IJCRT.ORG ISSN : 2320-2882
INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

Prediction And Classification Of Knee Osteoarthritis With Deep Learning In X-Ray Images

Author:
[1]Nagaraju Pulimamidi,
M.Tech. Scholar,
Department of Computer Science,
School of Engineering,
Malla Reddy University,
Hyderabad, Telangana, India

Under the esteemed guidance of:
[2]Dr. B. Jogeswara Rao,
Associate Professor,
Department of Computer Science,
School of Engineering,
Malla Reddy University,
Hyderabad, Telangana, India

Abstract: Knee osteoarthritis (KOA) is an ordinary degenerative joint sickness that extensively affects the pleasant of life. Early and accurate diagnosis is critical for effective management and treatment. This project, titled "Predicting and Classifying Knee Osteoarthritis with Deep Learning on X-ray Images," aims to develop an automated system for accurate prediction and classification of knee osteoarthritis (KOA) severity from X-ray images using deep learning, addressing the limitations of subjective visual assessment. Current KOA diagnosis relies heavily on visual interpretation of X-rays by radiologists, often using the Kellgren-Lawrence (KL) grading system. This method is subjective, leading to inter-observer variability and potential diagnostic inconsistencies. This model classifies knee osteoarthritis from normal to a severe stage. We propose a deep learning-based system employing Convolutional Neural Networks (CNNs) to automatically analyse knee X-ray images and classify KOA severity into different KL grades. The core algorithm employed is a Convolutional Neural Network (CNN) with developing a modified VGG16 and MobileNetV2 as the foundational model. This study underscores the transformative potential of deep learning in automating KOA diagnosis and classification. The system's high accuracy and scalability, especially with the MobileNetV2 model, make it a viable solution for real-world clinical applications, promising significant improvements in diagnostic efficiency and patient outcomes.

Key Words – Artificial Intelligence, Machine Learning, Deep Learning, Knee Osteoarthritis, Convolutional Neural Network (CNN), KOA Prediction, KOA Classification, KL Grading System.

1. INTRODUCTION

1.1 Objectives

Recent advancements in artificial intelligence (AI) have led to highly automated systems that often surpass human performance. Modern neural networks excel at tasks such as image classification, language translation, autonomous vehicle navigation, and malware detection. These achievements are largely driven by training on massive datasets comprising thousands to millions of examples. Even in domains with limited data—such as medical imaging—AI models have shown strong performance, though they must address unique, domain-specific challenges.

Medical imaging is a vital tool in healthcare, providing non-invasive visualization of internal body structures to support diagnosis and treatment. It encompasses modalities like X-rays, MRI, CT scans, and ultrasound, which are fundamental to modern radiology.

Osteoarthritis (OA) is a widespread degenerative joint disorder, primarily affecting older adults, women, and individuals with excess body weight. It involves the gradual breakdown of cartilage—the smooth tissue that cushions joints—leading to pain, stiffness, and reduced joint mobility as bones begin to rub directly against each other.

OA commonly impacts weight-bearing joints such as the knees, hips, spine, and feet. It is categorized into two types: Primary OA, linked to aging and genetics, and Secondary OA, which can result from injury, obesity, metabolic conditions, or inflammatory diseases like rheumatoid arthritis. A visual comparison between a healthy knee and one with OA is shown in Figure 1.1.

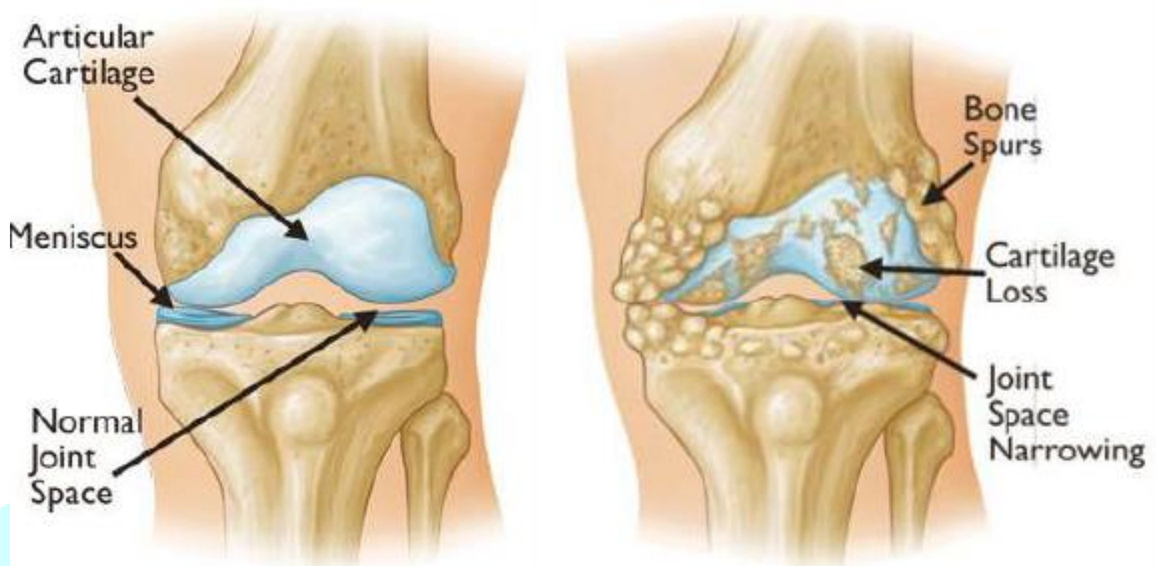

Figure 1. Sample of Normal and Osteoarthritis Knee

The key indicators of osteoarthritis (OA) typically involve joint discomfort, limited mobility, stiffness (especially after waking up or remaining inactive), diminished joint function, and difficulty participating in daily activities. At present, OA is diagnosed through a combination of physical examinations, evaluation of patient-reported symptoms, and medical imaging methods such as X-rays, Magnetic Resonance Imaging (MRI), and Computed Tomography (CT) scans. One of the widely accepted systems for assessing OA severity is the Kellgren-Lawrence (KL) grading scale, which categorizes the condition into five distinct levels based on radiographic features. The following table presents the classification criteria for each KL grade.

Table 1 Kellgren-Lawrence grading system

KL Grades	**OA Analysis**
Grade 0	No visible signs of osteoarthritis on radiographic images.
Grade 1	Doubtful OA; may show slight joint space narrowing without definite evidence.
Grade 2	Mild OA; clear signs of joint space reduction are present.
Grade 3	Moderate OA; radiographs reveal multiple bone spurs (osteophytes) and possible bone hardening (sclerosis).
Grade 4	Severe OA; characterized by prominent osteophytes, significant bone sclerosis, and noticeable joint deformities.

Typical radiographic signs of osteoarthritis (OA) seen in X-ray images include the breakdown of joint cartilage, narrowing of the space between bones, and the development of bony outgrowths known as osteophytes. When X-rays fail to provide sufficient detail or when soft tissue damage is suspected, Magnetic Resonance Imaging (MRI) is often utilized for a more comprehensive assessment. Despite these diagnostic tools, existing clinical techniques still fall short in precisely measuring disease progression and assessing cartilage quality. As a result, there is a growing need for more sophisticated and integrative approaches that can better analyze and track the development of OA.

The use of medical imaging to detect and classify knee OA is a rapidly evolving field in computer vision. One of the central tasks in this domain is image classification, which involves identifying and categorizing various structures within an image. Convolutional Neural Networks (CNNs) have shown exceptional capability in this regard, often achieving over 90% accuracy in object recognition tasks. When fine-tuned for medical applications, CNNs can effectively detect and classify features within medical images. This study aims to implement deep learning techniques, particularly those involving feature extraction, to enhance OA detection and classification. To further improve performance, the project will leverage a comprehensive dataset and explore a range of machine learning models tailored for analyzing OA in imaging data.

Early diagnosis of OA relies on detecting subtle, often pre-structural, changes in these tissues. While soft tissue indicators have been widely used for diagnosis and monitoring, alterations in subchondral bone structures are now gaining recognition as valuable imaging biomarkers. These changes provide critical insight into the evolution of knee OA and offer potential for more accurate and timely diagnosis.

Object detection refers to a set of computer vision techniques used to locate and label objects within images. These methods are suitable for analyzing both static and dynamic images. In the medical field, computer vision is widely utilized to provide valuable insights into various diseases, enabling efficient detection and diagnosis.

