SOLENT UNIVERSITY, SOUTHAMPTON

FACULTY OF BUSINESS, LAW AND DIGITAL TECHNOLOGY

MSC APPLIED AI AND DATA SCIENCE (MAIDS)
2022

"Intelligent Identification of Knee Osteoarthritis Severity Using Deep Learning"

STUDENT: VIRALBEN RAVAL(Q15826589)
COURSE: DESSERTATION(COM726)
COURSE LEADER: DR FEMI
SUPERVISIOR: MUNTASIR AI-AFSOOR

Acknowledgments

I want to thank Prof. Muntasir Al-Afsoor, who is my supervisor. And for their ongoing assistance and leadership over the course of our endeavour. I also want to thank Solent University for contributing to and participating in the study.

Abstract

The prevention and diagnosis of chronic diseases have been the top priority of the biomedical sector, & with the advancements in technology, computer-aided diagnostics (CADs) have aided a lot in this regard. Still, with the advancements in CADs, accurate diagnosis with the least amount of human error is always a challenging task for doctors, surgeons & medical staff for long period. With the recent advancements in computer science, and especially in the areas of artificial intelligence (AI) & Machine Learning (ML), the biomedical sector has truly aided and powered effective disease diagnosis. Knee Osteoarthritis is a chronic disease that happens in adults and it needs to be diagnosed at early stages to prevent harm to patients and their lives. With the advancements in AI, this disease can be diagnosed with the least amount of error and will aid in controlling the factor of human error that happens in manual diagnosis. This research aims in analysing, evaluating & developing an intelligent model using artificial intelligence that will be capable of identifying the severity of knee osteoarthritis in 5 possible classes using a dataset that contains X-ray scanned images of patients. The classes indicate the severity level of osteoarthritis from normal to severe. Deep Learning models like convolutional neural networks (CNNs) will be employed for the task since the dataset that is to be used is image related. The dataset for knee osteoarthritis will be obtained from Kaggle, which is an online platform for AI & machine learning researchers. For achieving higher accuracy & to put our research work competitive, pre-trained CNNs will be employed for image classification via transfer learning (TL). This research is initiated with a target not only to draw useful insights but also to highlight challenges for future practitioners & researchers working in this field.

Keywords: Machine Learning (ML), Artificial Intelligence (AI), Deep Learning (DL), Transfer Learning (TL), Convolutional Neural Networks (CNNs), X-ray images and Medical System

Contents

Figures:

Tables

1 – Introduction

1.1 Introduction & Background

Osteoarthritis (OA) of the knee, commonly referred to as degenerative joint disease of the knee, is frequently brought on by wear and strain as well as the gradual loss of articular cartilage. The elderly is most frequently affected. Primary and secondary osteoarthritis of the knee can be distinguished from one another. Articular degeneration without the need for a clear underlying cause is primary osteoarthritis. Secondary osteoarthritis results from either defective articular cartilage, such as in rheumatoid arthritis, or improper force distribution throughout the joint, as in post-traumatic reasons (Hsu and Siwiec, 2018[1]).

One of the most common chronic disorders is osteoarthritis (OA) of the knee. Pain, stiffness, and a reduction in participation in everyday activities are the results (Lawson et al., 2015[2]). Joint deformity affects muscle strength and bony alignment, which impairs proprioceptive sensory function and impairs the ability to recognize balance (Hatfield et al., 2016[3]). Between the ocular, vestibular, and brain systems, there are multiple neuromuscular interactions that contribute to balance. Therefore, any change in these systems has the potential to upset the balance (Liu et al., 2017[4]). KOA is strongly linked to a significant financial load on the healthcare system as well as an intolerable health burden on patients and their families (Toivanen et al., 2010[5]) (Gupta et al., 2005[6]).

Biomedical Development in the 21st Century: Technology-assisted chronic disease management has become an important area of biomedical study in the twenty-first century. Utilizing cutting-edge, modern methods for disease prevention is essential for healthcare and governmental agencies in developed countries. Such advanced technologies not only detect or identify chronic diseases but also strive to prevent them so that fewer people are affected by them. Due to developments in the medical profession, computer-aided diagnostics (CADs) in particular have shown to be quite beneficial in this area.

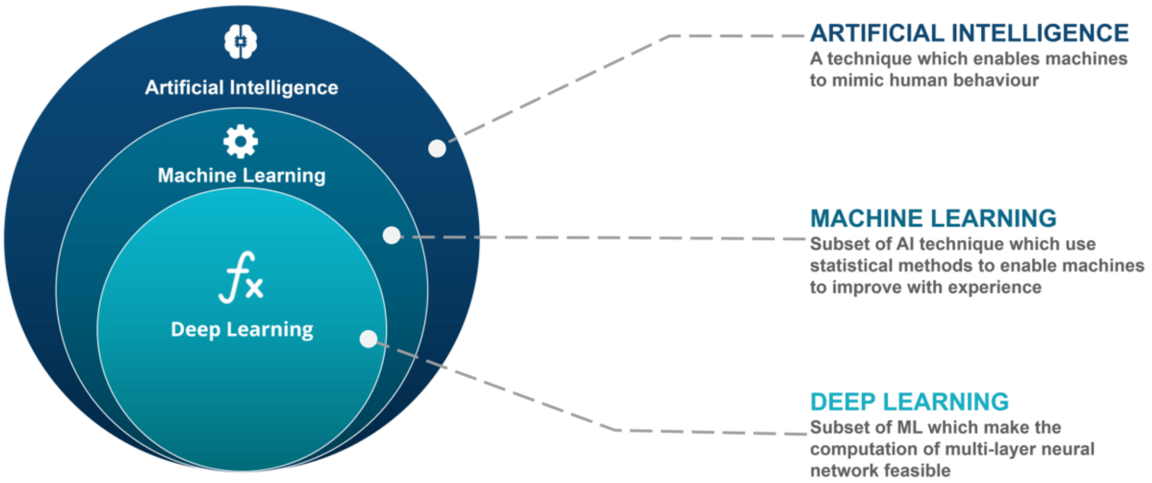


Figure 1 Comparison of Deep Learning

Artificial Intelligence has had a wide range of effects on the health care industry. The way people obtain healthcare is drastically changing, from online medical exams, phone lines, and automation surgeries. Machine learning has been increasingly important in modern healthcare image processing applications. Specifically, the deep learning area of machine learning that includes more sophisticated and effective techniques for conditioning a neural network for disease diagnosis. Models with numerous layers are used in deep learning to learn. The various methods used to represent data These methods have significantly improved the cutting-edge for many computing applications.

Knee Osteoarthritis: The fourth most prevalent cause of disability worldwide is knee osteoarthritis (OA), a widespread condition. Up to 14 million Americans are estimated to suffer knee symptoms, with tens of millions more impacted in Europe, South America, Asia, and the Middle East, according to the US National Health Interview Survey. The economic burden associated with OA is projected to equal up to 2.5% of Growth of National Product in Western countries as a result of ensuing healthcare costs and activity losses. Most often, older people are impacted. Osteoarthritis of the knee can be divided into primary and secondary forms. Primary osteoarthritis is articular deterioration without a recognized underlying cause.

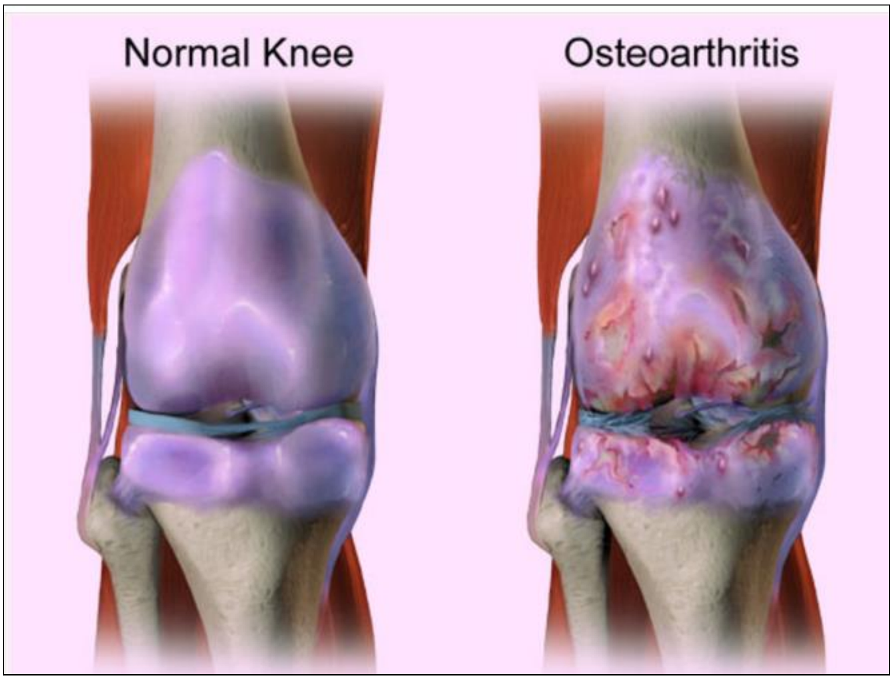


Figure 2 Comparison of Normal knee and osteoarthritis

Computer-Aided Diagnostics: Diagnosis medical images with the aid of computer-aided diagnostics (CADs), which are gadgets. Doctors can quickly identify a chronic condition with the aid of imaging tools including X-rays, CT scans, ultrasounds, and MRIs. To deliver data to aid a professional's decision-making, CAD systems examine digital pictures for common features and highlight salient areas, such as potential diseases. (2017) Digital Pathology Association.

CAD – Being an Interdisciplinary Technology: Computer-aided diagnostics, referred to as CAD, is an interdisciplinary technology combining the fields of artificial intelligence (AI) and computer vision (CV) for effective diagnosis of diseases. Typical applications of CAD include the detection of a brain tumour from MRI images. Hospitals make use of CAD techniques for the detection of breast cancer, brain tumour, COVID-19 from X-ray images, lung cancer detection, identification of Knee Osteoarthritis, etc. The development of computer-aided diagnostic systems has been boosted by the applications of artificial intelligence & computer vision. For successful identification of these chronic diseases, these two fields are combined to make the outcome. The diseases could also be diagnosed manually by medical staff but due to the intervention of AI and ML, thousands of such images could be diagnosed with the least amount of error, with human intervention**.**

Computer Vision (CV) is also an area of artificial intelligence directed toward the understanding of images by machines. The field of computer vision is headed toward how computers see and recognize images. Giving computers the capability of understanding the visual world is one of the hottest applications of AI. The applications of computer vision are applied in every sector: detecting the images of dogs and cats, intelligent image classification, and especially detecting disease or malfunction from a medical image. Computer vision has introduced cutting-edge performances in the areas of medicine and aiding medical staff to diagnose diseases efficiently and with the least amount of error.

This work seeks to undertake a thorough investigation, inspection, analysis, and insight-drawing procedure to assess the severity of knee osteoarthritis using deep learning techniques. This project aims to investigate the disciplines of deep learning, computer vision, etc. in artificial intelligence to control diseases and save patients' lives. The company also intends to use trained CNNs to build a deep learning model. Moreover, this finding will provide new prospects for young researchers, particularly those working in the disciplines of AI and ML, to support the biomedical sector in the efficient prevention of deadly illnesses.

1.2 Aims:

- ✓ To develop an intelligent model using deep learning for intelligent identification of the severity of Knee Osteoarthritis using deep learning.

1.3 Objectives:

- ✓ To assist medical staff and doctors in the identification of the severity of Knee Osteoarthritis by making a deep learning CNN model
- ✓ To increase the accuracy of an AI system employing transfer learning to detect the severity of Knee Osteoarthritis with minimum loss

1.4 Research Questions

- ✓ How to apply deep learning CNN techniques in the biomedical sector for intelligent classification of knee images obtained via X-ray scans?
- ✓ What is the significance of using deep learning models instead of employing classical machine learning models for intelligent image classification?
- ✓ How, with the use of transfer learning, the accuracy of an image classification system can be improved
- ✓ How an intelligent AI system for the biomedical sector can reduce the labour of medical staff in diagnosing the severity of Knee Osteoarthritis

1.5 Project Overview

The goal of the study is to assess the severity of knee osteoarthritis (KOA) using an image from an X-ray scan. The Kallgren-Lawrence (KL) grading system provides a scale that assesses the severity of KOA. This is a classification issue because the KL grading system classifies KOA severity in 5 classifications. Considering these factors, this project's philosophy is based on computer vision, deep learning, and artificial intelligence (AI) (CV). The X-ray image dataset will be entered into our workspace, transformed into numbers using computer vision methods, and then deep learning (DL) will be applied to those images for intelligent classification.

1.6 Ethical Considerations

Ethics is a crucial subject to discuss, especially when doing economical and error-free research. The UK Data Centre Department offers standards that include ethical issues for doing successful big data research, which also serves as the foundation for this research's ethics.

Chapter 2 Literature Review

2.1 Literature Review

Medical image analysis using machine learning has recently become a promising field. Solution to be used in the diagnosing field. Deep learning as a replacement for traditional machine learning Diagnostic and predicative techniques are being used to apply learning.so here I am going to explain related work regarding this model.

(Kokkotis et al., 2020) provided a systematic review of the development of machine learning in the diagnosis of knee osteoarthritis, in which key directions and future developments of machine learning in diagnosis & prediction of knee osteoarthritis are discussed. In this research, a survey was conducted using the research articles published in the period 2006-19, and the articles were divided into multiple categories based on the approaches used by researchers in those articles, such as regression, classification, planning techniques & image segmentation. After a brief systematic review of selected papers, this research indicates that knee osteoarthritis is a challenging & big data problem, in which data complexity and its size are the key factors. Machine Learning has attracted many researchers to diagnose this disease intelligently, and more developments are expected to be seen in the future.

(Schwartz et al., 2020) provided a research contribution to knee osteoarthritis by addressing the question of whether a convolutional neural network (CNN) can classify knee osteoarthritis using radiology images, competitive with those which are identified by surgeons. For answering this question, various knee radiology images obtained from different patients were gathered for experiments, and these images were graded by knee arthroplasty surgeons using the IKDC grading system. For calculating the skill of a CNN, the inter-correlation coefficient (ICC) of surgeons and the ones classified by CNN was compared with each other. The results of this experiment showed that a CNN can accurately classify knee osteoarthritis as well as compared to professional arthroplasty surgeons and this technique is much more useful to be applied along with computer-aided diagnostics to diagnose knee osteoarthritis.

