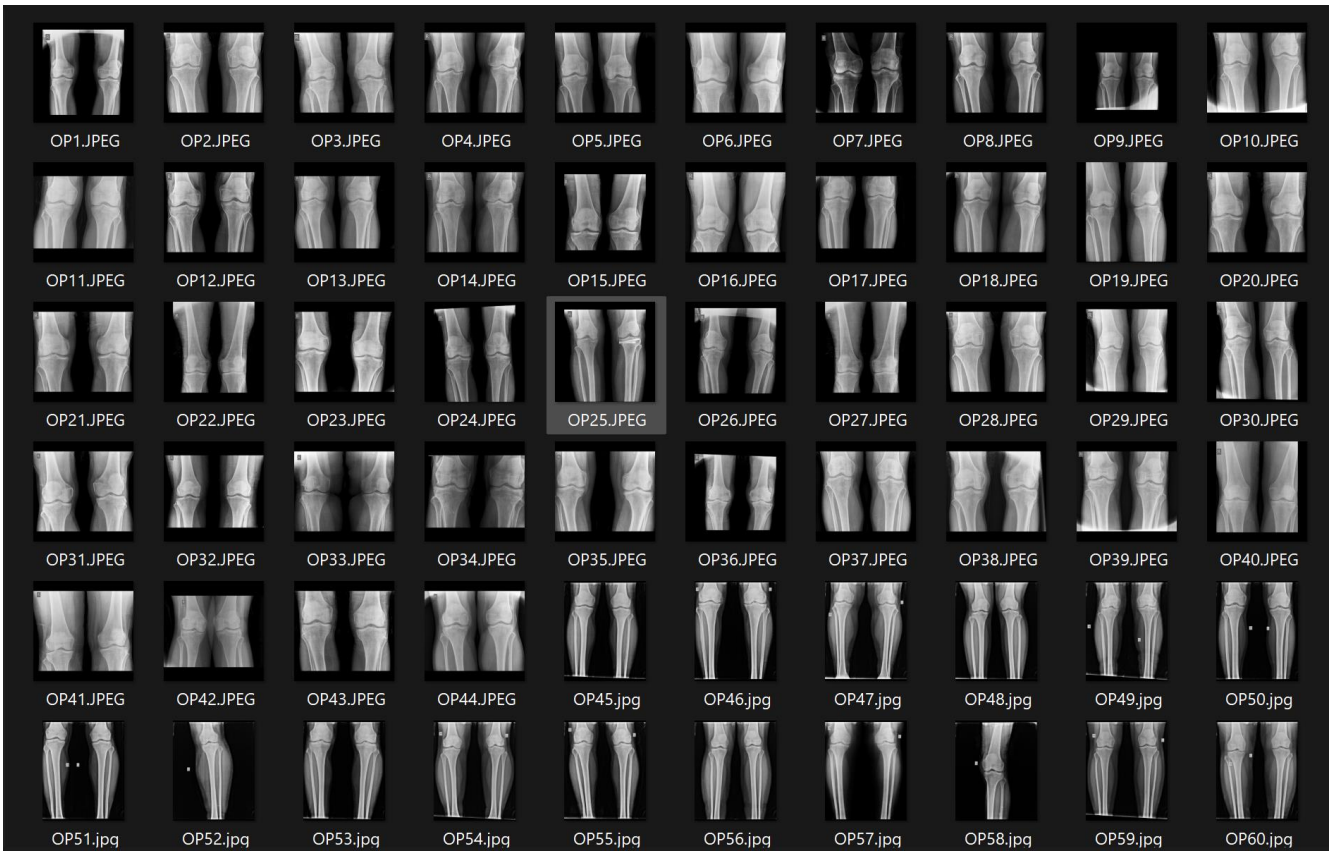
The dataset comprises of 3 classes: normal (healthy), osteopenia, osteoporosis onto which various models i.e., VGG16, VGG19, DenseNet169, ConvNeXTiny). These X-Ray bone images have been retrieved from XSITRAY database (a database for detection of Osteoporosis conditions). The dataset used for this project consists of knee X-rays collected from 52 females and 26 males in total 78 subjects.