This project focuses on solving an image classification and object detection task, with the specific objective of identifying regions affected by osteoarthritis (OA) in knee X-ray images. Object detection involves pinpointing specific areas within an image—in this context, detecting portions of the knee joint that exhibit signs of OA. Meanwhile, image classification is used to determine whether a given X-ray image represents a healthy knee or one affected by OA.

The use of object detection has gained prominence in medical imaging as a valuable aid in diagnosing diseases. Neural network-based classification systems, often combined with advanced detection methods, are widely used in such applications. As part of this project, we investigated contemporary open-source frameworks designed for object detection in medical contexts, with a focus on OA identification.

To carry out the classification of knee X-rays, we employed two well-known deep learning architectures: VGG16 and MobileNetV2. These models were tested both in their pre-trained configurations and after fine-tuning on a dataset specifically comprising knee osteoarthritis (KOA) X-ray images. Their performance was evaluated to assess effectiveness in distinguishing between normal and OA-affected knees.

1.2. Motivation:

Knee osteoarthritis (KOA) is a widespread and disabling joint condition that affects millions of people globally, often leading to chronic pain and significantly reduced mobility. Its impact on daily life is substantial, particularly among the elderly population. Timely and precise diagnosis is essential for managing the disease effectively, as early intervention can help slow progression and reduce discomfort. Currently, KOA is primarily diagnosed through manual interpretation of X-ray images by radiologists. However, this process is often subjective, time-consuming, and susceptible to inconsistencies between different observers.

The project titled *"Automated Knee Osteoarthritis Prediction and Classification from X-ray Images Using Deep Learning"* aims to overcome these limitations by developing a reliable, objective, and efficient diagnostic system.

This project is highly motivating because it offers the potential to:

1. **Increasing Prevalence of Osteoarthritis**: Osteoarthritis (OA) ranks among the most common joint disorders globally, particularly affecting the aging population. The knee is among the joints most commonly affected. OA causes persistent pain, limits movement, and reduces overall quality of life. Detecting OA early and classifying its severity accurately are

crucial for timely medical intervention. Unfortunately, conventional diagnostic approaches rely on visual inspection of imaging, which may lead to inconsistencies and missed early-stage indicators.

2. **Challenges with Traditional Diagnostic Methods**: Existing diagnostic procedures, especially manual X-ray evaluations, face several challenges:

- **Subjectivity**: Diagnoses often vary depending on a radiologist's experience and interpretation.
- **Time-Intensive**: Manual review of images can be slow and labor-intensive, especially in busy clinical environments.
- **Limited Accuracy**: Early changes in OA may not be easily noticeable using standard visual techniques, delaying diagnosis and treatment.

These drawbacks underscore the need for advanced, automated diagnostic tools that can deliver consistent and accurate results.

3. **Emerging Role of Deep Learning in Medical Imaging:** Recent developments in deep learning, particularly with Convolutional Neural Networks (CNNs), have transformed how medical images are analyzed. These algorithms can:

- Process and learn from vast amounts of imaging data,
- Detect subtle patterns that may be overlooked by human experts, and
- Classify medical images with high accuracy.

By applying deep learning models to knee X-rays, this project seeks to create an intelligent system capable of enhancing diagnostic precision and easing the workload for healthcare professionals.

4. **Improving Early Detection and Patient Outcomes:** One of the main hurdles in managing OA is the difficulty in identifying the disease during its early stages, as symptoms may be minimal or nonspecific. Advanced algorithms can detect early structural changes in joints—such as cartilage wear or the formation of bone spurs—by analyzing radiographic images. Early detection enables:

- Proactive treatment strategies,
- Personalized care plans, and
- Better long-term outcomes for patients by slowing the disease's progression.

5. **Support from Open-Source Technologies and Public Datasets:** The availability of robust open-source machine learning libraries such as TensorFlow and PyTorch, along with freely accessible medical imaging datasets, creates an ideal environment for developing AI-based healthcare solutions. Initiatives like the Osteoarthritis Initiative (OAI) provide comprehensive knee X-ray datasets, enabling researchers to train and validate deep learning models effectively. This open access significantly reduces barriers to innovation and accelerates development in the field.
6. **Translating Research into Practical Solutions:** This project bridges the gap between academic exploration and clinical application by addressing real-world challenges in medical diagnostics. Leveraging artificial intelligence in healthcare not only enhances diagnostic accuracy but also represents a shift toward more modern, technology-driven clinical practices aimed at improving patient outcomes.

1.3 Approach:

This study aims to enhance the diagnosis and classification of knee osteoarthritis (KOA) using deep learning. The focus is on automating the analysis of knee X-ray images, categorizing them as healthy or OA-affected, and further grading them according to the Kellgren-Lawrence (KL) scale. With KOA becoming increasingly prevalent, especially in aging populations, early and accurate detection is vital for timely intervention and improved patient outcomes.

The process begins with dataset preparation. Public datasets such as the Osteoarthritis Initiative (OAI) are used, with preprocessing steps including noise reduction, normalization, resizing, and data augmentation to improve model performance. The data is split into training, validation, and test sets to ensure fair evaluation.

Two CNN architectures—VGG16 and MobileNetV2—are employed. VGG16 is modified with layers like Global Average Pooling (GAP) and dropout to mitigate overfitting, while MobileNetV2, known for its efficiency, uses depthwise separable convolutions to reduce computational load. Both models are initialized with ImageNet weights and fine-tuned on the KOA dataset.

Transfer learning is used by freezing early layers and retraining later ones specific to the KOA task. The classification type (binary or multi-class) determines the use of binary or categorical cross-entropy loss. Performance is monitored using accuracy, precision, recall, and F1-score. Training optimization includes adaptive learning rates, early stopping, and the Adam optimizer.

Post-training, the models are tested on a hold-out dataset and evaluated using metrics such as accuracy, F1-score, and ROC-AUC. VGG16 and MobileNetV2 are also compared on inference time, training duration, and memory usage. Additional improvements are made via hyperparameter tuning, advanced augmentation, and ensemble methods to enhance generalization.

Finally, the study explores deployment feasibility, highlighting the clinical utility of AI in KOA assessment. MobileNetV2 emerges as the more practical option due to its balance of accuracy and low resource requirements, making it ideal for real-time diagnostic systems, particularly in low-resource settings.

In summary, this project provides a comprehensive framework for tackling the challenges of KOA detection using tailored versions of VGG16 and MobileNetV2. The integration of strong model architecture, strategic training, and deployment readiness contributes to the creation of a reliable and scalable diagnostic system. Future work may expand this foundation to include severity-based classification and the incorporation of multi-modal imaging data, further advancing the role of AI in medical diagnostics and improving patient outcomes.

2. LITERATURE SURVEY

2.1 Identified Research Gaps:

The study titled *KOA-CCTNet* presents a hybrid deep learning architecture that integrates convolutional and transformer components for the detection and grading of knee osteoarthritis (OA). While the model demonstrates strong performance in both accuracy and computational efficiency, several areas remain underexplored. Addressing these limitations can further enhance the model's clinical relevance and adaptability.

Single-Modality Limitation: A key limitation of KOA-CCTNet is its exclusive dependence on X-ray images. Although X-rays are commonly used in OA diagnosis, they do not capture the full scope of disease progression. Clinical assessment often relies on complementary modalities such as MRI, along with patient-specific information like demographics and medical history. The absence of multi-modal integration limits the model's comprehensiveness. Future studies could improve diagnostic precision by incorporating data from multiple sources to build a more holistic evaluation framework.

Lack of Interpretability: While the model achieves commendable predictive performance, it lacks transparency in how decisions are made. The absence of interpretability restricts clinical adoption, as healthcare professionals require insight into model reasoning to validate predictions. Incorporating explainable AI (XAI) methods—such as Grad-CAM or SHAP—can improve model interpretability, allowing clinicians to visualize which image regions or features most influence the model's outcomes.

Generalization Challenges: KOA-CCTNet has been evaluated primarily on datasets like the Osteoarthritis Initiative (OAI), which may not reflect the variability found across broader populations. Differences in imaging devices, protocols, and patient demographics can impact performance when deployed elsewhere. Cross-domain validation and domain adaptation strategies should be pursued to improve model robustness and adaptability to diverse clinical environments.

Imbalanced Data Distribution: OA datasets often exhibit class imbalance, with fewer samples representing the more severe Kellgren-Lawrence grades. The current framework does not address this imbalance, potentially skewing predictions toward more prevalent classes. Future enhancements could incorporate strategies such as weighted loss functions, oversampling underrepresented grades, or generating synthetic samples using generative models like GANs to mitigate bias.