(Swiecicki et al., 2021) used a deep learning-based approach for assessing the severity of knee osteoarthritis in radiology images. The purpose of this research was to develop an intelligent, deep learning algorithm that could automatically classify the security of knee osteoarthritis using radiology images, keeping by jointly using posterior-anterior & lateral views of radiology photographs. For training a deep learning model, a dataset of over 2500 images were used, in which 2000+ images were used for training & other images were used for validation & testing purposes. For applying a deep learning model, first knee joints in images were localized and then classification was applied to images using KL grading criteria. The research highlights that proper training of a deep learning model on radiology images yielded an accuracy of 71.90%. The trained model showed the desired accuracy using the KL criteria & the results were also shown to radiologists for confirmation. The radiologists agreed with the results of the model proposed in this research.

(Norman et al., 2019) applied a densely connected convolutional neural network (CNN) for identifying the stages of knee osteoarthritis (OA), using KL grading as a standard to identify the severity of knee OA. The neural network used in this research is fully automated, which identifies knee OA stages showing desired results. For applying a deep neural network for classification, knee radiology images were localized using image segmentation techniques. After the localization process, the localized images were then employed as an input for CNN for the training process. In this research, an ensemble of dense neural networks is used to make a model more accurate & effective, which yielded desired and projected accuracy & specificity scores on test datasets. Moreover, it was also found in this research that the features extracted by CNN from radiology images belonged to actual features that are necessary for classification with the help of saliency maps. The results of this research indicate that this model may be used by radiologists to get aided in the effective diagnosis of knee OA.

In a research contribution in the field of knee Osteoarthritis (OA), (Kim et al., 2020) identified if additional patient information may prove to be useful in increasing the performance of a neural network in identifying the security of knee OA, in which the diagnostic performance of the deep neural network was compared with previously employed networks, followed by KL grading criteria. For performing this experiment, over 4000 knee radiology images of patients ranging from the years 2003-17 were selected, and a deep learning model was trained on those images. The model was trained only by images at one stage & clinical information was added to training data at the second stage to check the model's performance. The results after training the model with 2 different scenarios show that by adding the patient's clinical information to the dataset, the performance of the deep learning classifier can surely be improved rather than using only images that were employed at one stage. The experimental findings of this research were comparable with previously held diagnostic performances using knee osteoarthritis KL grading criteria.

(Tiulpin and Saarakkala, 2020) implemented a deep learning algorithm for automatic grading of knee osteoarthritis using an image dataset, containing plain radiology images. Knee Osteoarthritis (OA) is one of the most common diseases in the world, and the examination of this type of disease is usually held via radiographs. In general, the security knee OA is calculated by a well-defined index, known as Kallgren-Lawrence (KL) grading, which rates the severity of knee OA in 5 possible classes. The proposed method in this research is composed of using residual (ResNet) CNN network via transfer learning (TL) approach, trained on OAI dataset which stands for osteoarthritis initiative. The model employed in this research yielded an outstanding AUC & precision score for detecting OA from radiographs. The results of this study are claimed to be state-of-the-art.

(Ahmed and Mstafa, 2022) provided a comprehensive survey on bone segmentation techniques in knee OA research, ranging from traditional methods to applying deep learning algorithms. This study tells that Knee Osteoarthritis (KOA) is a degenerated joint disease that happens to be the most in middle-aged and elder persons. Early diagnosis of KOA is crucial to keep someone avoided from being affected by this chronic disease. The prediction of development of KOA could help in making an effective early diagnosis of this disease, which is a big challenge. The authors of this study aim to provide a thorough overview of several cutting-edge approaches for segmenting the knee articular bone. From conventional methods to deep learning-based approaches, segmentation methods enable the calculation of the rate of articular cartilage loss, which is used in clinical practise for gauging disease progression and morphological change. Additionally, the goal of this effort is to provide scholars with a broad overview of the approaches that are currently used in the field.

(Guan et al., 2022) attempted to predict the pain progression in knee osteoarthritis (KOA) using deep learning methods, in which with the application of using demographics, clinical, and radiological risk indicators, an artificial neural network (ANN) was used to create a conventional risk assessment method to forecast the course of pain. The DL assessment of initial knee radiographs was coupled with demographics, clinical, and radiological risk variables in a combination model. The hold-out testing dataset was used to conduct an area under the curve (AUC) analysis to assess model performance. After applying these techniques, the traditional model depicted an AUC of closer to 0.7 (70%), while deep learning approach showed 0.77 AUC score (77%). The approach which was employed by combining traditional & deep learning (DL) methods showed the highest AUC of 0.807 (80.7%). Summarizing this study, the model proposed in this study outperforms all traditional biomedical approaches used of predicting pain progression in KOA.

(Yong et al., 2021) used an ordinal regression module for classification of severity of knee osteoarthritis into multiple classes, using Kallgren-Lawrence (KL) grading factor as the deciding metric to detect severity, taking this problem as an ordinary regression task. The model proposed in this research uses an input from an artificial neural network (ANN) to produce 4-cut partitions, in a way to classify a given picture of radiology knee image into one of 5 separate classes (KL grades). The performance of this approach promises notable improvements on knee OA severity prediction.

(Chen et al., 2019) suggested a fully automated framework for identification of knee osteoarthritis (KOA) by applying deep learning (DL) techniques. For the purpose of calculating neural network loss, this study also includes an ordinal loss function. According to this study, KOA is a serious illness that is among the leading causes of impairment in the elderly, and the early detection of this disease is crucial in order to prevent any sort of progression in KOA and could help in reducing this disease. Two convolutional neural networks (CNNs) have been used in this study for the intelligent identification of KOA in X-ray images based on the KL grading system. This work used a modified one-stage YOLOv2 algorithm to first analyse the size of knee joints scattered in X-ray pictures with low variability and then identify knee joints. The most well-known CNN models, including ResNet, VGG, and Dense Net as well as InceptionV3, are tweaked in the second stage to conduct classification of the observed knee joint pictures with a novel adjustable ordinal loss. The skill of applied models is tested on a famous dataset called as Osteoarthritis Initiative (OAI), and a Jaccard score of 0.85 and recall of 92%. After applying VGG-19 on KL-grading task for KOA, the best accuracy achieved in this research came out to be 69.7%, which is claimed to be state-of-the-art.

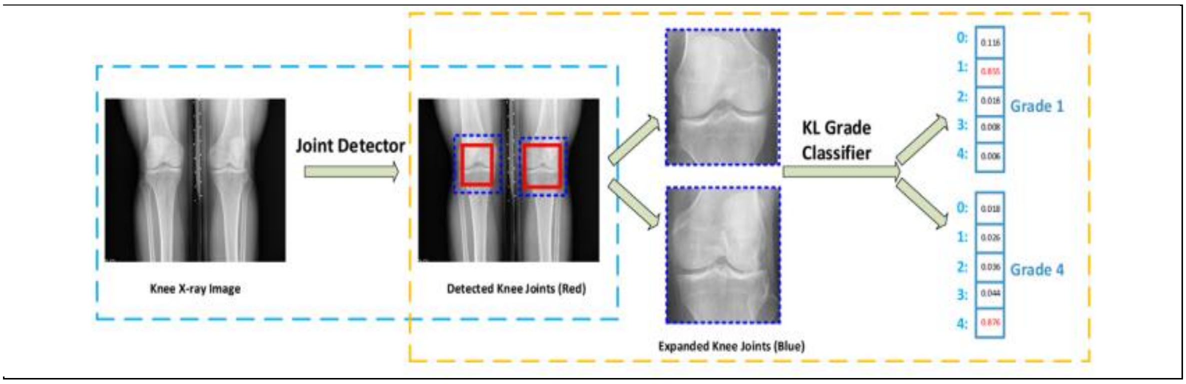


Figure 3Methodology Applied by Chen et al., (2019)

(Thomas et al., 2020) used deep neural networks (DNNs) for performing automatic classification of knee osteoarthritis (KOA) by using radiographic knee images. The study has applied deep convolutional neural networks (CNNs) for this task, and proper image data pre-processing steps such as data augmentation, are also applied. The model is developed by using more than 30,000 images, validated on 4074 image samples and for testing purpose, 4090 images are used. The results of the model are compared to individual radiologists by using the KL-grading system. The results of this approach indicate an average F1 score of 70%, a test score of 71% when full test set was used. After using just 50 images in the test set, the F1 and accuracy results both came out to be 70%. The authors of this study claim the results to be comparable as well as radiologists do.

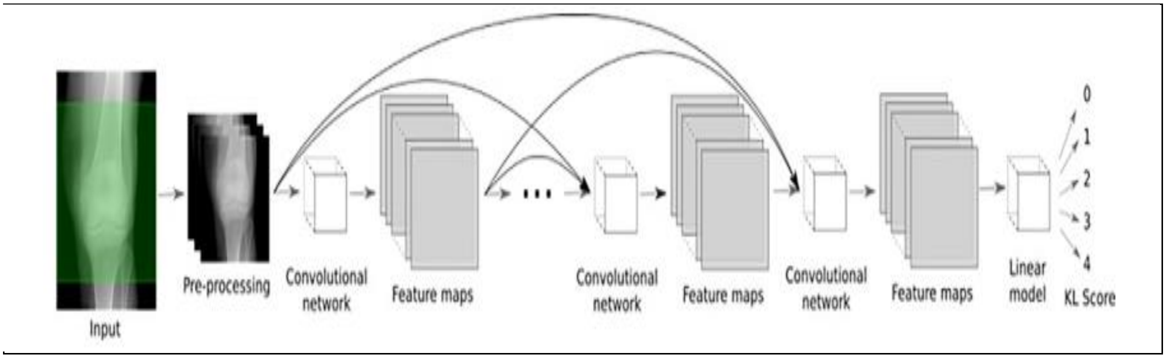


Figure 4: CNN Methodology Applied by Thomas et al., (2019)

(Jain et al., 2021) performed knee osteoarthritis (KOA) severity classification by applying an attentive deep learning framework, which was composed of multi-scale deep convolutional neural networks (CNNs). This study demonstrates that Knee Osteoarthritis (KOA) is a debilitating joint condition that affects millions of people worldwide and is characterized by joint stiffness, pain, and cognitive incapacity. It is often evaluated by taking into account physical symptoms, medical history, and further joint screening procedures such radiography, MRIs, and CT scans. Traditional approaches have not been proved enough in detecting severity at early stages, therefore there is a need of an advanced approach. Therefore, a new methos is presented in this study, called as OsetoNet in this study's terminology, which is a deep learning framework. The grading criteria for determining OA severity is based on KL-grading method. The High-Resolution Network (HRNet), one of the most current deep learning models, is the foundation for the new technique this study proposes to capture the multi-scale properties of knee X-rays. Additionally, a functional form is used to weed out unhelpful features and improve the effectiveness of the algorithm in use. For enhancing the learning of network, the researchers have also used gradient-based class-activation maps. When this neural network (HRNet) is used to the benchmark OAI dataset, it achieves an accuracy of 71.74 percent and an MAE of 0.311, which is stated to be superior results than those of earlier studies.

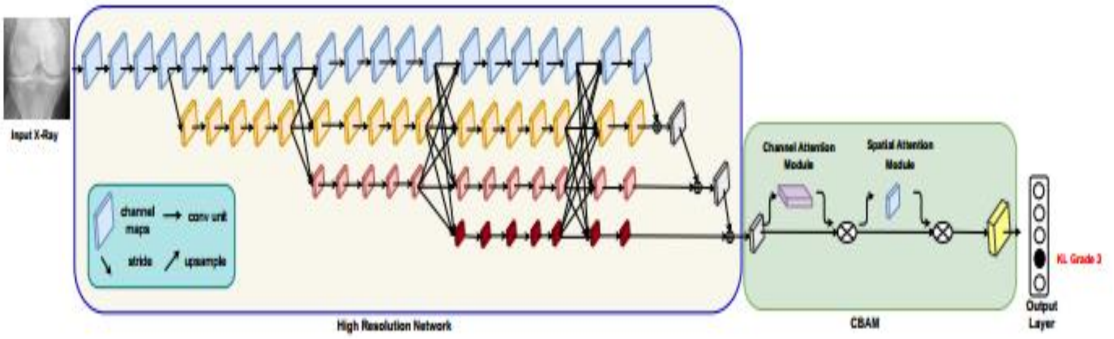


Figure 5: CNN Methodology Applied by Jain et al., (2021)

(Yeoh et al., 2021) highlighted the emergence and developments of deep learning techniques in knee osteoarthritis (KOA) detection, by indicating that traditional approaches used in computer-aided diagnosis have proved to be a benchmark in biomedicine, but it lacks the capability of detecting multiple variations at a time, especially when it comes to the point of X-ray scans. One of the most prevalent forms of arthritis, KOA is a devastating condition that severely impairs elderly persons. To overcome these limitations, these techniques need to be replaced with deep learning methods, such as applying convolutional neural networks (CNNs), which work best on image-based datasets. CNNs have shown benchmark performances on medical imaging datasets and these can also be used effectively for diagnosis of knee OA. In recent years, researchers have also created 3D CNNs to evaluate the joint issue for a more precise diagnosis, keeping in mind that the changes in a knee joint are a 3D complexity. This study has analysed the development of 2D and 3D CNNs for detecting knee OA. To complete their work, the researchers also reviewed more than 70 articles, some of which included deep learning techniques including segmentation and classification. Additionally, the development, growth, and potential applications of 3D CNNs' technical hurdles are briefly emphasised.

