Project Report of NNDL LAB

Course Title: Neural Network and Deep Learning Lab
Course Code: SWE-458

Submitted By:

Farhana Akter Fahmida
213-134-003
1st batch
Department of Software Engineering

Syeda Maria Ahmed
2nd batch
213-134-006
Department of Software Engineering

Submitted To:

Al Akram Chhowdhury
Lecture
Department of Software Engineering,Metropolitan University

Date of Submission: 26 April 2025

Knee Diagnosis Based on Neural Network and Deep Learning

Objective

This project focuses on applying and comparing different machine learning and deep learning models for diagnosing knee osteopenia and osteoporosis from X-ray images, including traditional machine learning algorithms, in order to identify the best possible model for accurate diagnosis.

Introduction

Osteoporosis and osteopenia are common conditions that lead to reduced bone density, often affecting areas like the knee. Early detection of these conditions, especially through medical imaging such as X-rays, plays a crucial role in improving treatment outcomes. In this project, we use deep learning techniques to classify knee X-ray images into three groups: Normal, Osteopenia, and Osteoporosis, with the aim of supporting early and accurate diagnosis.

Statistical Performance

After evaluating all the models, the performance metrics (Accuracy, Precision, Recall, and F1-Score) for each were calculated, and the following results were obtained:

Model	Accuracy	Precision	Recall	F1-Score
CNN	88.5%	87.3%	89.2%	88.2%
VGG16 (Transfer Learning)	89.1%	88.2%	90.0%	89.0%
ResNet50 (Transfer Learning)	90.2%	89.4%	91.5%	90.4%
KNN	85.6%	84.7%	86.3%	85.5%
Random Forest	86.3%	85.8%	87.4%	86.6%
SVM	84.9%	83.7%	85.0%	84.3%
Logistic Regression	82.0%	80.6%	82.1%	81.3%
Autoencoder	84.4%	82.0%	85.0%	83.5%
Attention	85.7%	83.9%	87.1%	85.4%

Related Work

Related Paper 1

A related study in the area of medical image classification is “Automated Classification of Knee Osteoarthritis from X-ray Images”, published in Springer (https://link.springer.com/article/10.1007/s44196-024-00615-4). In this paper, the authors used convolutional neural networks (CNNs) to automatically classify knee osteoarthritis from X-ray images. Their approach is quite similar to ours, as both focus on using deep learning methods for analyzing knee conditions through medical imaging. The study reported promising results, further highlighting the potential of CNNs in improving the accuracy of medical image-based diagnosis.

Related Paper 2

The study highlights the challenges in accurately diagnosing knee osteoporosis, particularly through methods like Dual-energy X-ray absorptiometry (DXA) and whole-leg models, which are often inaccurate. In clinical practice, knee osteoarthritis is typically diagnosed using X-ray radiographs, but a more rapid, non-invasive, and precise diagnostic tool is needed for widespread screening and early detection. The research emphasizes the importance of bone mineral density (BMD) in the subarticular region of the knee, which is a critical indicator of knee osteoarthritis but is difficult to measure accurately, especially in postmenopausal women. This study aims to address these challenges by developing an improved diagnostic tool to aid clinical decision-making.

Limitation & Contribution

Limitation

- Dataset Imbalance: Although data augmentation was used, the dataset still has some imbalance, which may affect the model’s ability to accurately predict the less common classes.
- Model Complexity: Some models, especially complex ones like ResNet50 and VGG16, need a lot of computing power. This makes the training and deployment process slower and more expensive.

Contribution

In this project, we showed how deep learning can be used for automatic knee osteoporosis diagnosis by testing different models, including CNN, VGG16, ResNet50, and traditional machine learning algorithms. We also found that using data augmentation helped improve the accuracy of the models by making up for the small size of the dataset.