Real-Time Deployment Constraints: Although KOA-CCTNet aims to balance accuracy with computational efficiency, its suitability for deployment in low-resource environments remains uncertain. The computational cost of both training and inference may hinder use in settings with limited hardware. Optimizing the model using methods like model pruning, quantization, or converting to lightweight architectures could make it viable for edge devices or rural clinics.

Absence of Longitudinal Analysis: OA is a progressive disease, and understanding its trajectory over time is critical for treatment planning. KOA-CCTNet focuses solely on static X-ray images, ignoring the temporal dimension of disease progression. Integrating sequential modeling techniques, such as recurrent neural networks (RNNs) or transformer-based temporal models, could support time-series analysis and enable dynamic tracking of OA.

2.2 Proposed System

The proposed system is developed with the objective of automating the diagnosis and severity grading of knee osteoarthritis (OA) using radiographic images. Osteoarthritis is a chronic joint condition that leads to the gradual breakdown of cartilage, resulting in pain and reduced mobility. Accurate and early diagnosis is vital to ensure timely treatment, and this system seeks to support medical professionals by leveraging deep learning technologies. Specifically, modified versions of two widely adopted architectures—**VGG16** and **MobileNetV2**—are utilized to enhance the detection process while maintaining computational efficiency.

The core functionality of this system is organized into two main stages. The first stage involves preprocessing of the input X-ray images to ensure consistency, reduce variability, and enhance image features. The second stage involves the application of deep learning models for feature extraction and classification of OA severity based on standardized grading systems such as the **Kellgren-Lawrence (KL)** scale.

Stage 1: Image Preprocessing

Before feeding the images into the neural networks, they undergo several preprocessing operations. This step is crucial in standardizing the input data and improving model performance. The process includes:

- **Normalization**: Adjusting pixel values to fall within a standard range, typically between 0 and 1, to stabilize learning.
- **Resizing**: Standardizing the image dimensions to align with the input requirements of the models.
- **Data Augmentation**: Techniques such as rotation, flipping, and contrast adjustment are applied to artificially increase the training dataset and help prevent overfitting.
- **Noise Reduction**: Filtering methods are used to suppress background noise, helping the model concentrate on relevant anatomical structures.
- **Correction of Artifacts**: Adjustments are made for common issues such as poor lighting, skewed orientation, and low resolution, which can interfere with accurate interpretation.

These preprocessing steps ensure the model receives high-quality, uniform inputs, thereby improving the reliability of its outputs.

Stage 2: Feature Extraction and Classification

Once the images are preprocessed, the system proceeds with feature extraction and classification. This is achieved using a two-model pipeline: Modified VGG16 for deep feature extraction and Fine-tuned MobileNetV2 for classification.

Modified VGG16 for Feature Extraction

VGG16, a deep convolutional neural network known for its uniform architecture, is utilized in this system after specific modifications:

- The number of fully connected (dense) layers is reduced to minimize model complexity and prevent overfitting.
- Batch normalization layers are introduced to enhance training stability and convergence.
- Dropout layers are added to improve generalization by reducing dependency on specific neurons during training.

VGG16 is responsible for identifying high-level visual features from the X-ray images, such as joint space narrowing, cartilage loss, bone spurs (osteophytes), and changes in bone structure. These features are essential in assessing the severity of OA.

MobileNetV2 for Efficient Classification

Following feature extraction, MobileNetV2 is used for classifying the images into OA severity grades. MobileNetV2 is particularly suitable for this task due to its:

- Depth wise separable convolutions, which reduce the number of parameters and computational cost.
- Compact architecture, enabling use in real-time and on devices with limited hardware resources.
- Flexibility for fine-tuning, making it adaptable to the specific patterns in medical imaging data.

The classification model outputs a prediction corresponding to OA severity grades, typically from Grade 0 (normal) to Grade 4 (severe), based on the patterns detected in the knee joint area.

Advantages and Impact of the Proposed System

The integration of these two models—VGG16 for extracting complex features and MobileNetV2 for efficient classification—provides a balanced system that is both powerful and lightweight. The system offers several advantages:

- **Accuracy**: Leveraging deep convolutional layers and transfer learning ensures precise grading of OA.
- **Speed**: The use of MobileNetV2 enables fast inference, which is critical in clinical environments.
- **Scalability**: The system can be adapted to handle X-rays from different sources and varying quality levels.
- **Deployability**: Due to its low computational requirements, the system can be implemented in rural or low-resource clinics.

In summary, the proposed system demonstrates a practical and effective approach to automating the detection and grading of knee osteoarthritis. By combining the strength of deep feature extraction through VGG16 with the computational efficiency of MobileNetV2, the system offers a solution that is both accurate and deployable in real-world clinical settings. It addresses key challenges in medical imaging, including limited dataset sizes, variability in image quality, and computational constraints. With further refinement, this framework has the potential to support healthcare professionals in making faster, more consistent diagnoses and improving patient outcomes through early detection and timely intervention.

2.2.1 Key Advantages of the Proposed System

The proposed deep learning-based framework for knee osteoarthritis (KOA) detection and classification introduces several notable benefits. It combines technical innovation with practical considerations to support clinicians in making timely and accurate decisions. Below is a detailed overview of its advantages:

- **Superior Prediction Accuracy:** The system demonstrates high diagnostic accuracy through the integration of powerful deep learning models—**VGG16** for feature extraction and **MobileNetV2** for classification. MobileNetV2, known for its computational efficiency, achieved a test accuracy of **96%**, indicating its strong capability to accurately categorize X-ray images into five distinct osteoarthritis grades. This level of precision is essential for reliable clinical assessment and grading of disease severity.
- **Automated Region of Interest (ROI) Detection:** An innovative technique is implemented to automatically extract the most relevant region in the X-ray—specifically, the cartilage zone—using pixel density analysis. This removes the need for manual cropping or selection, which not only saves time but also ensures consistent focus on diagnostically important features. By streamlining the preprocessing phase, this technique enhances workflow efficiency and model consistency.
- **Optimal Use of Expert-Labeled Data:** The system is trained on a curated dataset of 1,650 manually annotated knee X-ray images, which ensures high data quality and domain relevance. This strategic use of real-world clinical data supports robust model training and evaluation, reducing dependency on artificially augmented or synthetically generated samples that may not accurately reflect real patient conditions.
- **Comprehensive Evaluation and Validation:** To thoroughly assess model performance, the system utilizes a variety of evaluation metrics, including:
  - **Accuracy**: Overall prediction correctness.
  - **Precision and Recall**: Indicators of true positive and false negative rates.
  - **F1-Score**: Balancing precision and recall for more nuanced evaluation.
  - **Confusion Matrix**: Detailed visualization of predicted vs. actual classifications.

These metrics provide healthcare professionals with deeper insights into the system's diagnostic reliability.

- **Intuitive Web-Based User Interface:** Developed using the Flask framework for the backend and HTML, CSS, and JavaScript for the frontend, the system offers a clean and interactive web interface. Designed with usability in mind, it allows clinicians to upload X-ray images, receive predictions, and view diagnostic insights with minimal effort. The web-based deployment ensures ease of access across devices and operating systems without the need for local installations.
- **High Scalability and Deployment Flexibility:** The lightweight architecture of MobileNetV2 ensures the system can be deployed on a wide range of devices, including resource-constrained environments, without compromising performance. Improved Workflow Efficiency: The automation of KOA detection and classification reduces the burden radiologists and clinicians, saving time and enabling them to focus op patient care.
- **Improved Clinical Workflow:** By automating both detection and grading of KOA, the system reduces the burden on radiologists and orthopedic specialists. This automation helps minimize diagnostic delays, enabling faster decision-making and allowing healthcare providers to dedicate more time to patient interaction and care planning.
- **Visual Interpretation of Results:** The system presents its outputs with visually enriched content, such as graphical plots and bar charts representing model confidence and metric performance. These visualizations enhance the interpretability of predictions, aiding clinicians in understanding and communicating diagnostic outcomes more effectively.
- **Real-World Clinical Applicability:** The robust and adaptable nature of the system positions it well for practical deployment in various environments—ranging from urban hospitals and diagnostic labs to remote health centers and research institutions. Its ability to deliver consistent and actionable insights makes it a valuable tool for early diagnosis, treatment planning, and longitudinal monitoring of OA.