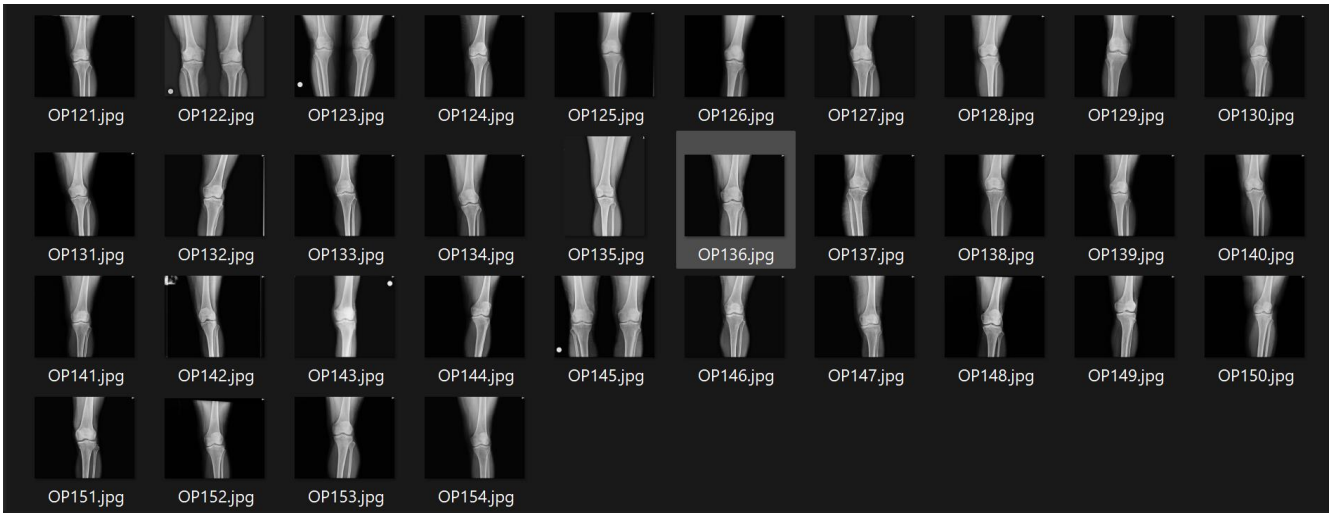
DATASET IMAGES

CLASS 1: NORMAL



CLASS 2: OSTEOPENIA





CLASS 3: OSTEOPOROSIS

METHODS AND MATERIALS

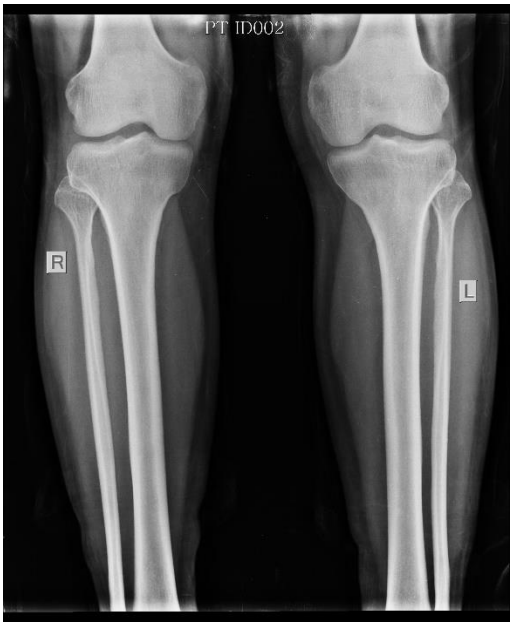
4.1 MATERIALS

The X-ray shows the bones, including the femur (thigh bone), tibia (shin bone), fibula (calf bone) and patella (kneecap), as well as the surrounding tissues. The X-ray may show joint space narrowing, bone spurs, or signs of inflammation. Soft tissues, such as ligaments and tendons, are not clearly visible on X-ray. These X-rays have been extracted from the XSITRAY database which have been collected from a research foundation center. The X-ray images belong to 52 female and 26 male subjects. The images are segregated into classes 1. Normal 2. Osteopenia 3. Osteoporosis. Each class has around 100-150 images.