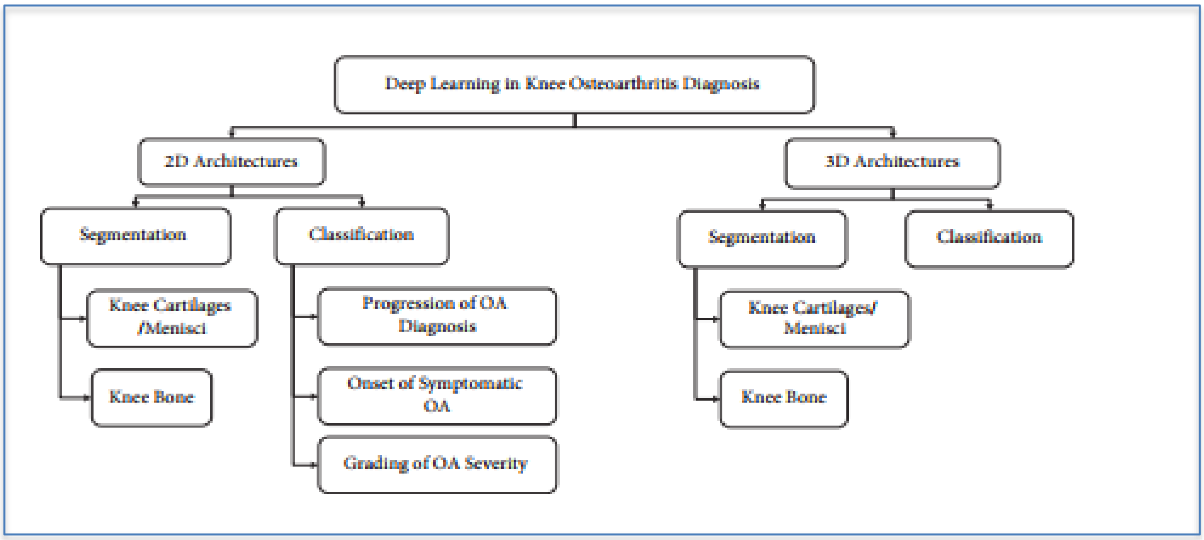


Figure 6: Brief Structure of Review Provided by Yeoh et al., (2021)

In order to diagnose knee osteoarthritis (KOA) and the evolution of OA in the knee, (Cheung et al., [17]) placed more emphasis on multiple-joint space width than single-joint space width. Using a deep learning (DL) technique, the researchers examined the prediction effectiveness of the multiple-joint space width (JSW) and the minimum-JSW on the severity and development of knee osteoarthritis (KOA). For the segmentation of knee X-ray images, a convolutional neural network (CNN) driven by Rest-Net architecture was created, and it was able to acquire an intersection over union (IoU) segmentation score of 98.9%. The difference between the radiologist's assessment and the CNN-based prediction was 0.7801 ($p < 0.0001$). The radiographic severity and development of KOA as determined by Kallgren-Lawrence (KL) grades were then predicted using the approximated JSWs by the XGBoost algorithm. The developed model in this study produced an AUC score of 0.621, which beats the previous approach that used minimum joint-space width (JSW) and achieved an AUC of 0.554. Summing all things at one place, the researchers have developed a system for accessing the severity of knee OA by showing that working on multiple JSW instead of minimum JSW can yield more demanding results for preventing KOA.

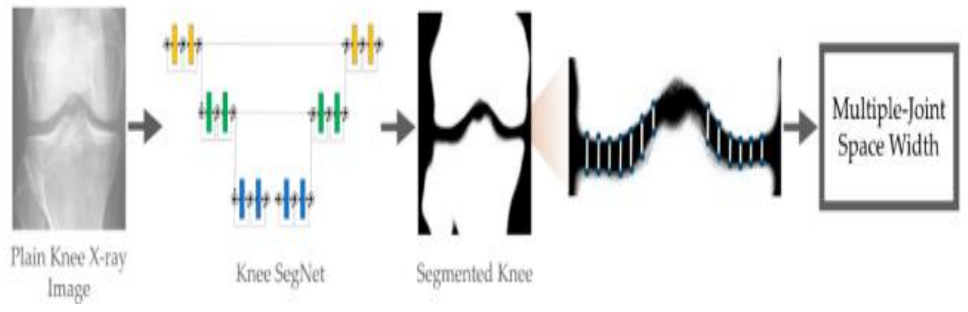


Figure 7: Multiple JSW Approach Used by Cheung et al., (2021)

(Brahim et al., [16]) provided a decision-support tool for making an early detection of knee osteoarthritis (KOA) by applying machine learning techniques using X-ray images. The dataset is pre-processed using a Fourier filter, followed by the use of a normalisation strategy using multi-variate linear regression (MLR). The purpose of applying MLR is to differentiate healthy samples from KOA images. To minimise the dimensionality of the data, a method known as independent component analysis (ICA) is employed for feature identification and extraction. In the end, the classification operation is accomplished by the Naive Bayes (NB) and Random Forest (RF) classification algorithms. The performance of this model has been tested on 1024 knee X-ray images that are obtained from OAI database and the results of this model have shown an accuracy of 82.98%, sensitivity and specificity scores of 87.15 percent and 80.65 percent.

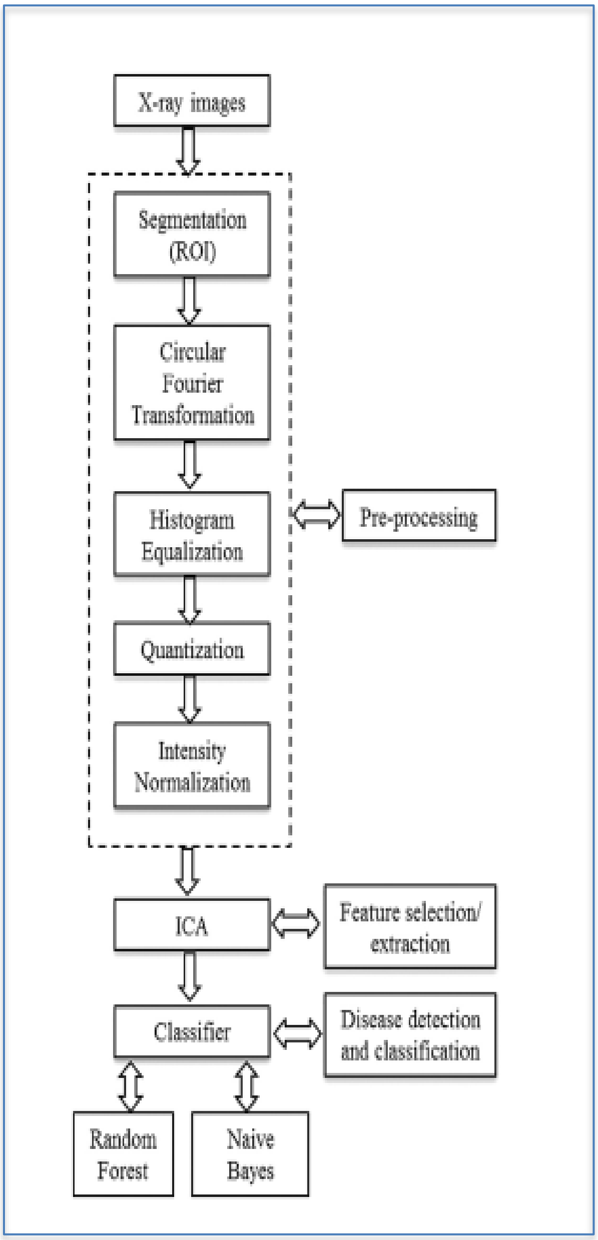


Figure 8: Approach Used by Brahim et al., (2019)

(Christodoulou et al., [21]) provided a study about the applications of deep neural networks for classification in dealing with complex health problems such as Knee Osteoarthritis. This study shows that osteoarthritis (OA) is the most prevalent joint condition, with a wide range in the severity, prevalence, and sequence of symptoms. A vast variety of characteristics and health conditions must be evaluated for OA, most of which are connected to medical hazards including advanced age, gender, hormonal state, body size or weight, family history of illness, etc. For dealing with this problem, this research has introduced a deep neural network-based framework for classifying knee OA based on different factors. Classifying several categories of comparison group using self-reported medical evidence and establishing a class of knee OA diagnosis demonstrate the methodology's effectiveness. The results of this deep learning approach came out to be with 79.39% accuracy, which is also compared with other machine learning techniques which show this methodology’s superiority.

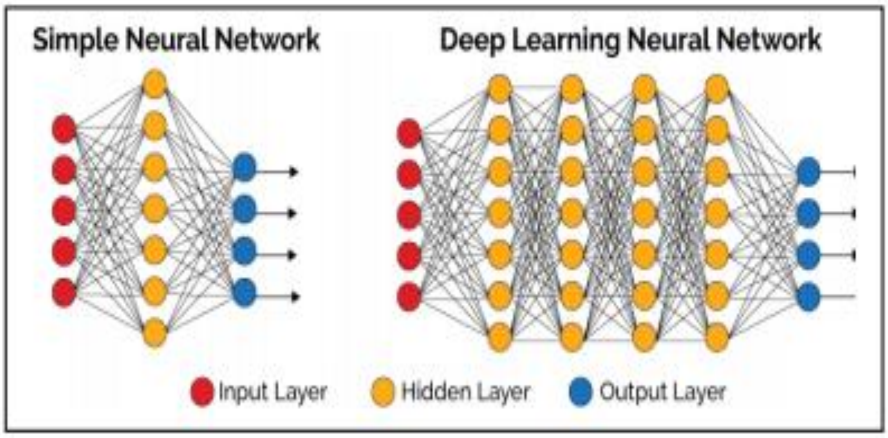


Figure 9: Deep Neural Architecture Used by Christodoulou et al., (2019)

(Primorac et al., [33]) provided a review of pathogenesis and non-operative therapeutic considerations in context of knee osteoarthritis (KOA). This research has provided indication that the most prevalent progressive musculoskeletal ailment, knee osteoarthritis, has drawn the attention of experts. Knee OA is one of the main causes of impairment in elderly individuals, with an estimated incidence of knee osteoarthritis (OA) in persons 60 years of age or older of around 10 percent in men and 13 percent in women. Today, it is recognised that osteoarthritis is a condition that impairs all of the tissues in the joint and results in observable changes in histology, metabolism, and function rather than a disease characterized by the loss of cartilage caused exclusively by mechanical stress. The epigenome, which controls all gene factors through Epigenetic modifications, histone changes, and mRNA interference, is another important factor in the course of osteoarthritis. Numerous efforts have been made to provide non-surgical therapeutic alternatives to reduce the progression of osteoarthritis and delay or possibly replace invasive procedures like total knee arthroplasty and it is hoped that gene therapy may eventually be available to orthopaedic sufferers.

(Kwon et al., 2019) Lim et al., [28] performed early detection of knee osteoarthritis (KOA) by applying deep neural network-based approaches. This study underlines that KOA is a serious disease that is affects elderly people which may end up in disability if not detected early. Early detection of osteoarthritis is crucial for active diagnosis and prevention of this disease. Experts routinely identify osteoarthritis by manual examination of patient medical photos that are often gathered in hospitals. For patients, determining whether osteoarthritis is present takes some time. Deep learning techniques are now being used to identify knee OA from medically produced photos, which are hard to get from hospitals. For this reason, neural networks are applied in this study using statistical data on patients, which includes information on their health behaviours, medicine usage, etc. The authors have extracted characteristics from statistical data for OA detection using Principal Component Analysis (PCA) with a quantile transformer network. This approach employed with deep neural network (DNN) yielded an AUC score of 76.8, and this approach is supposed to be promising in using feature information to detect knee OA.

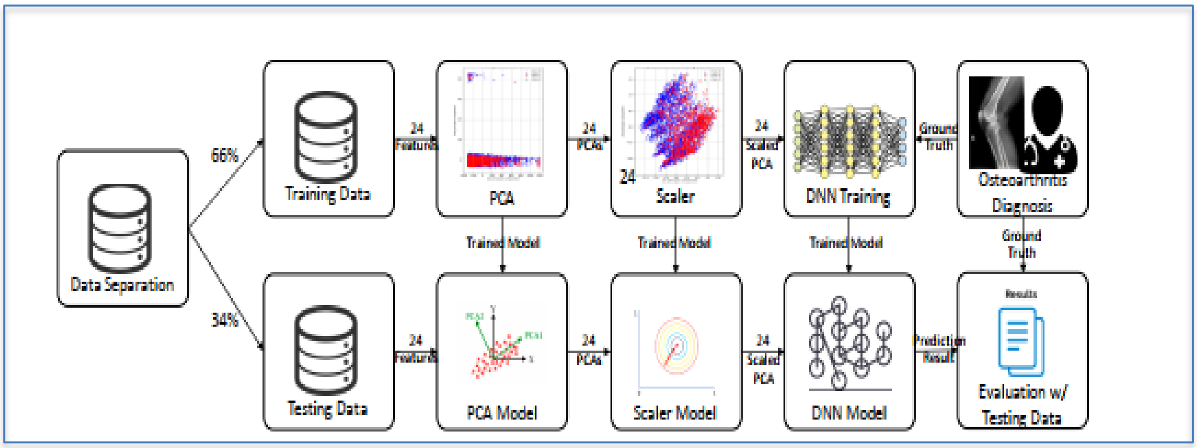


Figure 10: Architecture Used by Lim et al., (2019)

(Kwon et al., [13]) performed identification of gait features associated with radiological grade of knee osteoarthritis. The main focus of this work has been on characteristics that potentially serve as structural biomarkers for the severity and development of OA. After being statistically verified, the discovered characteristics were further investigated by creating a categorization model which is based on a machine-learning algorithm. This study has included 227 patients to get information about KOA, and a total of 165 features have been extracted. A machine learning classification model is subsequently developed utilising the Random Forest method after the feature selection procedure was carried out using neighbourhood component analysis. The KL-grading process has been used to determine the severity of osteoarthritis. The severity grading of knee osteoarthritis may be distinguished by a number of factors, including the joint score of 9 and other aspects from other body parts. The t-test also reveals that certain characteristics dramatically change across age and severity groups, while others are exclusive to each KL grade. The performance of the model was based on AUC score, which came out to be 0.97, 0.99, 0.84, 0.89, and 0.90, which stands for AUC scores of individual classes that range from 0 to 4 (KL grades of knee OA).