3. BACKGROUND STUDY

This chapter delves into the foundational theories surrounding neural networks and provides background on deep learning [14], particularly in the domains of image classification and object detection. Understanding these principles is essential for implementing accurate and efficient image analysis systems.

3.1 Image Classification

Image classification refers to the process of categorizing input data—specifically images—into predefined labels or classes based on distinguishing features. The objective is to construct a model that learns from labeled training data and generalizes well to unseen data. In supervised learning scenarios, this involves feeding the model images with corresponding labels so that it can recognize patterns and predict the label for new, unclassified inputs.

A well-trained classifier should be capable of correctly identifying the category of an unknown image with a high degree of accuracy. Several algorithmic approaches can be employed for classification tasks, including traditional statistical models and modern machine learning methods. Among these, **neural networks** have proven particularly effective due to their ability to capture complex, non-linear relationships within the data.

3.2 Machine Learning Methods

Traditional programming relies on explicitly defined rules written by developers. In such systems, human-engineered algorithms process input data and generate corresponding outputs based on hard-coded logic. This concept is illustrated in the classical programming pipeline, where the rules governing behavior must be manually crafted.

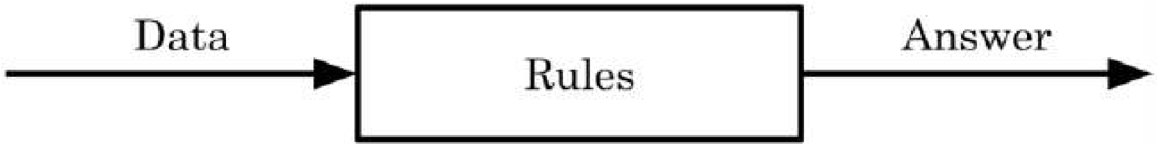

Figure 2. Classical Programming Pipeline

In the machine learning pipeline, instead of writing rules manually, algorithms learn to derive those rules automatically from existing data (shown in Figure 3). The process involves:

It can be used to calculate these rules automatically, so they do not have to be specified by hand. Three components are needed for such an approach:

- **Input Data**: Raw data to be analysed (e.g., images).
- **Target Output**: The correct result the model should produce (e.g., a label or classification).
- **Evaluation Metric**: A method to assess the accuracy or performance of the model (e.g., accuracy, precision, loss).

It works by feeding input and output data into a pipeline, which will learn to transform one into the other. With the advantage that no explicit programming is needed to generate the rules, comes the disadvantage that prior input and output data is required for the initial learning process. Machine learning may be applied as an effective method if it is not feasible or possible to define an algorithm by hand and sufficient data is available for training. How much "sufficient" is depends on factors like the type of task, the complexity of the data, the uniformity of the data, the type of machine learning algorithm and others.

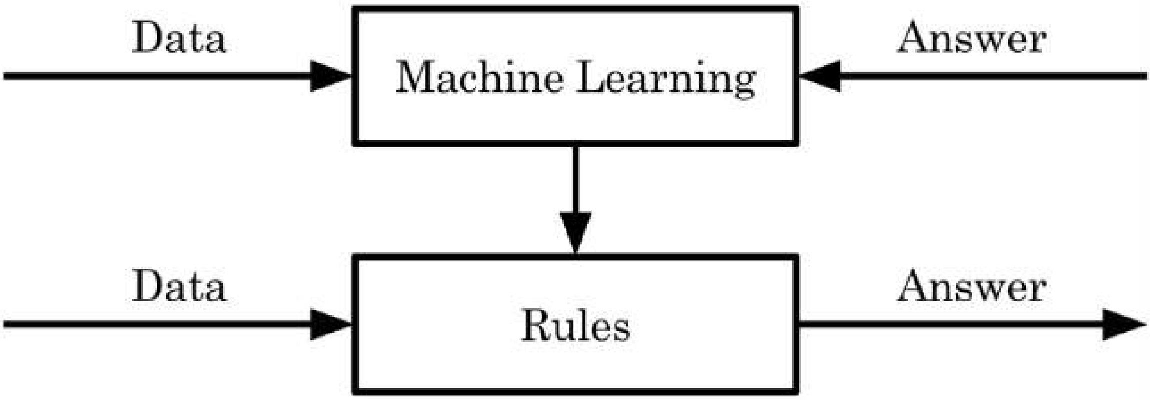

Figure 3. Machine Learning Pipeline

There are different subparts to machine learning like supervised and unsupervised learning. Supervised learning is used when it is clear what the output data looks like, whereas unsupervised learning can help to find unknown patterns in the data. Examples of supervised learning techniques include linear regression, naive Bayes, support vector machines, decision trees, random forests, gradient boosting and artificial neural networks (ANNs).

Supervised learning is typically preferred for classification tasks since the presence of labeled data helps the model learn specific decision boundaries. In contrast, unsupervised methods are better suited for exploratory data analysis, dimensionality reduction, and anomaly detection.

3.2.1 Artificial Neural Networks

Machine learning is a field within computer science focused on creating systems that learn from data, aiming to replicate human cognitive functions. A key method in this domain is the artificial neural network (ANN), modeled after the human brain's structure and function, as illustrated in Figure 4.

ANNs are composed of interconnected units called neurons, which act as the core computational elements. Each neuron receives inputs, processes them based on adjustable weights, and generates an output signal. These weights determine how strongly each input influences the neuron's activation. A standard ANN includes an input layer, one or more hidden layers, and an output layer. Neurons in one layer are linked to the next through weighted connections, forming a layered network. Learning in ANNs involves modifying these weights to minimize the gap between predicted and actual outputs.

Training is the process through which a neural network learns from a dataset of input-output pairs. By iterating over the data, the model adjusts its internal weights to capture meaningful patterns.

The most common training method is backpropagation, which updates weights by calculating the error between predicted and true outputs and distributing it backward through the network. Each weight is adjusted in proportion to its contribution to the error. Over multiple training cycles (epochs), the model improves in accuracy.

Backpropagation allows ANNs to solve complex tasks like image recognition, speech processing, and natural language understanding. With proper training and regularization, these models can generalize effectively to new, unseen data.

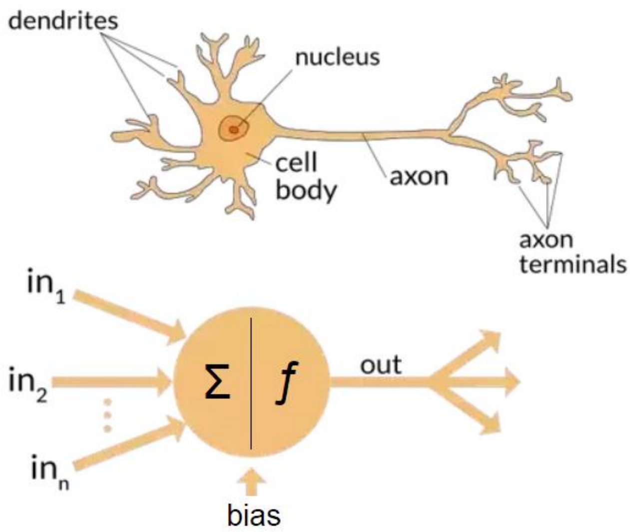

Figure 4. Comparison between a biological and an artificial neuron

3.2.2 Convolutional Neural Networks

Convolutional Neural Networks (CNNs) have emerged as a foundational component in modern deep learning, especially for tasks involving image and video analysis. They are also widely used in fields such as natural language processing, medical imaging, and recommender systems. The key strength of CNNs lies in their ability to capture local spatial patterns in data, which makes them particularly effective for visual recognition tasks.

Unlike traditional fully connected neural networks that treat all input features equally, CNNs utilize convolutional layers to learn localized features by applying filters (also known as kernels) across input data. These filters are trained to detect patterns such as edges, textures, or shapes in images, depending on the depth of the network. The output of the convolution operation is referred to as a feature map, which highlights the presence of specific features in different regions of the input.

The concept of convolution itself originates from mathematics, where it is defined as an operation involving two functions. In the context of image processing, one function represents the input image and the other represents the kernel. The convolution slides the kernel over the image, producing a new matrix where each value corresponds to the result of the operation at a specific location.

CNNs were first introduced by Yann LeCun in the early 1990s, particularly with the development of the LeNet-5 model for handwritten digit recognition. While their initial impact was limited due to computational constraints, the resurgence of deep learning and advances in hardware (especially GPUs) led to their widespread adoption and continued evolution.