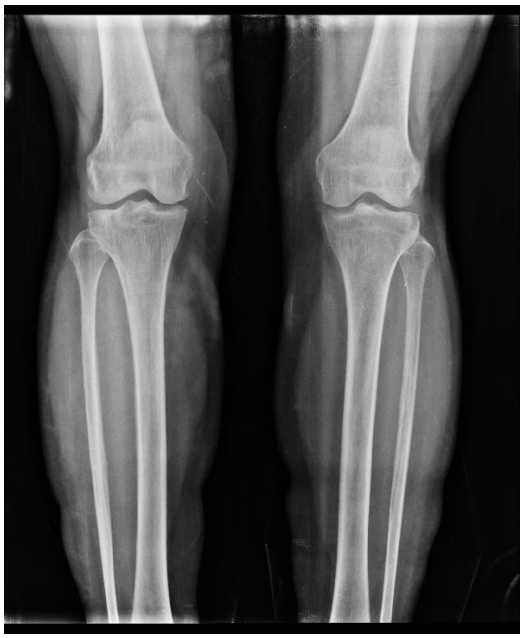
Osteoporotic bones appear lighter or more translucent on X-ray images due to a decrease in mineral content and bone mass. In contrast, x-rays of healthy bones show a denser and more opaque appearance.

Osteoporosis can cause changes in bone structure, such as thinning of cortical bone (outer layer of bone) and increased curvature of the spine (kyphosis). These structural changes may be visible on x-rays as thinning of bone cortex or an altered spinal alignment. Healthy bones, on the other hand, exhibit a more normal bone structure without such alterations.

Sample Images from 2 classes:



N21.jpg (NORMAL CLASS)



OS19.jpg (OSTEOPOROSIS CLASS)

4.2 METHODS

4.2.1. TRANSFER TECHNIQUES

Transfer learning is a machine learning technique that utilizes knowledge gained from pre-trained models to solve new tasks. Instead of starting from scratch, a pre-trained model, typically trained on a large dataset, is employed as a foundation. The learned features and representations are transferred to a new model, which is then fine-tuned on a smaller dataset specific to the target task. Transfer learning involves utilizing pre-trained models, such as InceptionV3, as a starting point and then customizing them for specific classification tasks.

4.2.2. INCEPTION V3 ARCHITECTURE

Inception-V3 is a deep convolutional neural network (CNN) architecture used for image classification and feature extraction tasks. Each inception module consists of multiple parallel convolutional layers with different filter sizes, allowing the network to capture both local and global information in the input image.

The key idea behind Inception-V3 is the concept of "inception modules." These modules are designed to efficiently capture features at different scales and spatial resolutions within the network.

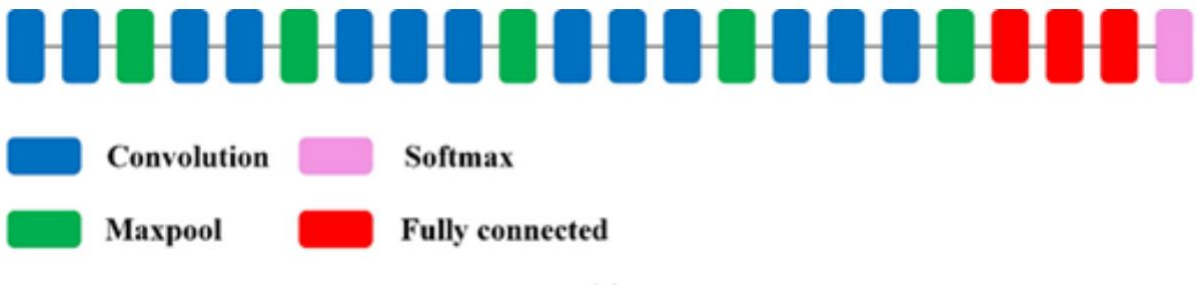
To adapt InceptionV3 for a specific classification task, we add additional layers on top of the pre-trained model. These new layers are designed to suit the unique requirements of the task at hand. These additional layers are trained on a smaller dataset specific to the task.

4.2.3. VGG16

The architecture consists of 16 layers, including 13 convolutional layers and 3 fully connected layers.

The input image is passed through a stack of convolutional layers with small 3x3 filters. These layers extract low-level features such as edges, textures, and basic shapes. The convolutional layers are followed by max pooling layers that reduce the spatial dimensions, allowing the network to focus on important information.

The extracted features are then passed through additional convolutional layers, deepening the network and capturing more abstract features. VGG16's depth plays a crucial role in learning intricate details and high-level representations. After the convolutional layers, the fully connected layers are responsible for the final classification. These layers receive the flattened features and apply a series of transformations, combining the learned representations to make predictions about the input image's class.