Conaghan et al., [38] provided a complete survey of the impact of knee osteoarthritis worldwide and the developments in its therapy, in which it is shown that osteoarthritis is one of the leading diseases causing disabilities worldwide. This study is based on exploring the effects of OA for better exploration of treatment strategies. A 52-question online survey of OA sufferers was done in the UK for this reason. The poll included the effects of OA, its diagnosis and treatment, the involvement of medical practitioners, and self-management. The number of responses were 2,001 out of the 4,443 participants, with males responding at a rate of 49 percent and women at a rate of 56 percent, with a mean age of 65. Of all of these people, 52% said that OA had a significant influence on their life, while 15% of respondents had taken retirement with an average of closer to 8 years than planned. According to the study's findings, individuals are still in disbelief about present treatment methods, which should be overcome with more sophisticated medical equipment. There are several opportunities to enhance knee OA therapy, particularly in information establishment and management techniques.

(Abdullah and Rajasekaran, [37]) performed automatic detection and classification of knee osteoarthrosis by applying deep learning approaches. The researchers have developed a tool for diagnosing knee OA by keeping in context the grading criteria given by Kellgren-Lawrence (KL) grading with a purpose to see the performance of AI based techniques in medical diagnosis. Patients having age above 50 having knee OA symptoms have been included in this study. To find the knee joint space width (JSW) area in digital X-ray images, researchers trained the Faster R-CNN algorithm. ResNet-50 with transfer learning is then used to extract features. Transfer learning is used to classify the severity of knee OA using AlexNet, another pre-trained network. The knee joint's X-ray pictures are graded using the Kellgren-Lawrence score by radiologists, and the region proposal network (RPN) is also trained using a manually extracted knee area as the ground truth image. The model proposed in this study has shown the accuracy of 98.5%

in identification of knee JSW area and it has also depicted an overall accuracy of 98.9% in osteoarthritis severity classification.

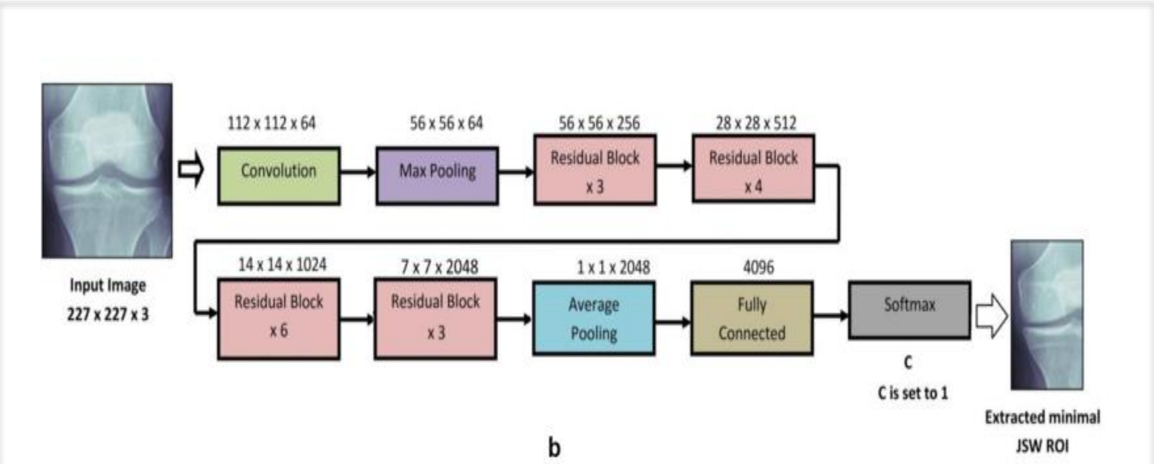


Figure 11: Feature Extraction Principle Defined by (Abdullah and Rajasekaran, 2022)

(Tiulpin et al., [9]) used a deep learning approach for automatic detection of knee osteoarthritis from plain radiology images. This study has underlined that Osteoarthritis (OA) of the knee is the most prevalent musculoskeletal condition. Its diagnosis is presently made by the evaluation of plain radiographs and symptoms, although this procedure is subjective. This study has presented a novel and enhanced computer-aided diagnosis technique using deep Siamese convolutional neural network (CNN) for automated grading of Knee OA severity. KL-grading criteria has been employed as a tool for calculating osteoarthritis severity. After training this neural network on OAI dataset by randomly selecting 3000 samples, this research has achieved a quadratic kappa value of 0.83 and an average accuracy score of 66.71% on all classes. The results were also compared to the comments of radiologists for confirmation. An AUC of 0.93 has also been achieved after training this model on knee OA dataset. This research claims that the results are useful in making clinical decisions in terms of knee OA classification.

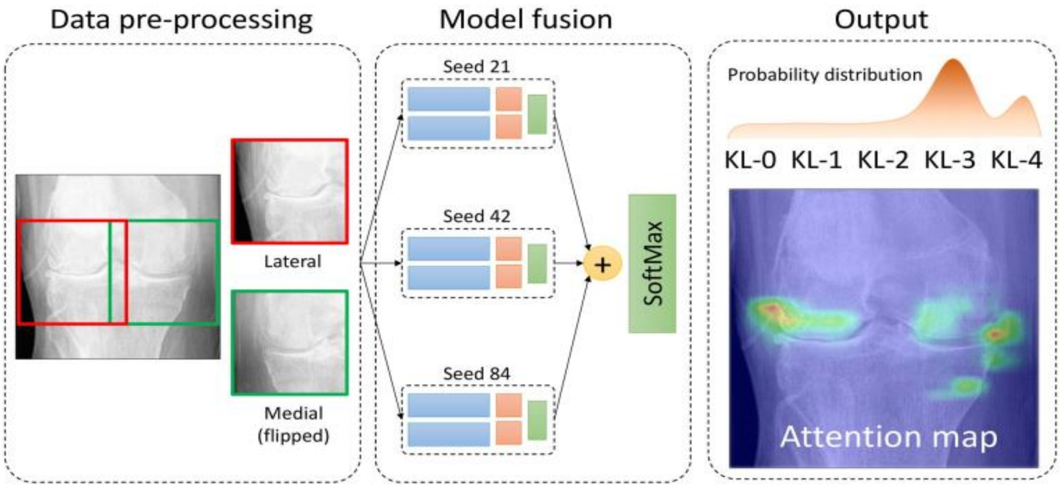


Figure 12: Proposed Model Framework Used by Tiulpin et al., (2018)

Convolutional Neural Networks (CNNs) are a special type of neural networks that are specifically designed to work well on image-based data. A basic convolutional neural network comprises of a convolutional layer which gathers the feature information from various regions of image, while pooling layer is the one which compresses the size of image to make use of only useful information that is required for the task. Pooling layers can be max-pooling, average pooling or can also be taken as global average pooling based on the nature of problem. CNNs have been widely used for a variety of image-related tasks such as image classification, image segmentation, object detection and recognition, facial recognition, and much more applications. After a bunch of convolutional and pooling layers, fully-connected layers come which flatten the input provided by previous pooling layer to get it forwarded to feed-forward network. The number of output neurons in a CNN depends on the number of classes the dataset has, i.e., if dataset has N number of classes, then the last layer will be comprised of N number of neurons. In case of binary classification, the number of neurons in final layer will be one. Below is a diagram that shows the basic structure of the working of a CNN.

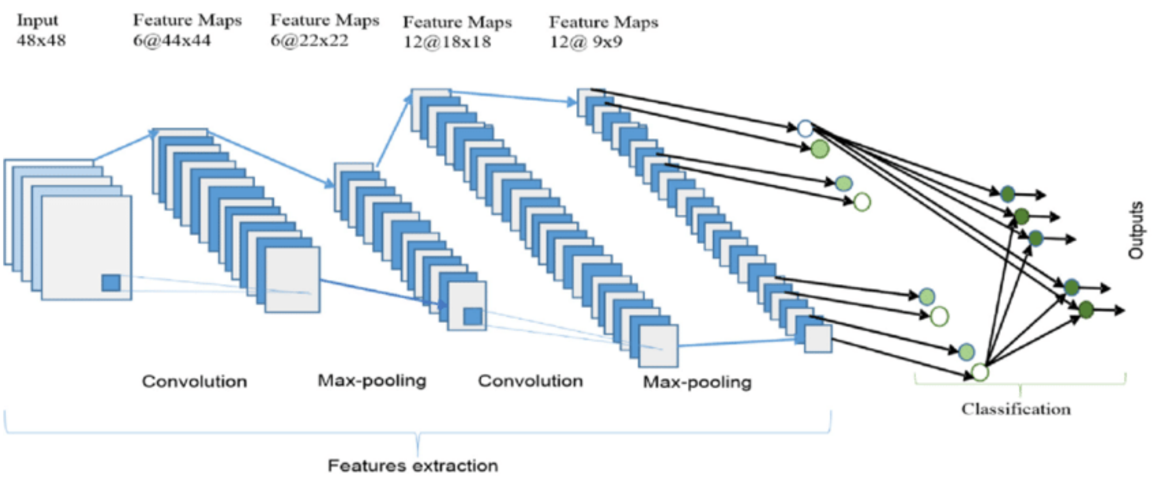


Figure 13: A Basic Functionality of a Convolutional Neural Network

Paper	Year	Headline	Accuracy	Precision (%)	Sensitivity (%)	Specificity (%)	Journal
Kokkotis et al., (2020)	2020	A systematic review of machine learning in knee osteoarthritis (KOA). Research outlines ML as recommended approach for diagnosis of KOA	-	-	-		Elsevier

Schwartz et al., (2020)	2020	Determination of if a CNN can address KOA on plain radiographs	Skill of model was decided on Interclass correlation coefficient (ICC), which was 0.703	-	-		Elsevier
Norman et al., (2018)	2018	Densely Connected CNN for KOA staging in plain radiographs	-	-	77.2	91.52	Springer
Guan et al., (2021)	2021	Deep Learning (DL) approach to predict pain progression in KOA	-	-	72.3	80.9	Springer
Yong et al., (2021)	2021	KOA severity classification using ordinal regression module. This approach	88.09	-	-	-	Springer

		takes the problem of KOA as ordinary regression problem.					

Table 1: Comparison of results from literature

Chapter 3 - Methodology

3.1 Research Methodology

There is 6 steps of the Methodology:

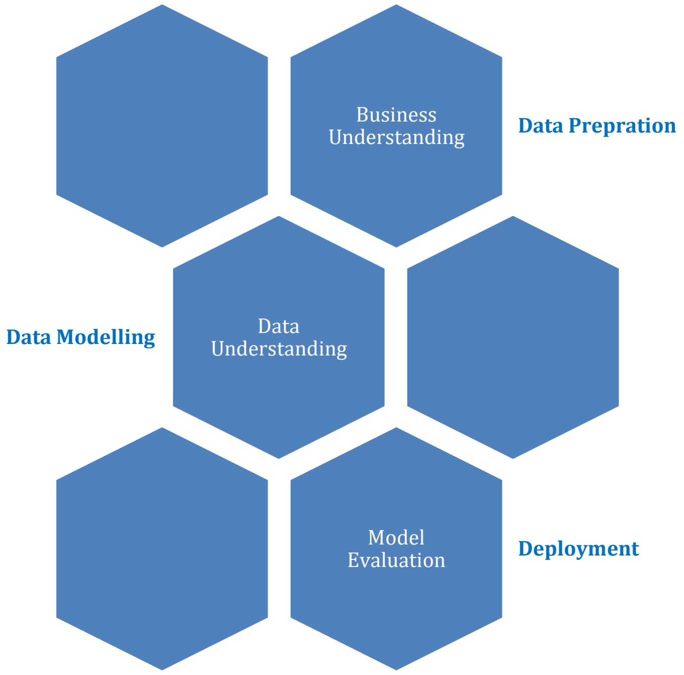


Figure 14: steps of the methodology

The proposed research methodology for undergoing this research is the way in which an AI system will be developed and researched to identify the severity of Knee Osteoarthritis. The research methodology comprises of a proposed intelligent system using artificial intelligence and neural networks (deep learning) that is able to diagnose diseases like Knee Osteoarthritis. This section is sub-divided into multiple sub-sections, which comprise of a proposed system that contains an input phase, the middle phase and the functional phase in which the data used to make intelligent system is fed into the neural network for intelligent classification or identification.

The first phase is the input phase, in which the knee osteoarthritis data used for intelligent classification is given as in input in the workspace. The dataset is composed of X-ray images belonging to 5 classes, with class 0 indicating a healthy scan & class 4 indicating a severe knee disease. After the data is input in the system, the data needs to be converted into real numbers using data pre-processing, since computers only understand numbers and not images or other representations. The next phase is the preprocessing step, also known as middle phase in our terminology.

The second phase, also known as the middle phase, is the preprocessing phase in which input data is pre-processed using data pre-processing techniques, with a purpose to make data ready to be modelled. The pre-processing techniques include data augmentation, which is a way to increase the number of image samples, applying necessary computations to images like horizontal flip, vertical flip, adjusting image ratios, etc., data scaling, wrangling, and so on. After the pre-processing step is completed, the images will be transformed into numerical format, and they now can be input to a neural network that takes them as input and then learns data representations to make predictions on unseen examples. Data available from online sources maybe unstructured, imbalanced and uncleaned, that's the reason pre-processing step plays a central role in data preparation for deep learning.

The final phase is the functional phase, in which the pre-processed data can be fed to convolutional neural network (CNN) to learn feature representations from dataset. The dataset learns the representations from training data to perform well on test data. For monitoring the skill and training of model, a loss function and the accuracy metric is used. Since the proposed system is a classification model, in which the target is to identify the knee osteoarthritis severity, therefore loss function will be chosen accordingly. Adaptive moment estimation (Adam) optimization and the cross-entropy loss function for the model may be utilized to get the best accuracy and least loss possible.

There are several reasons towards using deep convolutional neural networks (CNNs) instead of conventional machine learning classifiers. First, neural networks function in layers, with a large number of neurons communicating and sharing information to obtain the best weights. Convolutional and pooling layers in CNNs, which are created particularly for classifying images, extract the most important features from the images. These numerical features which are in matrix format are then passed to fully connected layers for classification.

A pre-trained CNN that has previously been trained on a much bigger dataset, like ImageNet Benchmark, is suggested for the proposed deep learning model in our research. Using a pre-trained network in a deep learning model is referred to as transfer learning (TL). The pre-trained networks have previously been trained on a much larger dataset like ImageNet, comprising of millions of images & 1000 classes. Using a pre-trained CNN model will be helpful because the study's objective is to increase the AI system's accuracy. The new CNN for our problem statement may be constructed using the weights of pre-trained convolutional neural networks like Inception Net, Exception Net, etc

From the above-mentioned steps, we have proposed this research methodology for undergoing an intelligent system that may be able to classify and predict the severity of Knee Osteoarthritis and to control this disease using deep learning techniques.