Key Components of CNNs

A typical CNN architecture consists of several types of layers arranged in a hierarchical structure:

1. **Convolutional Layers**: These layers apply filters to extract relevant features from the input image. Each filter focuses on different patterns, such as horizontal or vertical edges.
2. **Activation Functions**: Non-linear functions like ReLU (Rectified Linear Unit) are applied to the feature maps to introduce non-linearity into the model, allowing it to learn more complex patterns.
3. **Pooling Layers**: These layers reduce the spatial dimensions of feature maps, usually through operations such as max pooling or average pooling. Pooling helps in decreasing computational requirements and prevents overfitting by retaining only the most significant features.
4. **Fully Connected Layers**: After several layers of convolutions and pooling, the final output is flattened and passed through one or more dense layers for classification or regression.

By stacking multiple convolutional and pooling layers, CNNs progressively learn **high-level representations** of input data—from simple patterns in early layers to more abstract features in deeper layers.

CNN Workflow

Building a CNN typically follows four core steps:

1. **Preprocessing**: Input images are normalized and resized to ensure consistency.
2. **Feature Extraction**: Convolution and pooling layers extract hierarchical features from the images.
3. **Classification**: Fully connected layers use the extracted features to classify the input into defined categories.
4. **Model Evaluation**: Performance is measured using metrics like accuracy, precision, and recall, and the model is refined accordingly.

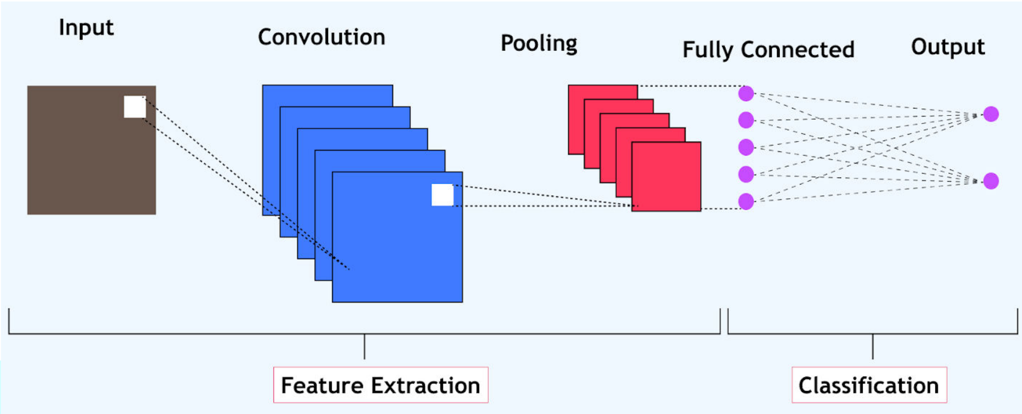

Figure 5. A General CNN layer hierarchy

Step - 1: Convolution
Step - 2: Pooling
Step - 3: Flattening
Step - 4: Full connection

The basic foundation of every Convolutional Neural Network is made up of these operations, so to develop a sound understanding of the working of these ConvNets, we need to comprehend thoroughly the working of these operations.

3.3 Object Detection

While traditional image classification models are designed to determine the presence or absence of specific objects within an image by assigning probability scores, they do not provide information about the location of these objects. This is where object localization comes into play.

Object localization refers to the process of identifying *where* an object appears within an image. It involves not just recognizing the object but also determining its position by outputting coordinates that define its location. In most computer vision applications, these positions are visually represented using bounding boxes—rectangular outlines that enclose the detected object within the image frame.

These bounding boxes are typically defined by the coordinates of the top-left and bottom-right corners (or center coordinates, width, and height) relative to the image dimensions. This approach enables algorithms to both recognize and spatially locate objects, which is essential in tasks like autonomous driving, video surveillance, and medical imaging.

Figure 6 illustrates an example of a bounding box being used to highlight an object within an image, effectively marking its location.

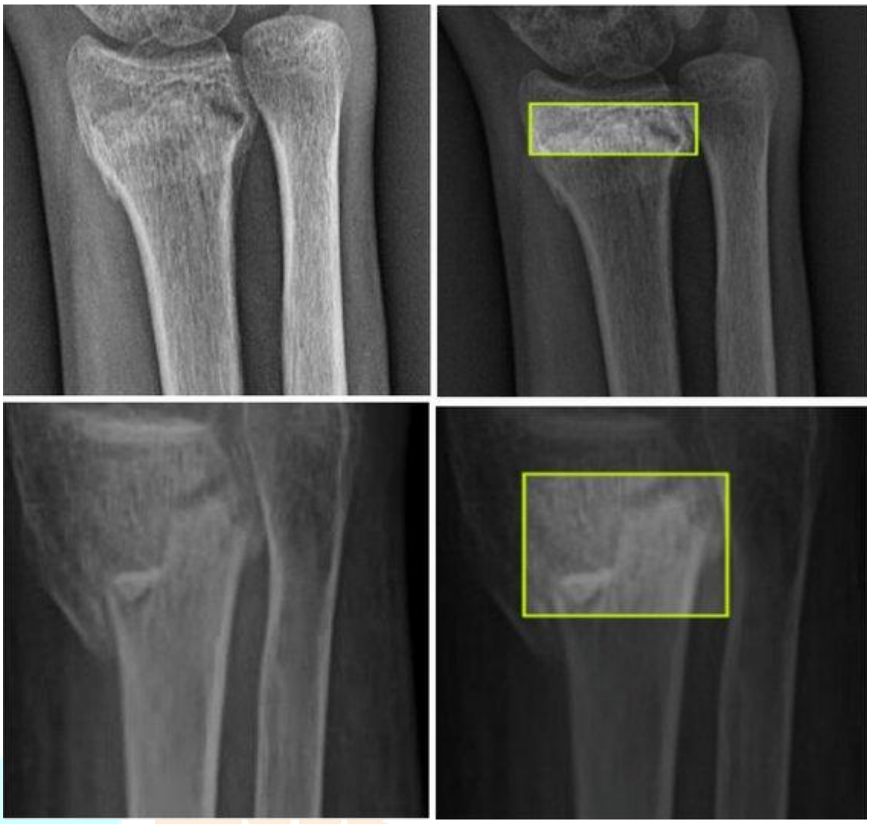

Figure 6. Sample of bounding box representation

3.3.1 Modern Methods in Object Detection

Recent advancements in object detection are largely driven by deep learning, particularly through the use of convolutional neural networks (CNNs) and transformer-based models. Modern detection frameworks generally fall into two main categories: two-stage detectors and one-stage detectors.

Two-stage detectors—such as Faster R-CNN—operate by first generating a set of region proposals that likely contain objects. These proposals are then passed through a second stage where classification and bounding box refinement are performed. Although these models are known for their high detection accuracy, they tend to have slower inference times, making them more suitable for applications where speed is not the primary constraint.

On the other hand, one-stage detectors, including YOLO (You Only Look Once) and SSD (Single Shot MultiBox Detector), take a more direct approach. These models predict object classes and bounding boxes in a single pass through the network, allowing for real-time detection. However, this speed often comes with a slight compromise in accuracy compared to two-stage methods.

In addition to these paradigms, transformer-based models have introduced new possibilities in object detection. A notable example is DETR (DEtection TRansformer), which utilizes self-attention mechanisms to detect objects without relying on traditional region proposal methods. This approach simplifies the detection pipeline and enables global context modeling, leading to competitive performance on benchmark datasets.

To further enhance detection, especially for small or overlapping objects, various supporting techniques have been introduced:

- Feature Pyramid Networks (FPNs) help models detect objects at multiple scales.
- Anchor-free methods eliminate predefined anchor boxes, reducing computational overhead and improving localization flexibility.
- Multi-scale training augments model robustness across different object sizes.

Moreover, modern detection systems increasingly incorporate self-supervised learning to leverage unlabeled data, knowledge distillation to transfer performance from large models to lightweight ones, and domain adaptation to maintain accuracy across different image environments or acquisition conditions.

Together, these innovations make current object detection systems more scalable, efficient, and applicable to real-world scenarios, including autonomous driving, robotics, medical diagnostics, and surveillance.

3.3.2 Speed-Accuracy Tradeoff

In object detection, one of the central challenges lies in balancing detection accuracy with inference speed a compromise commonly referred to as the speed-accuracy trade-off. This trade-off reflects the difficulty of achieving both real-time performance and high-precision results within a single model, especially in complex or resource-constrained environments.