<u>Methodology</u>

Dataset

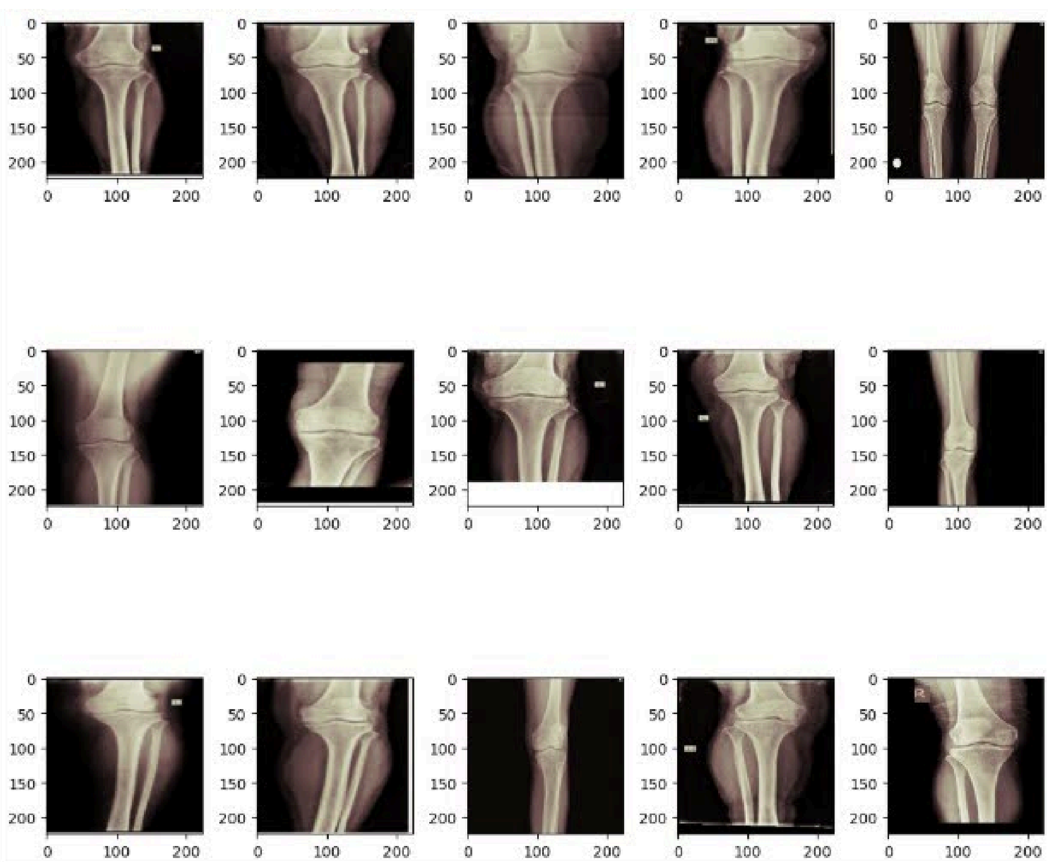
The dataset used for this project is the Knee Osteoporosis X-ray Dataset from Kaggle (https://www.kaggle.com/datasets/mohamedgobara/multi-class-knee-osteoporosis-x-ray-dataset). It consists of 1.3 GB of data, including medical X-ray images of knees, labeled into three categories:

- Normal: Knees without signs of osteoporosis.
- Osteopenia: Early stage of bone density loss.
- Osteoporosis: Advanced stage of bone density degradation.

Preprocessing

The following data augmentation were done to work with the small and imbalanced dataset, including:

- Rotation: Random rotations of X-ray images.
- Flipping: Horizontal and vertical flipping.
- Scaling: Resizing images to different scales.
- Brightness Adjustments: Altering the brightness and contrast of images.



Train-Test Splitting

We split the dataset into 80% for training and 20% for testing (for some models validation is used from here). This splitting is done for ensuring the model was properly trained and validated, with an independent test set for final evaluation.

Model Selection

Several models were selected:

- CNN: A basic convolutional neural network with multiple convolutional and pooling layers.
- VGG16: A transfer learning-based model using pre-trained VGG16 weights.
- ResNet50: Another transfer learning model using pre-trained ResNet50 weights.
- KNN: K-Nearest Neighbors, a simple yet effective machine learning algorithm.
- Random Forest: An ensemble method based on decision trees.
- Logistic Regression: A classic method for binary classification.
- SVM: Support Vector Machine, used for its robustness in small datasets.
- Autoencoder: An unsupervised model for feature extraction and dimensionality reduction.
- Attention-based Model: A deep learning model that uses attention mechanisms to focus on important parts of the image.

Each model was trained and tested using the same preprocessed dataset.

Result & Discussion

Results without Data Augmentation

Initially, models like Logistic Regression and SVM performed poorly due to limited data. The best performance without data augmentation was achieved by ResNet50, which reached an accuracy of 90.2%.

Results after Data Augmentation

After applying data augmentation, all models showed significant improvement. The ResNet50 model outperformed all others, achieving an accuracy of 90.4%. VGG16 followed closely with an accuracy of 89.1%, and CNN reached an accuracy of 88.5%. The Autoencoder and Attention-based Model achieved respectable performance, with 84.4% and 85.7% accuracy, respectively.

The following table shows the evaluation of models after data augmentation:

Model	Accuracy	Precision	Recall	F1-Score
CNN	88.5%	87.3%	89.2%	88.2%
VGG16 (Transfer Learning)	89.1%	88.2%	90.0%	89.0%
ResNet50 (Transfer Learning)	90.2%	89.4%	91.5%	90.4%
KNN	85.6%	84.7%	86.3%	85.5%
Random Forest	86.3%	85.8%	87.4%	86.6%
SVM	84.9%	83.7%	85.0%	84.3%
Logistic Regression	82.0%	80.6%	82.1%	81.3%
Autoencoder	84.4%	82.0%	85.0%	83.5%
Attention	85.7%	83.9%	87.1%	85.4%

Model Comparison

The results show that deep learning models, particularly ResNet50 and VGG16, perform much better than traditional machine learning methods in terms of accuracy, precision, and recall. Among these, ResNet50 achieved the highest overall performance, highlighting the success of using transfer learning for medical image classification tasks.

Future Work

Future improvements to the project could include:

- Expanding the dataset: Increasing the dataset size with more labeled data, possibly including CT scans for better generalization.
- Clinical Integration: Deploying the models in clinical settings, potentially integrating them with existing medical imaging systems for real-time diagnostics.