3.2 project Timeline:

The estimated time period for completing the tasks of Intelligent Identification of Knee Osteoarthritis Severity Using Deep Learning model in this study is shown in the Gantt chart below:

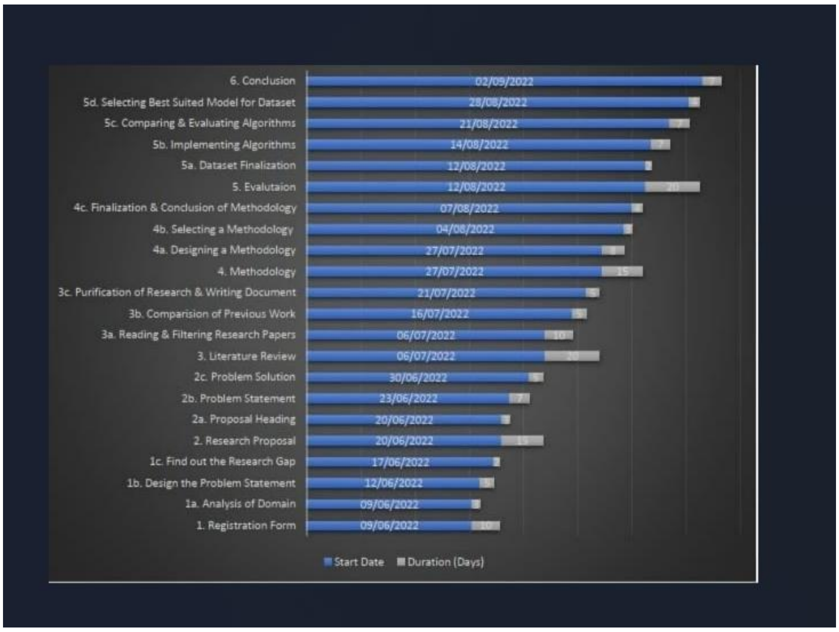


Figure 15: Project Timeline

3.3 Algorithm Explanation

The performance, assessment, and experimental findings of the deep learning model used to perform image classification on the knee osteoarthritis dataset will be performed in this section. The performance of a deep learning CNN model on train and test datasets will be discussed. Since the problem to be dealt is a classification problem, so accuracy score is chosen as the metric for measuring the model's performance. The goal of this study is to implement a system with a design that will demonstrate that the suggested architecture outperforms existing architectures.

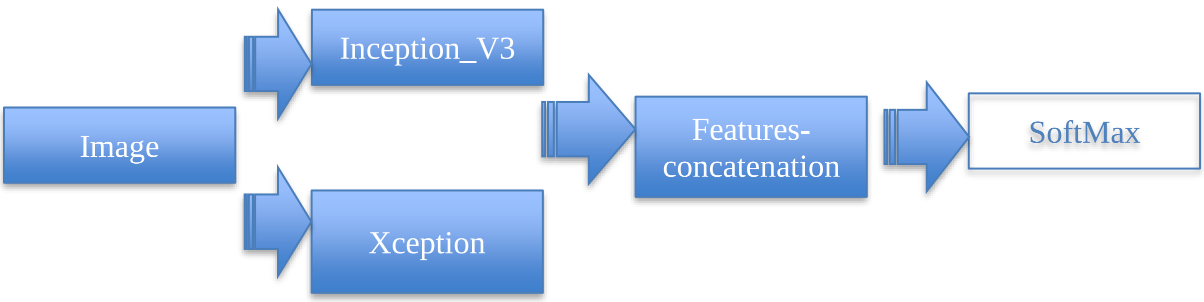


Figure 16: Workflow of model

3.4 Introduction to Problem Statement

Degeneration of the articular cartilage, a flexible, slick substance that typically shields bones from joint friction and impact, is what constitutes knee osteoarthritis. The disorder can also damage neighbouring soft tissues and results in alterations to the bone that lies beneath the cartilage. By far the most prevalent form of arthritis to result in knee discomfort is knee osteoarthritis, which is frequently referred to as just knee arthritis. Rheumatoid arthritis, reactive arthritis, and many other uncommon kinds of arthritis can also hurt the knees.

Human knee osteoarthritis (OA) is a disorder that mostly affects cartilage. The cartilage in the legs plays a big part in how the legs move. In OA, the top layer of cartilage breaks down and degrades, causing excruciating pain. The patient with knee pain must see a doctor, who will then examine the patient and test their clinical symptoms before recommending radiographic imaging of the knee. Osteoarthritis treatment must take clinical symptoms into account.

The correct diagnosis of the condition is made by looking at both the clinical symptoms and the radiological criteria. The width of the joint area, osteophytes, sclerosis, etc., are important radiological criteria. Considering the radiological characteristics, Using the Kallgren and Lawrence (KL) grading method, the severity level of the illness is assessed. The most popular way for grading knee joint OA into 5 different grades—basically to determine the severity of the condition—is the KL system.

Radiographic pictures of the knee are particularly susceptible to unintentional errors that could complicate the analysis of bone structures. It can lead to the professionals analysing the knee x-ray and determining the presence of OA taking longer. This work focuses on the early diagnosis and grading of Knee OA using Hu's invariant moments to comprehend the geometric transition of the cartilage region in Knee X-ray images in order to circumvent these issues.

3.5 Dataset Used:

Degeneration of the articular cartilage, a flexible, slick substance that typically shields bones from joint friction and impact, is what constitutes knee osteoarthritis. The disorder can also damage neighbouring soft tissues and results in alterations to the bone that lies beneath the cartilage.

By far the most prevalent form of arthritis to result in knee discomfort is knee osteoarthritis, which is frequently referred to as just knee arthritis. Rheumatoid arthritis, reactive arthritis, and many other uncommon kinds of arthritis can also hurt the knees.

Inception_v3 and Xception : Dataset for this project will collect from this sourcehttps://www.kaggle.com/datasets/shashwatwork/kneeosteoarthritis-dataset-with-severity?select=train The basic purpose of collecting data to develop an intelligent model using deep learning that can detect class. in 5 possible class using deep learning CNN model.

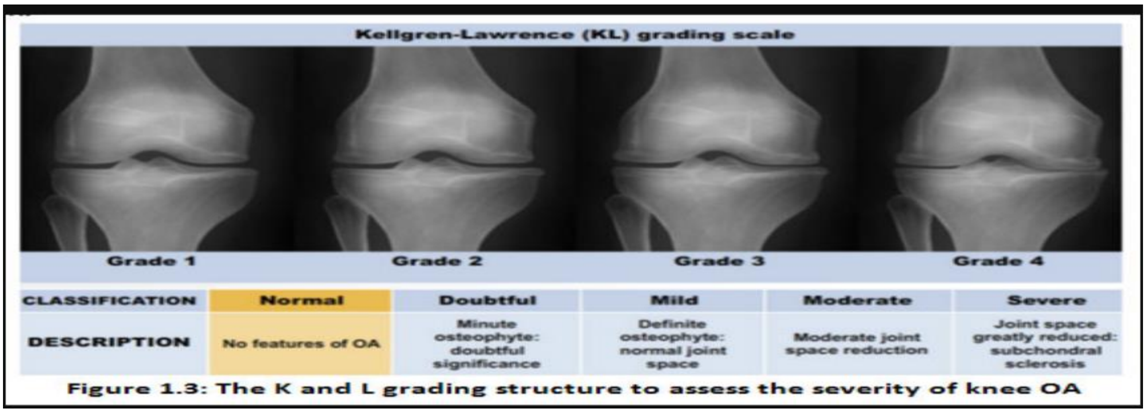


Figure 17: Grading of the 5 classes

- **Class 1**: Healthy knee image.
- **class 2** (Doubtful): Doubtful joint narrowing with possible osteophytic lipping
- **class 3** (Minimal): Definite presence of osteophytes and possible joint space narrowing

- **class 4** (Moderate): Multiple osteophytes, definite joint space narrowing, with mild sclerosis.
- **class 5** (Severe): Large osteophytes, significant joint narrowing, and severe sclerosis.

This Image dataset of 8300 x-ray images. I distributed in 3 types .

Table 2: Distributed Images

Training	5788	5 Classes
Validation	856	5 Classes
Test	1656	5 Classes

3.6 Algorithms Used:

As our problem is concerned with image classification, we have used the following deep learning algorithms:

- InceptionNetV3
- Xception

Algorithms Detail:

1. InceptionNetV3:

The Inception-v3 convolutional neural network architecture uses Label Smoothing, factorized 7 x 7 convolutions, and the inclusion of an auxiliary classifier to transport label information lower down the network, among other advances (along with the use of batch normalization for layers in the side head).

In figure 18, showing that there are layers used in inception_V3 model summary listed below:

Conv2d, activation, global_average_pooling2d, batch_normalization, mixed, concatenate, average_pooling_2d, max_pooling_2d, dense.

```
mixed9_1 (Concatenate)          (None, None, None,    0          ['activation_87[0][0]',
                                768)                              'activation_88[0][0]']

concatenate_1 (Concatenate)     (None, None, None,    0          ['activation_91[0][0]',
                                768)                              'activation_92[0][0]']

activation_93 (Activation)      (None, None, None,    0          ['batch_normalization_93[0][0]']
                                192)

mixed10 (Concatenate)           (None, None, None,    0          ['activation_85[0][0]',
                                2048)                             'mixed9_1[0][0]',
                                                                  'concatenate_1[0][0]',
                                                                  'activation_93[0][0]']

global_average_pooling2d (Glob  (None, 2048)          0          ['mixed10[0][0]']
alAveragePooling2D)

dense (Dense)                   (None, 1024)          2098176    ['global_average_pooling2d[0][0]'
                                                                 ]

dense_1 (Dense)                 (None, 5)             5125       ['dense[0][0]']

==================================================================================================
Total params: 23,906,085
Trainable params: 13,218,181
Non-trainable params: 10,687,904
```

Figure 18: InceptionV3 model summary

2.Xception:

- The Xception convolutional neural network has 71 layers.
- A network that has been pretrained using more than a million images is available in the ImageNet database.
- A keyboard, mouse, pencil, and a few animals are among the 1000 different object categories that the pretrained network can classify images into.
- As a result, the network now includes rich feature representations for many different types of photos.
- The network's picture input measures 299 by 299 pixels.

In figure 19, showing that there are layers used in Xception model summary listed below:

block13_sepconv, conv2d, batch normalization, add, global_average_pooling2d, dense.

```
block14_sepconv1_bn (BatchNorm  (None, None, None,   6144       ['block14_sepconv1[0][0]']
alization)                      1536)

block14_sepconv1_act (Activati  (None, None, None,   0          ['block14_sepconv1_bn[0][0]']
on)                             1536)

block14_sepconv2 (SeparableCon  (None, None, None,   3159552    ['block14_sepconv1_act[0][0]']
v2D)                            2048)

block14_sepconv2_bn (BatchNorm  (None, None, None,   8192       ['block14_sepconv2[0][0]']
alization)                      2048)

block14_sepconv2_act (Activati  (None, None, None,   0          ['block14_sepconv2_bn[0][0]']
on)                             2048)

global_average_pooling2d (Glob  (None, 2048)        0           ['block14_sepconv2_act[0][0]']
alAveragePooling2D)

dense (Dense)                   (None, 1024)        2098176     ['global_average_pooling2d[0][0]'
                                                                 ]

dense_1 (Dense)                 (None, 5)           5125        ['dense[0][0]']

==================================================================================================
Total params: 22,964,781
Trainable params: 0
Non-trainable params: 22,964,781
```

Figure 19: Xception model summary

Chapter 4 – Implementation

4.1 Skills Required:

For this study, since it requires the applications of deep learning and machine learning, a thorough understanding of these fields is needed. Plus, for implementing these skills on a programming language, skill of Python Programming will be a crucial step, plus with hands on skill on following libraries of Python:

- NumPy (for manipulating with arrays)
- pandas (for analysing the data)
- os (for interacting with an operating system)
- random (for generating random numbers)
- matplotlib (for data visualization)
- pyplot (for data visualization and graphs plotting)
- counter (for counting hash able objects)
- cv2 (for image processing and computer vision tasks)
- glob (used for returning all the file paths, matching a specific pattern)
- TensorFlow (for model building)
- Deep learning libraries for model building and evaluation.

4.2 Complete Workflow:

Explained below is the complete workflow, which will elaborate the complete methodology used to make this deep learning model. The complete explanation of this task has been mentioned above, and in the below section we will be describing each component of this task step-by-step.

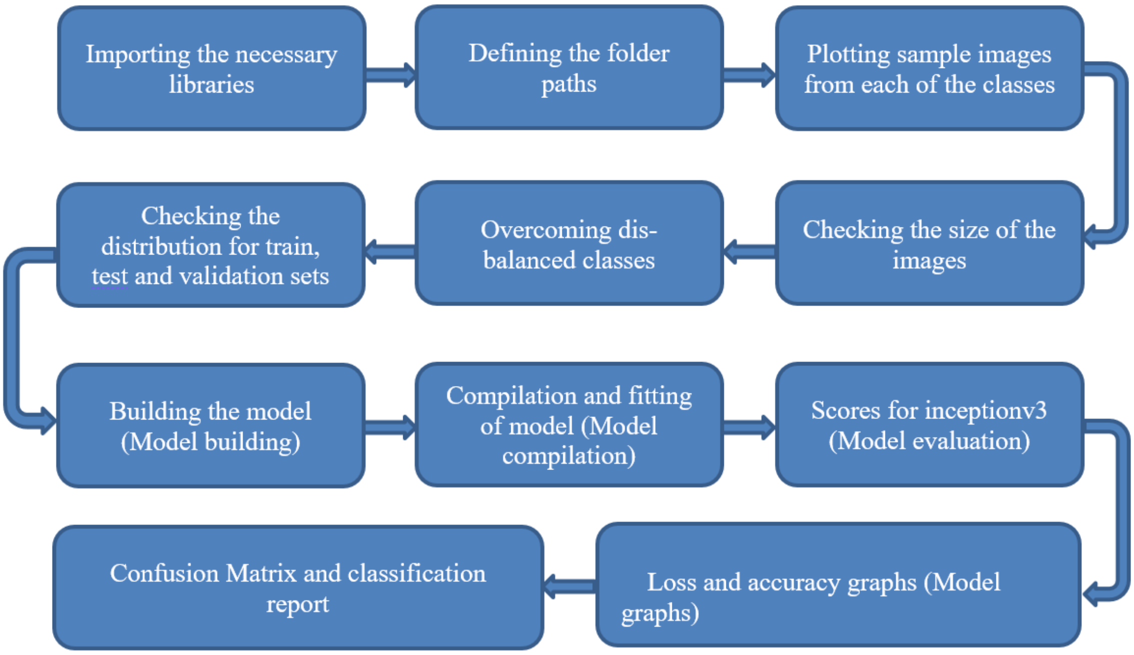


Figure 20: Workflow of inception_v3 and Xception model

1) Importing the libraries & data loading: In the first step, we have imported the necessary libraries for manipulating with image data. [Appendix-F] Below is the attached source code for it.