Detection models with a focus on accuracy, such as Faster R-CNN and Mask R-CNN, typically employ multi-stage pipelines, deeper networks, and extensive computation. These models are capable of detecting small, overlapping, or low-contrast objects with high precision, making them well-suited for applications where detection quality is critical. However, this comes at the cost of higher inference times, limiting their practicality in time-sensitive scenarios.

Conversely, real-time detection models like YOLO (You Only Look Once) and SSD (Single Shot MultiBox Detector) prioritize speed by adopting simplified, one-stage architectures and fewer computational layers. While they offer impressive performance in terms of latency, they may experience a drop in accuracy, particularly when dealing with densely populated scenes, occlusions, or fine-grained object boundaries.

To optimize this tradeoff, researchers employ several strategies, including:

- Model pruning, which removes redundant parameters to reduce computation.
- Quantization, which compresses model weights by reducing precision (e.g., from 32-bit to 8-bit).
- Knowledge distillation, where a smaller "student" model learns from a larger "teacher" model to retain performance while minimizing complexity.
- Hardware acceleration, leveraging devices like GPUs, TPUs, or NPUs to speed up processing without changing the model architecture.

In addition, lightweight architectures such as MobileNet, EfficientDet, and Tiny-YOLO have been developed specifically for environments with limited processing capabilities, such as smartphones, drones, or embedded systems.

Ultimately, the optimal model choice is guided by the specific requirements of the application. For example, autonomous vehicles demand rapid, real-time detection to ensure safety and responsiveness, even if some precision is compromised. In contrast, medical imaging applications often tolerate slower inference in favor of achieving the highest possible accuracy, especially when diagnosing critical conditions.

3.3.3 Object detection in medical images

Computer vision techniques have become increasingly integral to modern healthcare, particularly in the field of medical imaging. Medical object detection refers to the process of identifying and localizing disease-related structures or regions of interest within medical images. This capability is essential in various diagnostic applications, including the detection and classification of diseases such as osteoarthritis, which is the focus of this study.

With the rise of advanced medical imaging technologies, tools for medical image processing are playing a pivotal role in supporting clinicians across tasks such as diagnosis, treatment planning, and image-guided surgical interventions. One critical application is the accurate and efficient tracking of anatomical structures—especially those that are dynamic or deformable, like the heart. A major challenge in this domain is precisely locating the region of interest within noisy or low-resolution images.

Today, medical object detection is being successfully applied across a range of imaging modalities, including:

- **Ultrasound (US):** for identifying tissues or fetal anatomy,
- **Computed Tomography (CT):** particularly for analyzing cardiovascular or skeletal structures,
- **Magnetic Resonance Imaging (MRI):** widely used for brain and musculoskeletal analysis,
- **X-ray fluoroscopy:** for continuous imaging during interventional procedures.

By leveraging object detection models, healthcare systems can enhance diagnostic accuracy, reduce manual effort, and ultimately improve patient care through early and reliable disease identification.

3.4 Object Detection Approach

Object Detection with VGG16

VGG16 is a widely used deep convolutional neural network architecture known for its simplicity and strong performance in feature extraction. It plays a crucial role in many object detection systems, often serving as the feature extraction backbone in models like Faster R-CNN. This architecture consists of 16 layers, primarily convolutional layers that use small 3×3 filters. These filters are effective at capturing detailed and hierarchical patterns within an image, enabling the model to learn rich feature representations. Such features are critical in object detection tasks, where accurate localization and classification of objects depend on deep spatial understanding. Despite its strengths in extracting meaningful features, VGG16 is computationally intensive, with a large number of parameters and high memory consumption. This makes it less suitable for applications requiring real-time inference or deployment on devices with limited computational resources. Nonetheless, VGG16 remains a strong choice for tasks where accuracy is prioritized over speed, especially in research and clinical environments where processing time is less critical than diagnostic precision.

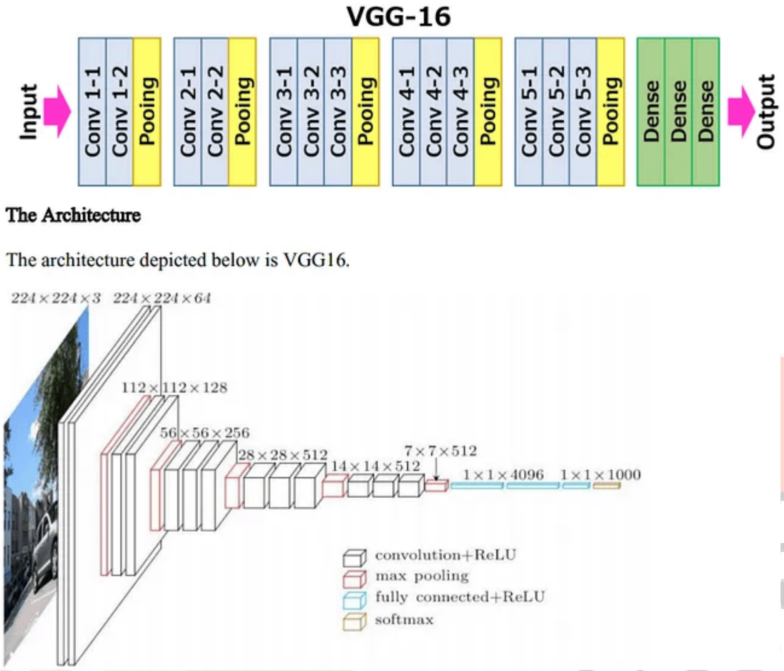

Figure 7. VGG16 Architecture

Object Detection with MobileNetV2

MobileNetV2 is a lightweight convolutional neural network architecture optimized for speed and efficiency, making it ideal for real-time object detection on mobile and edge devices. It achieves this by incorporating depthwise separable convolutions and inverted residual blocks, which significantly reduce the number of parameters and computational requirements compared to traditional CNN models. This streamlined design allows MobileNetV2 to be effectively integrated into single-stage detection frameworks, such as SSD (Single Shot MultiBox Detector), where quick inference is essential. The architecture is particularly well-suited for deployment in environments with limited resources, such as smartphones or embedded systems. While MobileNetV2 offers significant advantages in terms of speed and energy efficiency, it may compromise slightly on accuracy when compared to deeper and heavier models like VGG16. However, for many real-time applications, this trade-off is acceptable, especially when the priority is low latency and minimal hardware demand.