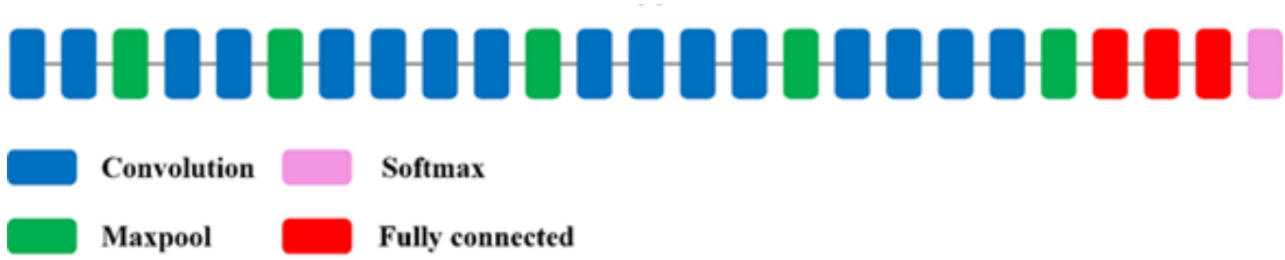


Structure of VGG16 network model

4.2.4. VGG19

The architecture consists of 16 layers, including 13 convolutional layers and 3 fully connected layers.

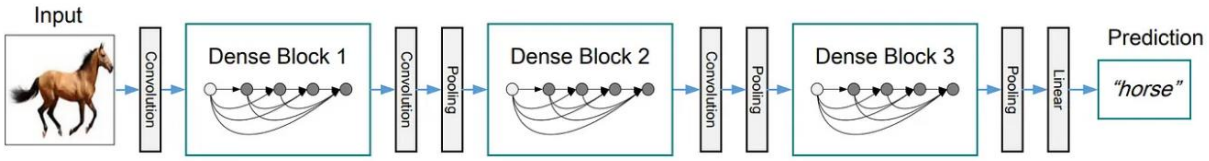
VGG19 expands on VGG16's architecture by adding more convolutional layers, further deepening the network for enhanced feature extraction and understanding. The depth of VGG19 enhances its capability to model fine details and high-level representations. The additional layers compared to VGG16 allow for a richer and more comprehensive understanding of the input images.



Structure of VGG19 network model

4.2.5. DenseNet169

DenseNet169 introduces the concept of dense connectivity, where each layer is connected to every other layer in a dense manner. DenseNet169 is composed of multiple dense blocks. Each dense block consists of a series of densely connected layers. Within a dense block, the output feature maps from all previous layers are concatenated and passed as input to the subsequent layers.



Structure of DenseNet169 network model

4.2.6.ConvNeXTiny

ConvNeXTiny is a compact convolutional neural network (CNN) architecture designed for efficient image classification on resource-constrained devices. It utilizes lightweight convolutional layers, depth wise separable convolutions, and downsampling techniques to reduce computational complexity while preserving important features. Despite its compact size, it maintains a satisfactory level of performance and accuracy. Downsampling helps in compressing the information and capturing the most salient features while reducing computational requirements.

DATA PREPARATION

The code snippet uses the glob module to retrieve the file paths of X-ray images corresponding to different bone health conditions (normal, osteopenia, and osteoporosis). The images are loaded using mpimg.imread() and stored in separate lists: normal_images, osteopenia_images, and osteoporosis_images. The matplotlib library is then utilized to visualize the first image from each category. The first image from each category is displayed using matplotlib's imshow() function to verify the loading process and assess the visual characteristics of the images.

DATA AUGMENTATION AND GENERATOR SETUP

Data Augmentation and Data Generation: We employ the ImageDataGenerator class from the Keras library to perform data augmentation. The augmentation techniques applied include rescaling, shearing, zooming, and horizontal flipping. The train_datagen.flow_from_directory() function is used to set up a data generator for the training dataset, specifying parameters such as the target image size, batch size, class mode (binary in this case), and subset for training. Similarly, a data generator is created for the validation dataset using the same function.

Training Data Generator Setup: The train_datagen.flow_from_directory() function is used to set up a data generator for the training dataset. The function takes the data directory path, target image size, batch size, class mode (binary in this case), and a subset parameter for training. It automatically generates batches of augmented images during model training.

Validation Data Generator Setup: Similarly, the train_datagen.flow_from_directory() function is employed to create a data generator for the validation dataset. The shuffle parameter is set to False to maintain the order of the images during validation. This generator is used to evaluate the model's performance on unseen data

Data preprocessing and augmentation techniques applied to the image dataset before feeding it into the deep learning model. It ensures that the model receives a varied and augmented dataset during training, enhancing its ability to learn and generalize from the data.