2) Defining the folder paths: In this step, we have defined our dataset locations. We have stored our dataset on google drive. We have made a base dir. which contains the directory path for the whole dataset. Train directory contains the training images. Similarly testing and validation directories contains the testing and validation images in the given dataset. We have total 5 no of classes in this dataset. [Appendix-G]Below is the attached source code for it.

3) Plotting sample images for each class: In this step, as we have total 5 no of classes, we want to visualize each image from 5 classes. [Appendix-H] Below is the attached screenshot for it

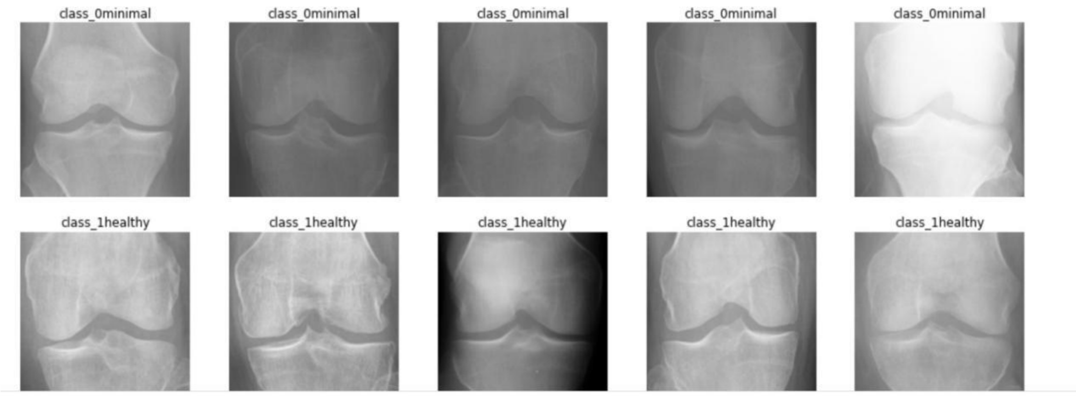


Figure 21: Distribution images

4) Checking distribution for train, test and validation sets:

In this step, we will show the distribution for train, test and validation datasets. There are 5784 images in the training dataset belonging to 5 classes. There are 834 images in the validation dataset belonging to 5 classes. There are 1666 images in the testing dataset belonging to 5 classes. All these are visualized using pie charts

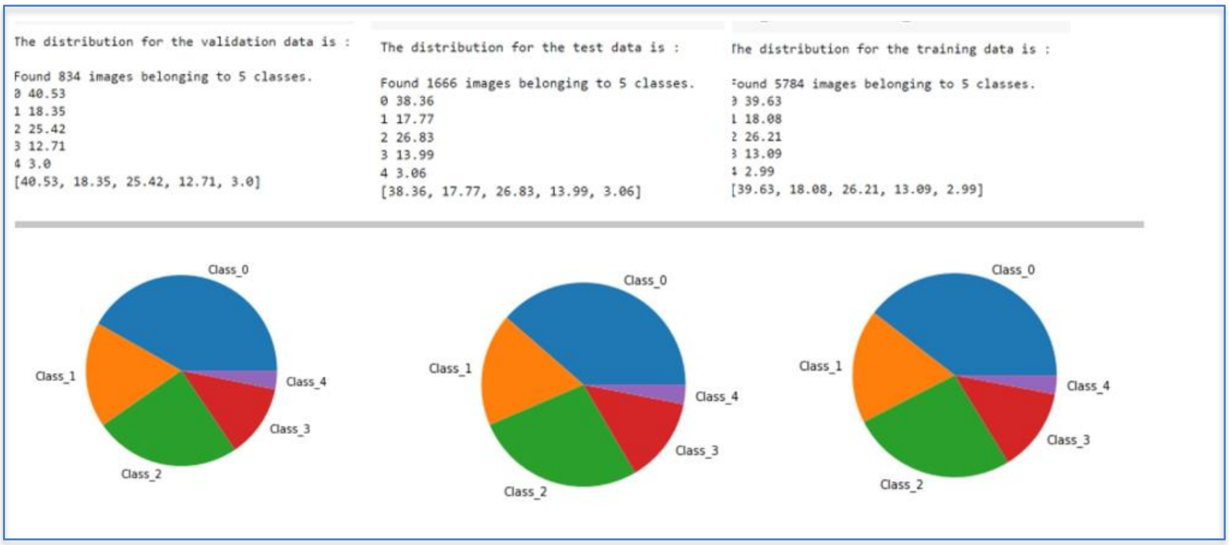


Figure 22: class distribution

5) Overcoming dis balanced classes:

In this step, we have managed to eliminate disbalanced classes by using class weights. We are performing class balancing for accurate results[Appendix-I] Below is the attached code screenshot.

6) Checking the size of images:

In this step, we have checked the size of all the images present in our dataset for further processing. [Appendix-J]The default image size is (224, 224, 3). Below is the attached source code for it.

7) Data Pre-processing: The collected data are typically disorganised and originate from several sources. They need to be cleaned up and standardised before being fed into the ML model (or neural network). Pre-processing is frequently utilised to carry out procedures that lessen the complexity and boost the precision of the given algorithm. Since there isn't a specific algorithm that can be written for every situation where a picture is collected, when we acquire an image, we usually transform it into a format that a generic algorithm can use to solve it**.** In pre-processing I used pipeline to balance the imbalance classes through weights using ImageDataGenerator's. [Appendix-K] This image data generator is very important for image dataset

8) Model building:

In this step, we have first defined some model hyperparameters such as the total no of classes present in our dataset, the image size of all the images presents inside our dataset, the batch size our dataset is divided into. We have created the data generators for loading the data into our deep learning model. We have used call-backs also. The purpose of using call-backs is to apply actions at various stages of the training. For example, early stopping will eventually stop training if our evaluation metric stopped improving. Similarly, reducelronplateau will reduce the learning rate if our evaluation metric stopped improving. [Appendix-L]Below is the attached code screenshot for it.

9)Adding reducelronplateau & early stopping:

- **Reducelronplateau class:**
When a metric stop improving, lower the learning rate. Once learning begins to decline, models frequently gain by decreasing the learning rate by a factor of 2–10.[Appendix-L]

- **Early stopping:** Stop exercising when an observed statistic stops improving.

10) Building the Model

For Building the model I divided in class in 5 and image size and batch size.

NUM_CLASSES = 5

IMAGE_SIZE= [224, 224]

BATCH_SIZE=64

11) Imagedatagenerator: This function handles the original data and convert it to a new transformed data. The Imagedatagenerator constructor's dictionary only lists a few parameters. [Appendix-M] Here is an explanation of them.

- **Rotation range**: integer, range of degrees for random rotations.

- **Height_shift_range** shifts the picture along the height dimension with height shift range. It accepts a range of inputs. The picture must be adjusted by for float.

- if less than 1, a percentage of the whole height; else, a number of pixels.

- **Width_shift_range** shifts the picture along the width dimension with width shift range.

- **Rascal**: Rescaling to None by default. Otherwise, we multiply the data by the given value if None or 0 is returned (after applying all other transformations).

- **Fill mode**: One of "constant," "nearest," "reflect," or "wrap" for fill mode. Nearest is the default. Points outside the input's limits are filled using the specified mode. The ImageDataGenerator's dictionary only lists a few parameters.

Using data generator function I divided classes in 3: Train data generator, test data generator, and validation data generator to check how many images belongs to 5 classes. We got it results like:

Train data generator: Found 5788 images belonging to 5 classes
Test data generator: Found 1656 images belonging to 5 classes
Validation data generator: Found 856 images belonging to 5 classes

4.3 Inception_v3 Model (Background and Motivation)

48-layer convolutional neural network Inception-v3 is a pre-trained version of the network that has previously been trained on more than a thousand pictures from the ImageNet collection. This network can categorise photos into 1000 different item classes, including keyboard, mouse, pencil, and several animals. Consequently, the network has picked up detailed image features for a variety of pictures. The network accepts images of a 299 by 299 resolution. In the first section, the model collects generic characteristics from input photos and then classifies them in the second section using those features.

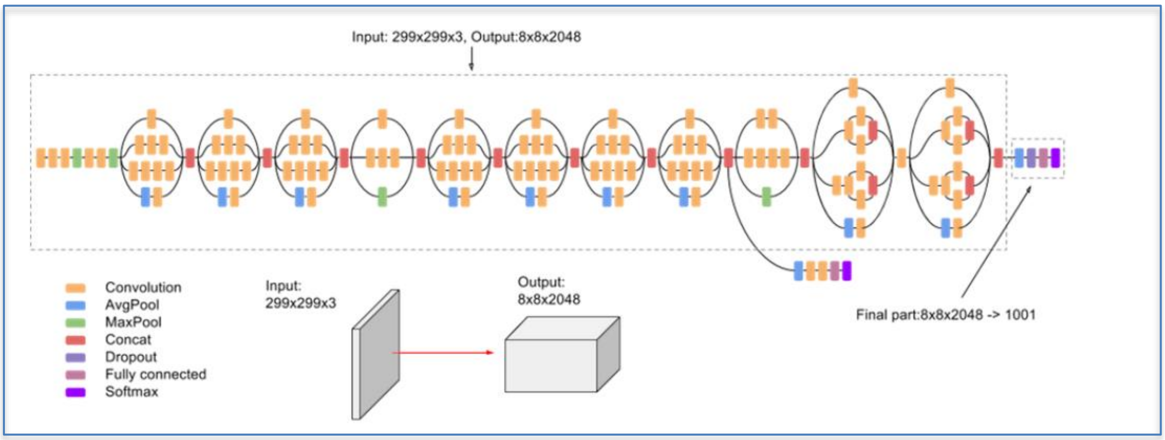


Figure 23: Model architecture of inception_v3 (background)

A popular image recognition model called Inception v3 has been demonstrated to achieve more than 78.1% accuracy on the ImageNet dataset and almost 93.9% accuracy in the top 5 results. The model is the result of several concepts that have been established by various scholars throughout the years. It is based on the original paper: "Rethinking the Inception Architecture for Computer Vision" by Szegedy, et. al.

4.4 Xception Model(Background and Motivation)

Depth wise Separable Convolutions are used in the deep convolutional neural network architecture known as Xception. Researchers from Google created it. According to Google, Inception modules in convolutional neural networks serve as a transitional stage between the depth wise separable convolution operation and ordinary convolution (a depth wise convolution followed by a pointwise convolution). A depth wise separable convolution can be viewed in this context as an Inception module with a maximum number of towers. With Inception modules replaced with depth wise separable convolutions, they suggest an unique deep convolutional neural network design based on this result.

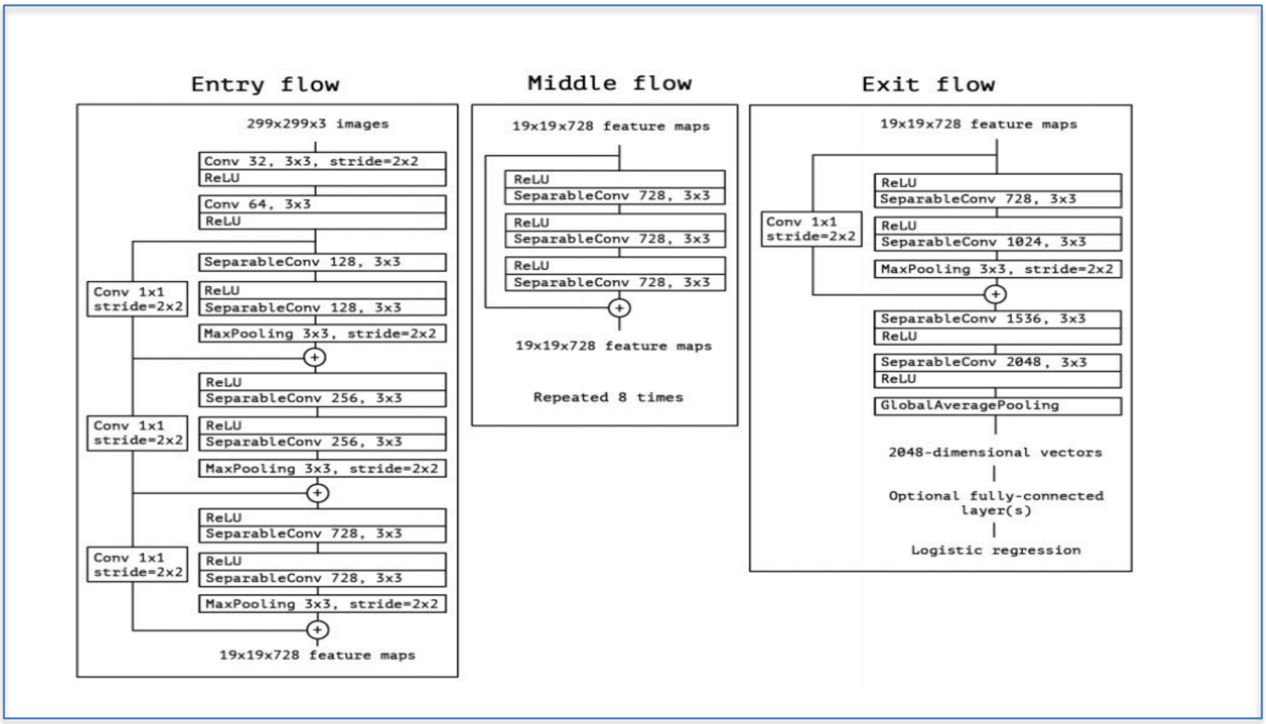
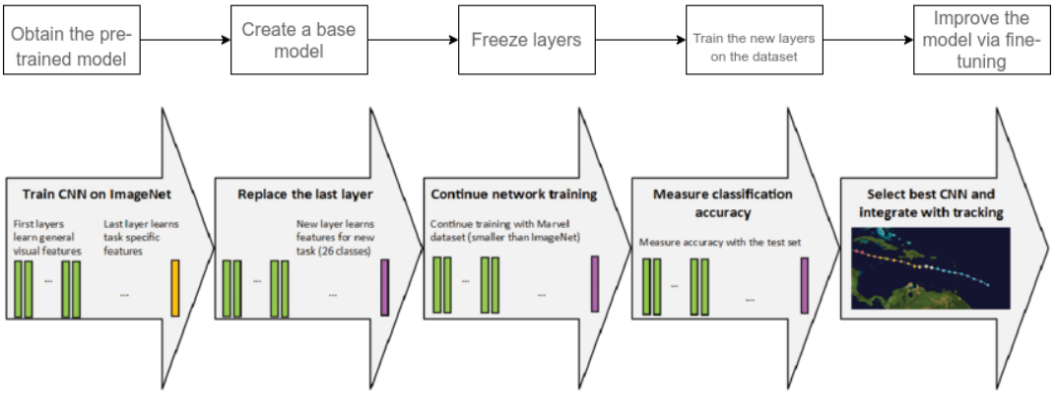


Figure 24: Workflow of Xception model

The entry flow, the middle flow, which is repeated eight times, and the exit flow are all the steps that the data must initially go through. Keep in mind that batch normalization comes after every Convolution and Separable Convolution layer." Chollet, F., 2017. Xception: Deep learning with depth wise separable convolutions."