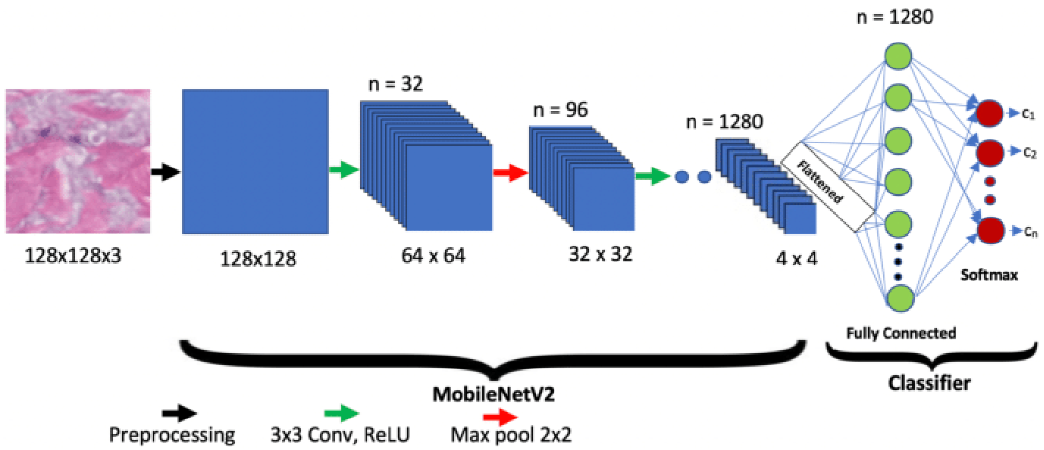

Figure 8. MobileNetV2 Architecture

4. SYSTEM DESIGN

4.1 Design

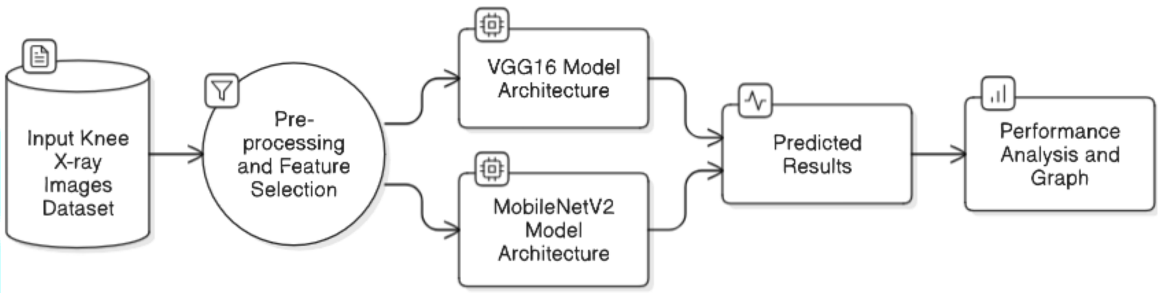

Figure 9. Design of KOA classification

Building a deep learning system for predicting and classifying knee osteoarthritis (OA) from X-ray images involves several critical stages: data acquisition, preprocessing, model training, and deployment. The process begins with obtaining a dataset of annotated knee X-rays, such as those from the Osteoarthritis Initiative (OAI), which includes images labeled using the Kellgren-Lawrence (KL) grading scale. A diverse dataset covering all OA severity levels—from healthy to severely affected joints—is essential for robust model performance.

Data preprocessing plays a crucial role in preparing the images for training. Images are typically resized to a uniform resolution, such as 224x224 pixels, and pixel intensities are normalized, usually to a 0–1 scale. To prevent overfitting and improve generalization, augmentation techniques like rotation, flipping, and zooming are applied. Each image is accurately labeled based on OA severity—either in discrete grades for classification or continuous scores for regression.

Convolutional Neural Networks (CNNs) are the preferred choice for image-based analysis due to their ability to extract spatial features. In this study, transfer learning is leveraged using pre-trained models like ResNet, EfficientNet, and DenseNet, originally trained on ImageNet. These models are fine-tuned for OA detection by modifying their final layers to match the classification or regression task—using softmax for class prediction or linear activation for continuous score estimation.

The training phase involves selecting appropriate loss functions—categorical cross-entropy for multi-class classification or mean squared error (MSE) for regression—and optimizing the model using algorithms such as Adam or SGD. Performance metrics like accuracy, F1-score, and mean absolute error (MAE) are used for evaluation, and cross-validation helps ensure the model generalizes well to unseen data.

Once trained, models are evaluated on a separate test set using relevant metrics. In this project, both VGG16 and MobileNetV2 were fine-tuned to classify knee X-rays into five KL grades: Normal, Doubtful, Mild, Moderate, and Severe. These models are compared based on accuracy, inference time, and computational efficiency.

For deployment, the trained model is integrated into a user-accessible system—commonly via a Flask web application—enabling users to upload X-ray images and receive automated OA severity assessments. The complete pipeline is developed using frameworks like TensorFlow or PyTorch for model training, OpenCV or PIL for image preprocessing, and Flask, FastAPI, TensorFlow Lite, or ONNX for deployment. This modular architecture ensures the system is scalable, efficient, and suitable for real-world clinical use.

4.2 Dataset Collection

The data set is a collection MRIs of Right and Left Knees. It includes data samples that came from multiple sources. The dataset includes 9,786, 399 and 2645 high-quality 8-bit grayscale X-ray images collected from reputable hospitals and diagnostic centers by the University of Florida. It is available for download at the following URLs:

- Knee Osteoarthritis Dataset with Severity Grading, "https://www.kaggle.com/datasets/shashwatwork/knee-osteoarthritis-dataset-with-severity/data"
- cgmh-oa "https://www.kaggle.com/datasets/tommyngx/cgmh-oa"
- Knee Xray "https://www.kaggle.com/datasets/jayaprakashpondy/knee-xray"