DATA GENERATOR

It involves using the ImageDataGenerator class from the Keras library to perform data augmentation and generate batches of image data for training and validation.

```
train_datagen = ImageDataGenerator(rescale=1./255,
    shear_range=0.2,
    zoom_range=0.2,
    horizontal_flip=True,
    validation_split=0.3)

train_generator = train_datagen.flow_from_directory(
    dataDirectory,
    target_size=(IMG_H, IMG_W),
    batch_size=BATCH_SIZE,
    class_mode='binary',
    subset='training')

validation_generator = train_datagen.flow_from_directory(
   dataDirectory,
    target_size=(IMG_H, IMG_W),
    batch_size=BATCH_SIZE,
    class_mode='binary',
    shuffle= False,
    subset='validation')
```

```
Found 169 images belonging to 3 classes.
Found 70 images belonging to 3 classes.
```

The validation_split parameter specifies that 30% of the data will be used for validation, while the remaining 70% will be used for training.

LOAD DATA FROM DIRECTORY

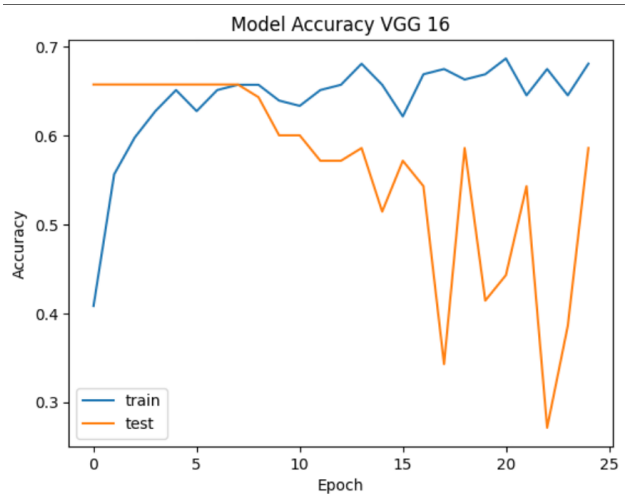
```
In train_generator
normal :         26
osteopenia :     108
osteoporosis :   35

In validation_generator
normal :         10
osteopenia :     46
osteoporosis :   14
```

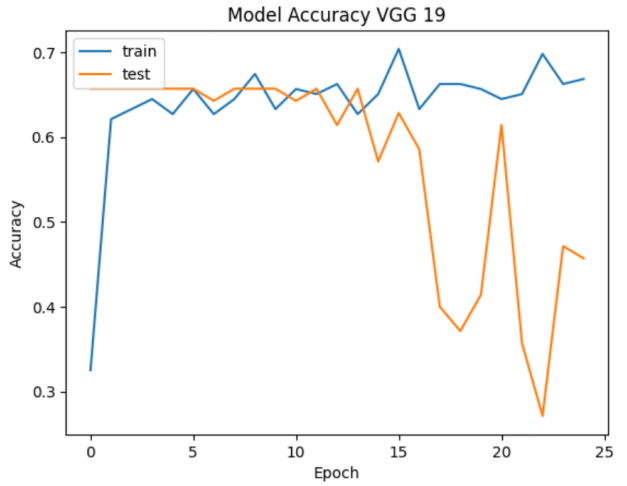
The train_generator and validation_generator are created by calling the flow_from_directory method on the train_datagen.

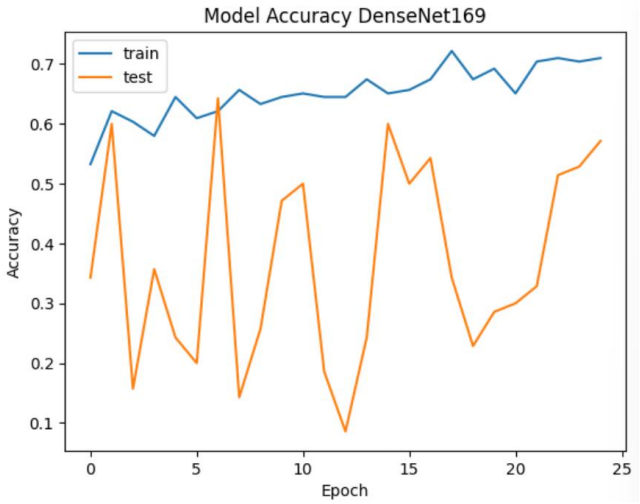
Class Distribution Analysis : This segment provides insights into the class distribution within the training and validation data generators. For each bone health condition, the code counts the number of instances belonging to that class within the generators. The results are printed, allowing researchers to assess the balance and representation of different classes in the datasets.