Transfer Learning has become immensely popular because it considerably reduces training time and requires a lot less data to train on to increase performance.



Source

Figure 25: Implement transfer learning in these six general steps

SoftMax Layer: Our approach to the Inception architecture retrained the last layer of Inception using SoftMax Regression, and we generated probabilities based on evidence we gathered from the pictures. The evidence are computed using a bias-added sum of weights determined by pixel intensity.

$$\text{evidence}_i = \sum_j W_{i,\ j} x_j + b_i$$

The formula for the evidence for a class I given an input of x is shown below. Here, Wi stands for weights, bi for bias, and j for the index used to add up the pixels in the input.

After that, by running the evidence through the SoftMax function, we calculate our probabilities.

$$y = \text{softmax}(\text{evidence})$$

Given an n-dimensional vector, the SoftMax function, also known as the normalised exponential, produces same-dimensional vectors with values ranging from 0 to 1.

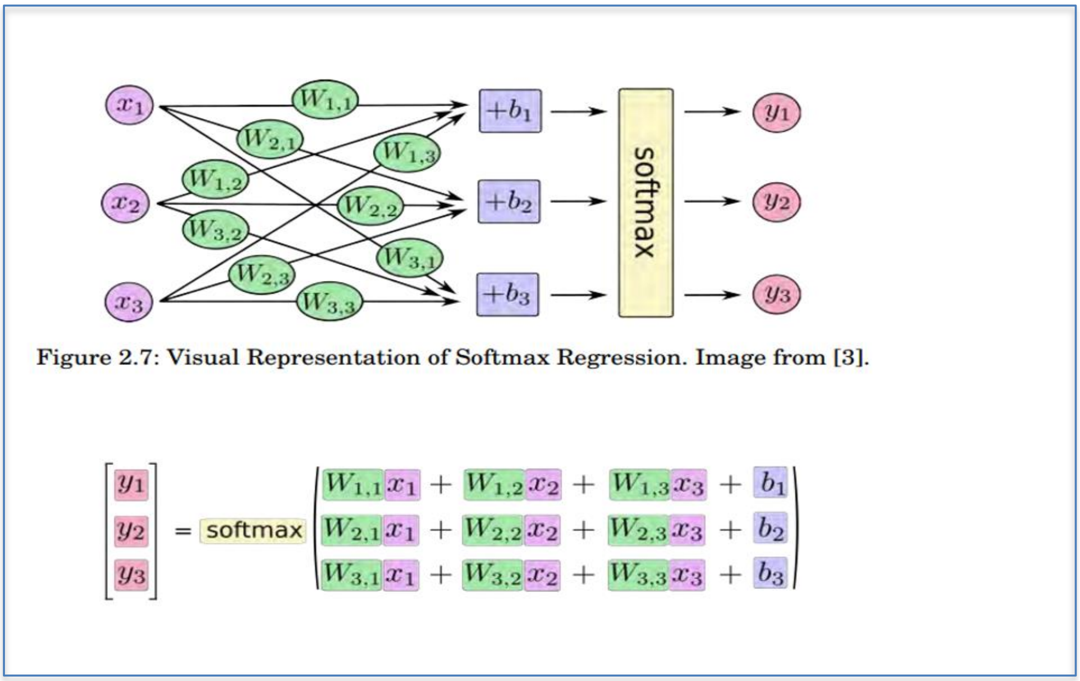


Figure 26: The visualization from can be transformed into above representation showing Probability Matrix in SoftMax

ImageNet Dataset:The model must be trained using a sizable collection of labelled photos before it can be used to recognise images. The dataset ImageNet is often used.

Over ten million URLs of captioned photos are available on ImageNet. Additionally, one million of the photos contain bounding boxes that give the named items a more specific position.

1)Importing libraries for Implement Inception_v3 and Xception:

We will import all necessary libraries before downloading, building, and initialising the model. Below is a snapshot of the code for this stage.To create pre-train model, I used ImageNet weights. All the model's layers will be used, [Appendix-N] with the exception of the last fully connected layer, which is unique to the ImageNet competition.

2)Load the Model:

Here we have loaded inception v3 pretrained model using ImageNet weights. Moreover, for a better result of vectors we have used global average pooling instead of flatten layer. To recognize and classifying images in better way we have added fully connected layer with relu activation function. In the output layer we have 5 classes so for multiple probabilities we used SoftMax activation function.

3)Train model: Freeze layer: During exercise, freeze layers to prevent them from changing. It is crucial to freeze the layers from the pre-trained model. You do not need the weights in any of those layers to also be re-initialized, therefore this is why. You will lost all of the knowledge you have previously acquired if they are. This will be the same as starting from scratch when training the model.

```
for layer in base_model.layers:
    layer.trainable = False
```

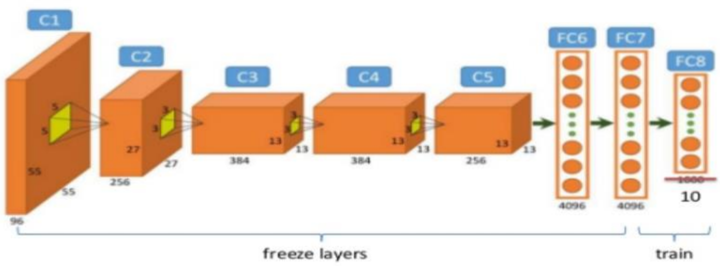


Figure 27: how Freeze layers works

After that we started model training, and we trained only few top layers and freeze convolutional layers. Additionally, after setting layers to non-trainable we started compilation of model with optimizer Adam and metrics used accuracy. Next to it, we have trained our model using model fit on our dataset with epochs of 100 and checked for reducing learning rate. Hereafter, we started fine tuning, for that we have freeze the bottom layer and train the remaining top layers. we chose 2 inception block, and we freeze first 249 layer and we used other layers for training. Next to it, we recompile the model and trained model once again. **Initially,** we trained the data on a basic Inception model using all of the parameter default values, which are:

Table 3: Training prospective parameters description

Training Prospect	**Parameters**	**Description**
Patience	5	By including a delay in the trigger equal to the number of epochs on which we would want to see no progress, we can account for this.
Initial Learning Rate	0.001	First Learning rate using for training
Learn Rate Drop Factor	0.2	Factor causing the learning rate to decline. =0.2 in this instance.
Maximum Epochs	100	The highest number of training epochs.
Minimum Batch Size	64	The size of the mini batch that will be used for each training cycle.
Shuffle	Every Epoch	Before training, mix up the training and validation data once. preventing the same data from being lost each epoch.
Dense(g)	(1024, activation = 'relu') (5, activation = 'SoftMax')	Also known as fully connected connections, uses a linear operation in which each input is connected to each output by a weight.

Chapter 5 – Result and discussion

5.1 Model Evaluation (Inception_v3):

In this step detection includes the Five classes. And the knee arthritis acceptable ratio for it. The effectiveness of CNN's suggested model throughout training and validation based on Knee x-ray scans as shown in figure. Each one's validation was tracked epoch used to calculate performance depending on accuracy at various iterations and epochs. Table displays the training's validation accuracy and Accuracy.

Table 4: Validation accuracy based on epochs

Model Name	Dataset	Epochs	Validation Accuracy	Accuracy
InceptionV3	Collection of Knee X-Ray	1	49%	49%
		25	56%	65%
		50	53%	74%
		100	50%	86%

we have evaluated our model on the training and testing sets. We are achieving an accuracy of 86% on the training dataset and 54.1% on the testing dataset. The test loss is 1.288 and the test accuracy is 0.537. Below is the attached code screenshot.

Loss Vs Accuracy Graphs (Inception_v3):

Figure 28: graphs of training Loss vs accuracy

It can be inferred from the loss graph:

- The red lines show the training loss whereas the green line shows the validation / test loss. The blue dot shows the best epoch
- The x-axis shows the no of epochs whereas the y-axis shows the model loss.
- It is observed that the training loss has started from 1.19 and eventually reached end with some minor ups and downs at 0.61.
- It is observed that the validation / testing loss has started from 1.23 and eventually reached end with some major ups and downs at 1.13 following training loss. The best epoch is 9.
- The interval on x-axis is 10 whereas the interval on y-axis is 0.1.

It can be inferred from the accuracy graph:

- The x-axis shows the no of epochs whereas the y-axis shows the model accuracies.
- It is observed that the training accuracy has started from 0.50 and eventually reached end with some minor ups and downs at 0.75.
- It is observed that the validation / testing accuracy has started from 0.52 and eventually reached end with some major ups and downs at 0.57 following training accuracy. The best epoch is 10.
- The interval on x-axis is 10 whereas the interval on y-axis is 0.1.

- The red lines show the training accuracy whereas the green line shows the validation / test accuracy. The blue dot shows the best epoch.

- **Confusion matrix and classification report (Inception_v3):**

Confusion Matrix is a performance metric used to describe the performance of classification model on data. It’s an N x N matrix, where N is the number of labels. Confusion Matrix is a function that shows a N x N matrix to show the actual vs predicted
outcome of our data. It is a summarized table that shows correct vs incorrect predictions
with count values. The structure of a confusion matrix contains 4 parameters:

1. True Positive
2. True Negative
3. False Positive
4. False Negative

Outcomes that are positive and predicted as True.

1.True Negative:

Outcomes that are negative and predicted as True.

2. False Positive:

Outcomes that belong to negative class and are falsely predicted as of positive class

3. False Negative:

Outcomes that belong to positive class and are falsely predicted as of negative class

4. True Positive:

Outcomes that are positive and predicted as True. Below is a visualization of confusion matrices of Inception_v3 model that we have used blow:

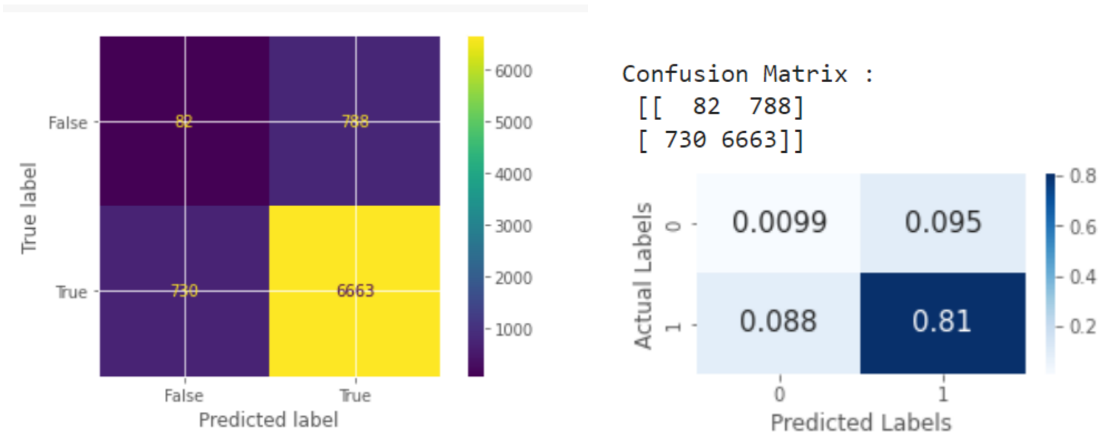


Figure 29: Confusion matrix of Inception_v3 model

Classification Report of the Inception_v3 model:

It is one of the metrics used to assess the performance of classification-based machine learning models. The precision, recall, F1 score, and support of your model are shown. It gives us a clearer picture of our trained model's overall performance. You must be familiar with each statistic shown in the classification report of a classification model in to understand it. I have clarified all of the metrics below so you can quickly comprehend the categorization report of your machine learning model.

Metrics	Definition
Precision	Precision is defined as the ratio of true positives to the sum of true and false positives.
Recall	Recall is defined as the ratio of true positives to the sum of true positives and false negatives.
F1 Score	The F1 is the weighted harmonic mean of precision and recall. The closer the value of the F1 score is to 1.0, the better the expected performance of the model is.
Support	Support is the number of actual occurrences of the class in the dataset. It doesn't vary between models, it just diagnoses the performance evaluation process.