5. RESULTS AND CONCLUSION

5.1 Results

Figure 10. Welcome Page

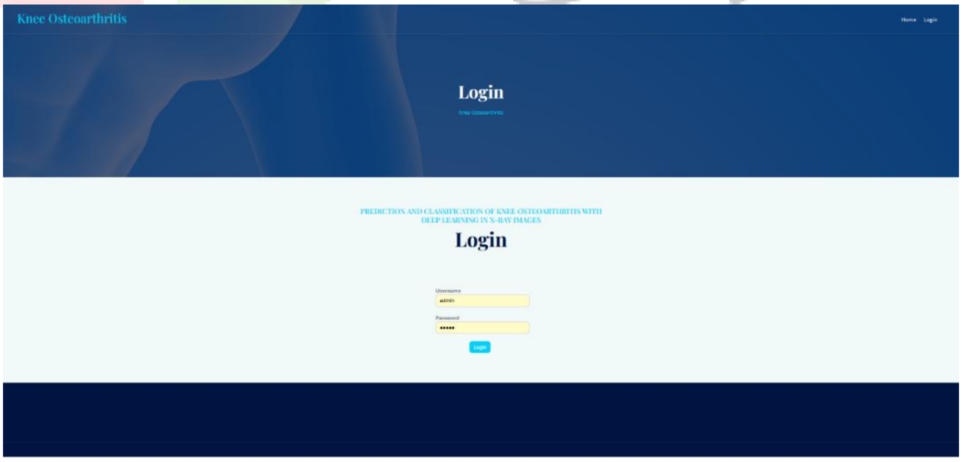

Figure 11. Login Page

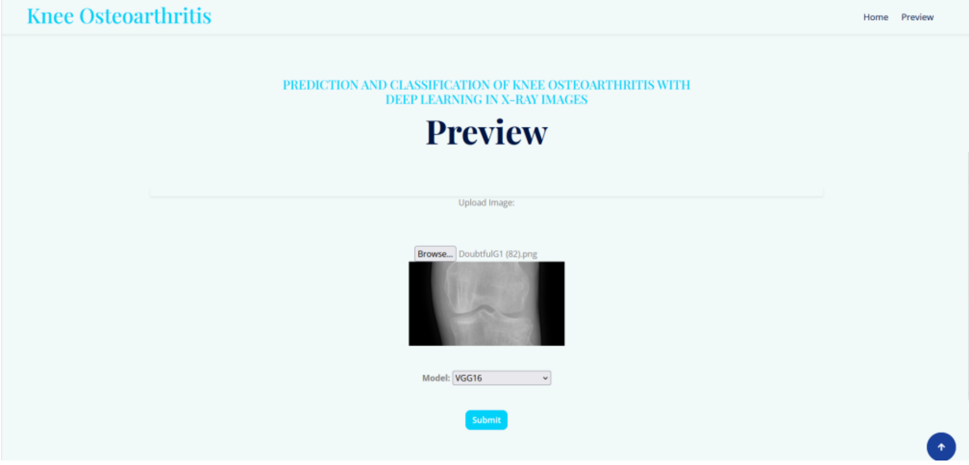

Figure 12. Image Uploading Page

Figure 13. Image Results Page

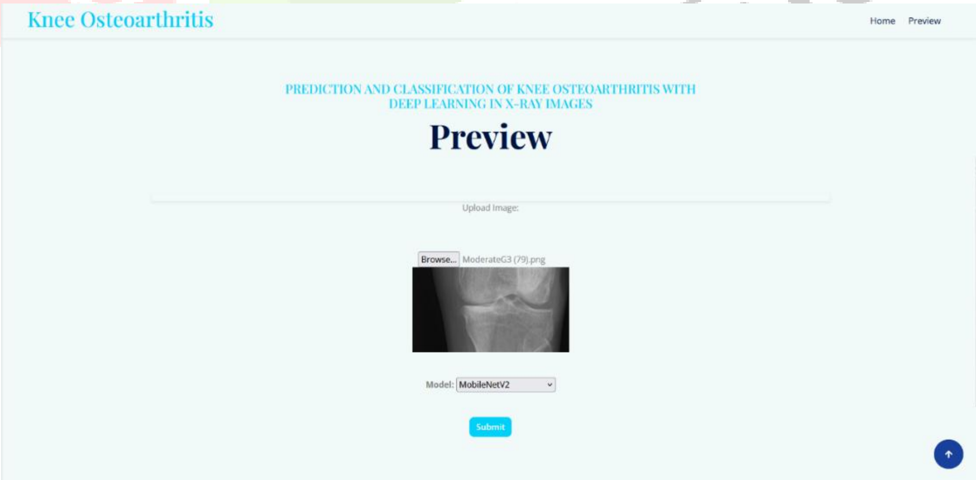

Figure 14. Image Uploading Page

Figure 15. Image Results Page

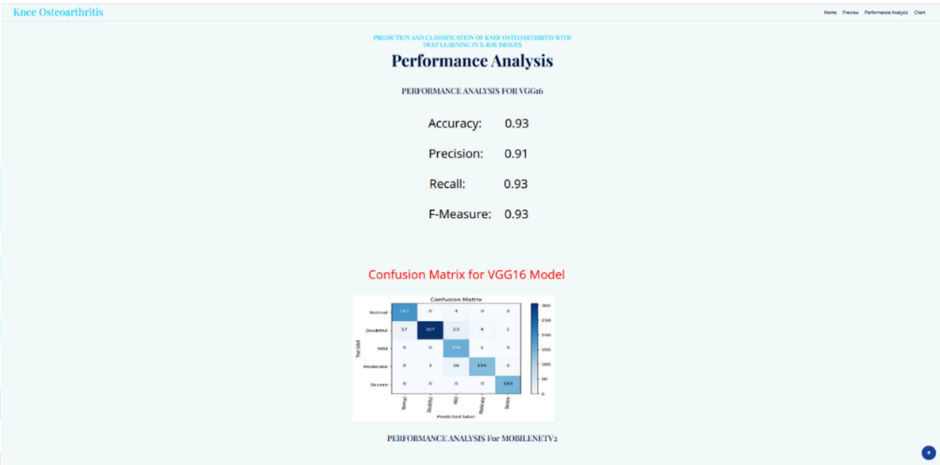

Figure 16. Performance Analysis – VGG16

Figure 17. Performance Analysis – MobileNetV2

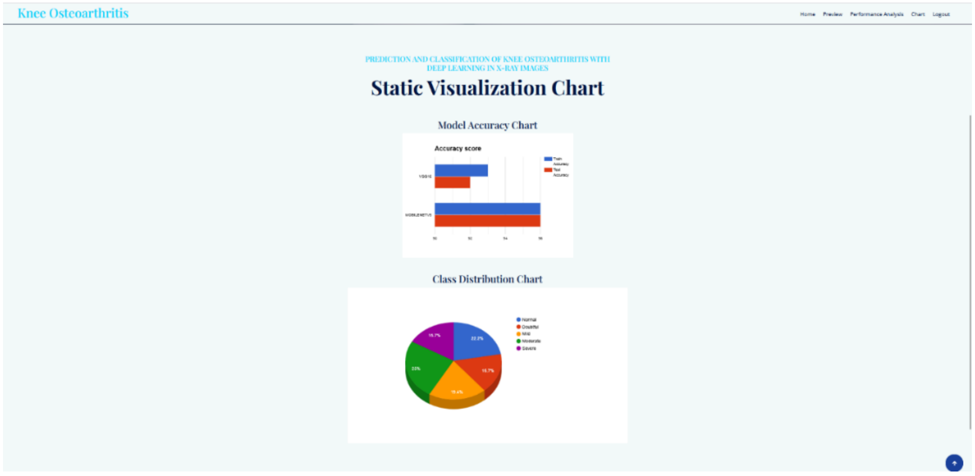

Figure 18. Accuracy and Distribution Chart

5.2 Conclusion

The project titled " Predicting and Classifying Knee Osteoarthritis with Deep Learning on X-ray Images" effectively showcases how deep learning can streamline and enhance the diagnostic process for knee osteoarthritis (OA). By utilizing cutting-edge convolutional neural network architectures such as VGG16 and MobileNetV2, the system demonstrates impressive accuracy levels, with MobileNetV2 achieving a test accuracy of 96%.

To enhance the model's precision, advanced preprocessing techniques are employed, including innovative cartilage region extraction, which helps the model concentrate on the most diagnostically relevant areas of the knee X-ray. A carefully selected and expert-annotated dataset reinforces the reliability of the model's predictions. In addition, thorough performance evaluations and an intuitive web-based interface contribute to the system's readiness for use in real-world clinical environments.

This work bridges the gap between modern artificial intelligence technologies and pressing healthcare needs, delivering a scalable, accurate, and efficient solution for diagnosing and grading knee osteoarthritis. The results underline the transformative role of deep learning in medical imaging, especially in improving diagnostic workflows and supporting clinical decision-making.

Future Enhancements:

- **Incorporating Multimodal Data**: Future versions of the system could integrate complementary data types, such as demographic details, patient history, and genetic markers, to enhance diagnostic accuracy and deliver more holistic assessments.
- **Real-Time Inference Capability**: Adding support for live, real-time X-ray analysis could improve clinical usability by enabling immediate diagnostic insights, thus minimizing wait times for patients.
- **Mobile and Edge Deployment**: Optimizing the system for mobile platforms and edge devices would expand access in rural or resource-limited environments, making the tool more inclusive.
- **Support for Additional Imaging Modalities**: Extending compatibility to include other imaging technologies like MRI or CT scans could broaden the diagnostic scope, particularly for complex or advanced OA cases.
- **Expanding and Diversifying the Dataset**: Incorporating a wider range of X-ray images from different ethnic groups, age brackets, and health backgrounds would enhance the model's robustness and generalizability.
- **EMR System Integration**: Designing the system to interface with Electronic Medical Record (EMR) systems would enable seamless adoption into clinical workflows, improving both efficiency and patient outcomes.
- **Longitudinal Tracking**: Embedding functionality to monitor OA progression over time using serial imaging would assist clinicians in evaluating treatment effectiveness and adjusting care plans accordingly.

REFERENCES

[1] Pirkko Mustamo. Object detection in sports: Tensorow object detection api case study. Bachelor's Thesis Degree Programme in Mathematical Sciences, January 2018.

[2] Zhifeng Chen Quoc V. Le Mohammad Norouzi Wolfgang Macherey Maxim Krikun Yuan Cao Qin Gao Klaus Macherey Je Klingner Apurva Shah Melvin Johnson Xiaobing Liu Lukasz Kaiser Stephan Gouws Yoshikiyo Kato Taku Kudo Hideto Kazawa Keith Stevens George Kurian Nishant Patil Wei Wang Cli Young Jason Smith Jason Riesa Alex Rudnick Oriol Vinyals Greg Corrado Macdu Hughes Yonghui Wu, Mike Schuster and Jerey Dean. Google's neural machine translation system: Bridging the gap between human and machine translation. Sep 2016.

[3] Anna Choromanska Krzysztof Choro-manski Bernhard Firner Lawrence Jackel Mariusz Bojarski, Philip Yeres and Urs Muller. Explaining how a deep neural network trained with end-to-end learning steers a car. Apr 2017.

[4] Joshua Saxe and Konstantin Berlin. Deep neural network based malware detection using two dimensional binary program features. Aug 2015.

[5] Justin Francis. How to create your own custom object detection model. O'reilly link- https://www.oreilly.com/, October 2017.

[6] Kurt. Object detection tutorial in tensorow: Real-time object detection. edureka, 7, April 2019.

[7] Pooja U. Patravali Ramesh R. Manza, Shivanand S. Gornale. A survey on exploration and classification of osteoarthritis using image processing techniques. International Journal of Scientific Engineering Research, 7, June 2016.

[8] Introduction and use - tensorow object detection api tutorial, pythonprogramming.net.

[9] A. Cui, H. Li, D. Wang, J. Zhong, Y. Chen, and H. Lu. Global, regional prevalence, incidence and risk factors of knee osteoarthritis in population-based studies, April 2020.

[10] P. P. F. M. Kuijer, H. F. van der Molen, and S. Visser. A health-impact assessment of an ergonomic measure to reduce the risk of work-related lower back pain, lumbosacral radicular syndrome and knee osteoarthritis among floor layers in The Netherlands, March 2023.

[11] V. K. V, V. Kalpana, and G. H. Kumar. Evaluating the efficacy of deep learning models for knee osteoarthritis prediction based on Kellgren lawrence grading system, September 2023.

[12] S. Castagno, M. Birch, M. van der Schaar, and A. McCaskie. Prediction of the Rapid Progression of Knee Osteoarthritis Using Automated Machine Learning: A Novel Precision Health Approach for Chronic Degenerative Disease, January 2024.

[13] A. A. S. Afroze, R. Tamilselvi, and M. G. P. Beham. Machine learning based osteoarthritis detection methods in different imaging modalities: A review, March 2023.