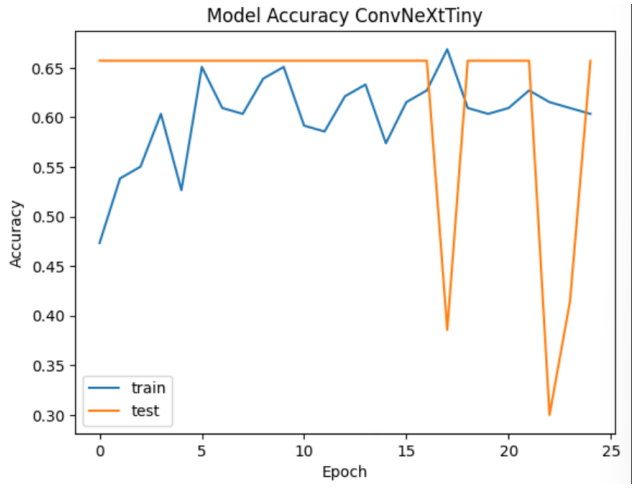
MODEL ACCURACY:

The model accuracy graph serves as a crucial visualization in evaluating the performance of different models in a research paper. It presents a comparison of the accuracy achieved by various models across different experiments or conditions. The graph showcases the effectiveness of each model in accurately predicting the target variable and demonstrates the model's ability to generalize well to unseen data.









RESULTS:

Accuracy and F1 score are both metrics used to evaluate the performance of classification models, but they measure different aspects. Accuracy alone can be misleading when the dataset is imbalanced, meaning that some classes have significantly more samples than others.

The F1 score balances these two measures and provides a single metric that takes into account both precision and recall. It is particularly useful when the dataset is imbalanced or when both precision and recall are important.

ConvNeXTiny give the best accuracy score and f1-score among all the trained models with a percentage of 65%

ACCURACY:

```
Accuracy scores :

VGG 16 :
0.5714285714285714
VGG 19:
0.4714285714285714
DenseNet169:
0.5714285714285714
ConvNeXtTiny:
0.6571428571428571
```

F1-SCORES:

The F1 score is a statistical measure used in machine learning and information retrieval to evaluate the performance of a classification model. It combines precision (the ability to correctly identify positive samples) and recall (the ability to retrieve all positive samples) into a single metric. The F1 score balances the trade-off between precision and recall, providing an overall assessment of a model's accuracy.

VGG16

```
              precision    recall  f1-score   support

      normal       0.00      0.00      0.00        10
  osteopenia       0.62      0.87      0.73        46
osteoporosis       0.00      0.00      0.00        14

    accuracy                           0.57        70
   macro avg       0.21      0.29      0.24        70
weighted avg       0.41      0.57      0.48        70

Recall score:
0.5714285714285714
Precision score:
0.4107142857142857
```

VGG19

```
              precision    recall  f1-score   support

      normal       0.00      0.00      0.00        10
  osteopenia       0.58      0.72      0.64        46
osteoporosis       0.00      0.00      0.00        14

    accuracy                           0.47        70
   macro avg       0.19      0.24      0.21        70
weighted avg       0.38      0.47      0.42        70

Recall score:
0.4714285714285714
Precision score:
0.3804511278195489
```

DenseNet169

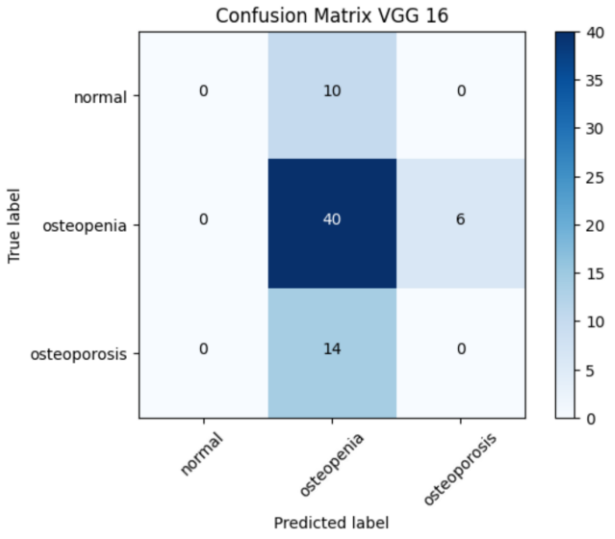
```
              precision    recall  f1-score   support

      normal       0.00      0.00      0.00        10
  osteopenia       0.62      0.87      0.73        46
osteoporosis       0.00      0.00      0.00        14

    accuracy                           0.57        70
   macro avg       0.21      0.29      0.24        70
weighted avg       0.41      0.57      0.48        70

Recall score:
0.5714285714285714
Precision score:
0.4107142857142857
```

ConvNeXTiny

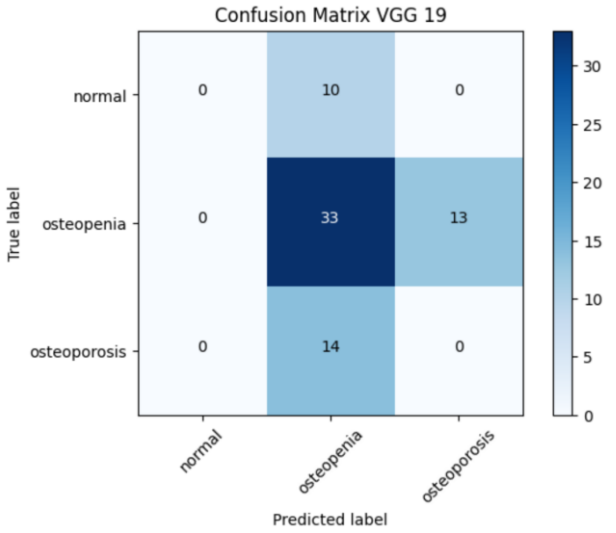
```
              precision    recall  f1-score   support

      normal       0.00      0.00      0.00        10
  osteopenia       0.66      1.00      0.79        46
osteoporosis       0.00      0.00      0.00        14

    accuracy                           0.66        70
   macro avg       0.22      0.33      0.26        70
weighted avg       0.43      0.66      0.52        70

Recall score:
0.6571428571428571
Precision score:
0.43183673469387757
```

CONFUSION MATRIX OF ALL MODELS:

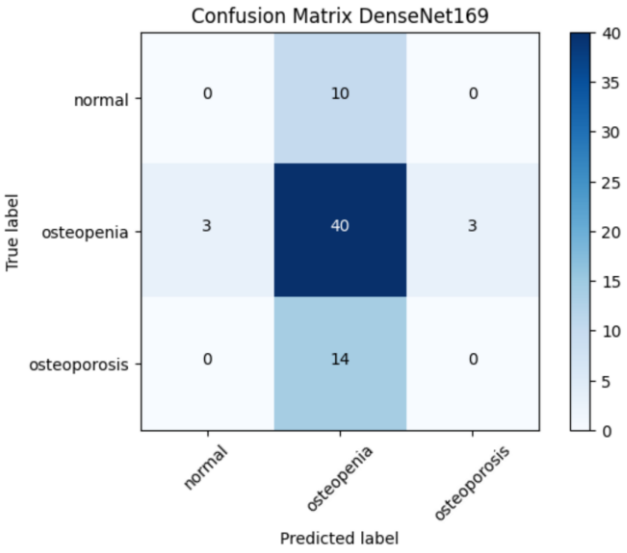
VGG16



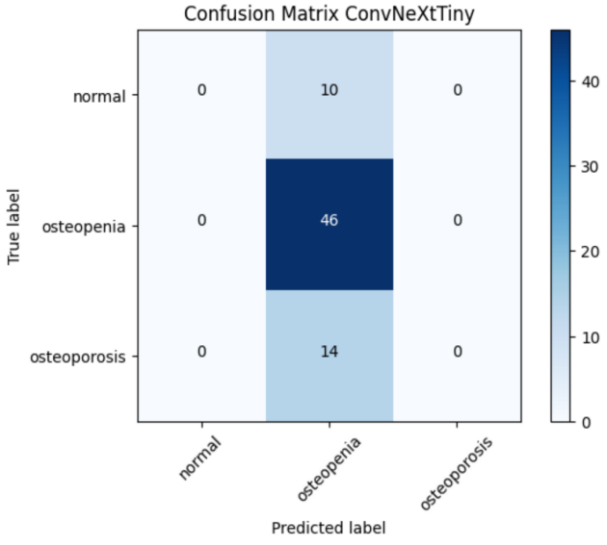
VGG19



DenseNet169



ConvNeXTiny



CONCLUSION:

In conclusion, this research paper explored the application of deep neural networks in the context of image classification. Four different models, namely VGG16, VGG19, DenseNet169, and ConvNeXTiny, were implemented and evaluated for their performance.

We were successfully able to build and train the 4 mentioned models on our dataset of knew X-Rays and finally compare all trained models through the F1-Score to find the best trained model. The results indicate that ConvNeXTiny, a compact architecture, achieved the best overall performance in terms of accuracy and efficiency. This model was able to achieve an accuracy of 65% which is much better than any other model

The study also utilized Inception V3 as the base architecture for fine-tuning and transfer learning. This approach allowed leveraging the pre-trained weights and learning capabilities of Inception V3 to adapt the models to the specific classification task at hand.

Challenges of Interpretability still exist for ordinary people/ medical professionals to understand the output of the machine learning models along with limitations of small dataset for train a deep neural network

BIBLIOGRAPHY

The XSITRAY database can be observed and downloaded at the institutional web address: http://www.sethu.ac.in/ XSITRAY/.

Dzierżak, R.; Omiotek, Z. Application of Deep Convolutional Neural Networks in the Diagnosis of Osteoporosis. *Sensors* **2022**, *22*, 8189. https://doi.org/10.3390/s22218189

1.) Johnell, O., & Kanis, J. A. (2006). An estimate of the worldwide prevalence and disability associated with osteoporotic fractures. Osteoporosis International, 17(12), 1726-1733.

2.) NIH Consensus Development Panel on Osteoporosis Prevention, Diagnosis, and Therapy. (2001). Osteoporosis prevention, diagnosis, and therapy. JAMA, 285(6), 785-795.

3.) Cummings, S. R., Black, D. M., Rubin, S. M. (1989). Lifetime risks of hip, Colles', or vertebral fracture and coronary heart disease among white postmenopausal women. Archives of Internal Medicine, 149(11), 2445-2448.

4.) Kanis, J. A., Bergstrom, F., De Laet, C., Johansson, H., Johnell, O., Jonsson, B., Oden, A., Zethraeus, N., & Pfleger, B. (2005). Assessment of fracture risk. Osteoporosis International, 16(6), 581-589.

5.) NIH Consensus Development Panel on Osteoporosis Prevention, Diagnosis, and Therapy. (2001). Diagnosis of osteoporosis in postmenopausal women: 2001 clinical practice guidelines. Journal of the American Medical Association, 285(6), 785-795.

6.) Reginster, J. Y., Burlet, N., & Delmas, P. D. (2006). Evaluation of patients with osteoporosis. Best Practice & Research Clinical Rheumatology, 20(4), 807-826.

7.) Kanis, J. A., McCloskey, E. V., Johansson, H., Oden, A., Leslie, W. D. (2012). FRAX® with and without bone mineral density. Calcified Tissue International, 90(1), 1-13.

8.) Siris, E. S., Chen, Y. T., Abbott, T. A., Barrett-Connor, E., Miller, P. D., Wehren, L. E., Berger, M. L. (2004). Bone mineral density thresholds for pharmacological intervention to prevent fractures. Archives of Internal Medicine, 164(10), 1108-1112.

9.) Eastell, R., Reid, D. M., Compston, J., Cooper, C., Fogelman, I., Francis, R. M., Hosking, D. J., Purdie, D. W., Ralston, S. H., Reeve, J., Russell, R. G., Stevenson, J. C., & Torgerson, D. J. (2001). A UK Consensus Group on management of glucocorticoid-induced osteoporosis: an update. Journal of Internal Medicine, 250(6), 459-468.

10.) Kanis, J. A., Johansson, H., Oden, A., Johnell, O., de Laet, C., Melton, L. J., & Tenenhouse, A. (2004). A meta-analysis of prior corticosteroid use and fracture risk. Journal of Bone and Mineral Research, 19(6), 893-899.

11.) Sambrook, P., Cooper, C. (2006). Osteoporosis. Lancet, 367(9527), 2010-2018.