Figure 30: classification report term explanation

In table 00, classification report of inception_V3 model explains precision, Recall, F1 Score, support and accuracy:

Table 5: classification report of Inception_v3

	Precision	Recall	F1-score	Support
0	0.10	0.09	0.10	870
1	0.89	0.90	0.90	7393
Accuracy			0.82	8263
Micro Avg	0.50	0.50	0.50	8263
Weighted Avg	0.81	0.82	0.81	8263

- **Loss Vs Accuracy Graphs (Inception_v3):**This below graph is showing that Loss vs accuracy in Inception_v3 model.

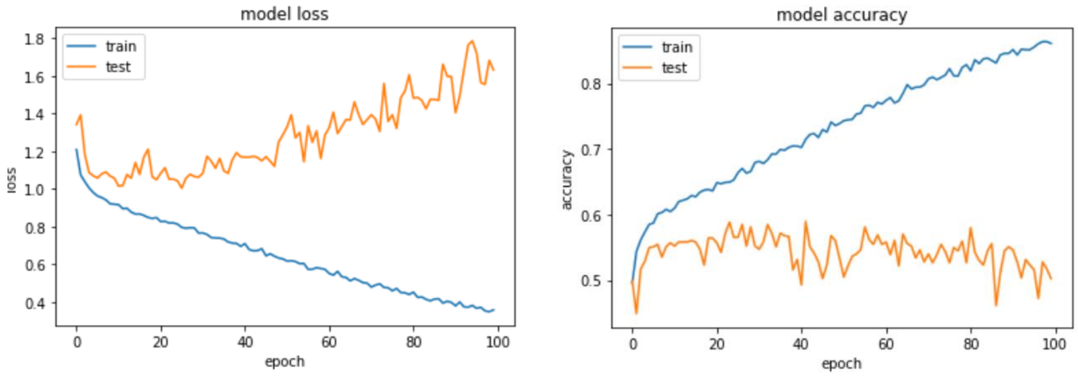


Figure 31: graphs of model loss vs accuracy of inception_v3 model

It can be inferred from the graph

- **Model accuracy**

❖ The blue line shows the training loss whereas the orange line shows the testing / validation loss.

❖ The x-axis shows the no of epochs whereas the y-axis shows the model loss

❖ It is observed that the training loss has started from 1.4 and eventually reached end with some ups and downs at 1.2.

❖ It is observed that the testing / validation loss has started from 1.2 and eventually reached end with some ups and downs at 0.6 following training loss.

❖ The interval on x-axis is 1 whereas the interval on y-axis is 0.01

- **Model Loss**

❖ The blue line shows the training accuracy whereas the orange line shows the testing / validation accuracy.

❖ The x-axis shows the no of epochs whereas the y-axis shows the model accuracy

❖ It is observed that the training accuracy has started from 0.510 and eventually reached end with some ups and downs at 0.78.

- ❖ It is observed that the testing / validation accuracy has started from 0.45 and eventually reached end with some ups and downs at 0.53 following training accuracy.
- ❖ The interval on x-axis is 1 whereas the interval on y-axis is 0.01.

5.2 Model Evaluation(Xception):

In this step, The Five classes are part of detection. Additionally, the knee arthritis appropriate ratio. Figure 1 illustrates how well CNN's recommended model performed during training and validation using knee x-ray data. Performance was calculated based on accuracy at different iterations and epochs using each validation's monitored epoch. The training's validation accuracy and accuracy are shown in the table.

Model Name	**Dataset**	**Epochs**	**Validation Accuracy**	**Accuracy**
Xception	Collection of Knee X-Ray	1	44%	41%
		25	48%	51%
		50	47%	53%
		100	50%	56%

Table 6: Epochs wise validation accuracy of Xception model

we have evaluated our model on the training and testing sets. We are achieving an accuracy of 51.3 % on the training dataset and 50.4 % on the testing dataset. The test loss is 1.19 and the test accuracy is 0.49 .

- **Training loss Vs Validation loss graph (Xception):** This below graph is showing that Loss vs validation accuracy in Xception model.

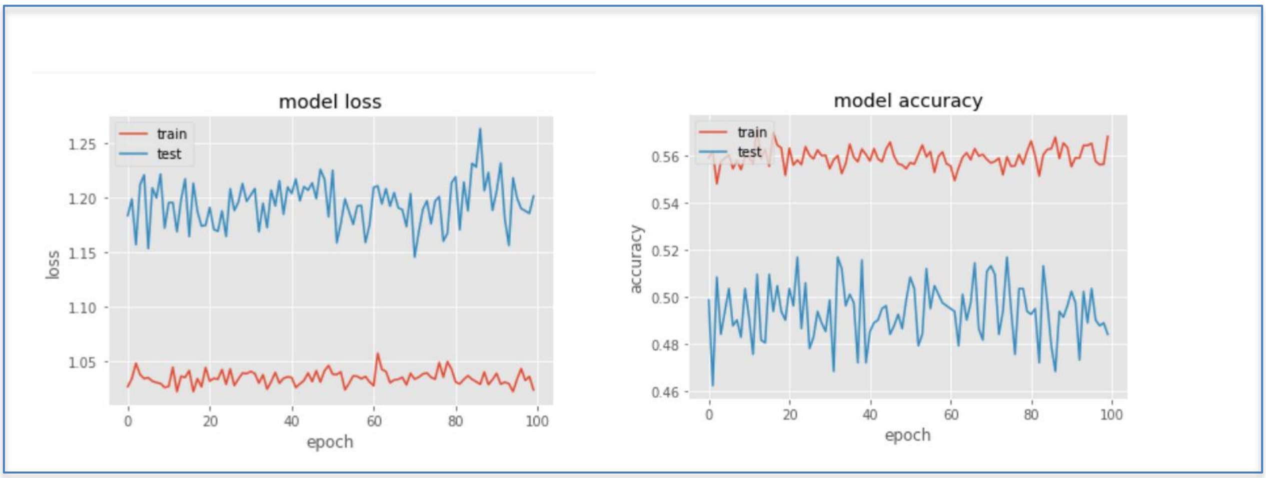


Figure 32: Graphs of model loss and accuracy of Xception model

It can be inferred from the graph:

- The blue line shows the training loss whereas the orange line shows the testing / validation loss.
- The x-axis shows the no of epochs whereas the y-axis shows the model loss
- It is observed that the training loss has started from 1.161 and eventually reached end with some ups and downs at 1.149.
- It is observed that the testing / validation loss has started from 1.188 and eventually reached end with some ups and downs at 1.161 following training loss.
- The interval on x-axis is 1 whereas the interval on y-axis is 0.01.

Training accuracy Vs Validation accuracy (Xception):

- The blue line shows the training accuracy whereas the orange line shows the testing / validation accuracy.
- The x-axis shows the no of epochs whereas the y-axis shows the model accuracy
- It is observed that the training accuracy has started from 0.510 and eventually reached end with some ups and downs at 0.514.

- It is observed that the testing / validation accuracy has started from 0.494 and eventually reached end with some ups and downs at 0.517 following training accuracy.
- The interval on x-axis is 1 whereas the interval on y-axis is 0.01.

Confusion matric and classification report (Xception):

The confusion matrix for Xception model is shown below:

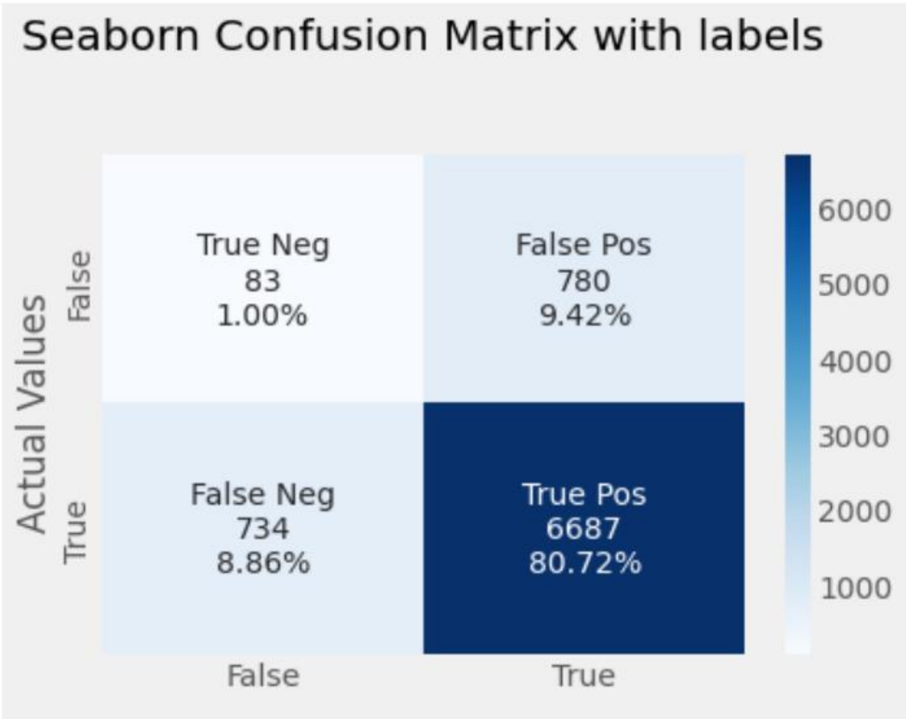


Figure 33: Confusion matrix of Xception model

We will analyse the results using conventional measurement to judge the effectiveness of the suggested CNN model's detecting capabilities.The first measurement we compute is an accuracy of a technique, which establishes how accurately the values are predicted.

accuracy is defined as TP+TN TP+FP+TN+FN............. where TP, FP, and FN stand for really positive, truly negative, false positive, and false negative, respectively. The confusion matrix was calculated as well. The confusion matrix of the Knee x-ray dataset is displayed in Figure . resulting in a 80.72% True positive rate. the second class of Normal had a false positive rate of 9.42%, while the third had false nagative rate of 8.86% and last had true nagative rate of 1%.

In table 00, classification report of Xception model explains precision, Recall, F1 Score, support and accuracy:

Classification Report of Xception model:

Table 7: classification report of Xception model

	Precision	Recall	F1-score	Support
0	0.10	0.10	0.10	863
1	0.90	0.90	0.90	7421
Accuracy			0.82	8284
Micro Avg	0.50	0.50	0.50	8284
Weighted Avg	0.81	0.82	0.82	8284

Training and Validation Loss VS Accuracy: This is below graph is showing that training and validation loss and accuracy coparision**.**

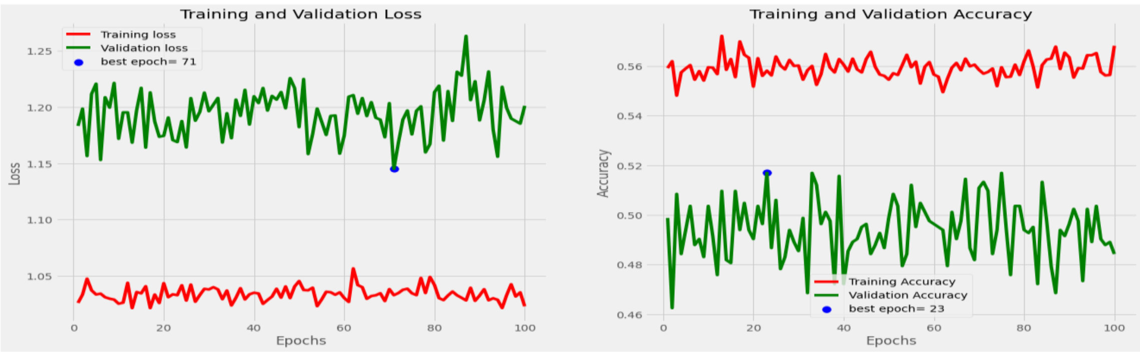


Figure 34: Graphs of training and valodation loss and accuracy of xception model

It can be inferred from the loss graph:

- The red line shows the training loss whereas the green line shows the validation / test loss. The blue dot shows the best epoch.
- The x-axis shows the no of epochs whereas the y-axis shows the model loss.
- It is observed that the training loss has started from 0.5 and eventually reached end with some minor ups and downs at 0.61.
- It is observed that the validation loss has started from 1.18 and eventually reached end with some major ups and downs at 1.20 following training loss. The best epoch is 71.
- The interval on x-axis is 10 whereas the interval on y-axis is 0.1.

It can be inferred from the accuracy graph:

- The x-axis shows the no of epochs whereas the y-axis shows the model accuracies.
- It is observed that the training accuracy has started from 0.50 and eventually reached end with some minor ups and downs at 0.75.
- It is observed that the validation / testing accuracy has started from 0.52 and eventually reached end with some major ups and downs at 0.57 following training accuracy. The best epoch is 23.
- The interval on x-axis is 10 whereas the interval on y-axis is 0.1.
- The red lines show the training accuracy whereas the green line shows the validation / test accuracy. The blue dot shows the best epoch.

Comparison of Accuracy Using Different Methods Summary:

Table 8: comparison of model

Model	**Dataset**	**Accuracy**
Inceptio_v3	X-ray Images	86%
Exception	X-ray Images	56%

From above tables and figures, it can be shown that, for the Knee x-ray picture we chose from the "dataset," the suggested model outperformed pre-trained Inception_v3 and Xception in terms of accuracy. The Inception_v3 model comparatively given best Result with 86 %.

Chapter 6 – Conclusions & Future work

6.1 Conclusion

The detection of chronic diseases, such as knee OA, on time is crucial and also to obtain useful information about knee osteoarthritis from raw healthcare data. In the long run, it has the potential to save lives. The identification & prevention of such diseases was difficult before the development of technology, but technological improvements have made it much simpler to diagnose and prevent such diseases. Artificial Intelligence (AI) and Deep Learning (DL) have demonstrated outstanding benchmark performance in the field of biology in recent years. The goal of this study is to develop an intelligent classification model that uses deep learning techniques such as CNNs to identify the security of knee osteoarthritis, so that this disease may be controlled in time to save patients' lives. Keeping in context the nature of training data (radiology images), convolutional neural networks (CNNs) powered by transfer learning (TL) will be employed for image classification.

6.2 Limitation

Since It is having large number of image data set, it requires high configuration system which is limitation of this model. If you don't have enough resources to run the code then it can be crashed. It requires in depth domain knowledge to execute the project. I ran into execution issue since there were limited cloud resource like memory, CPU. Since Deep learning is very waste topic, I couldn't achieve as per my expectations because of lack of time and resources.

6.3 Future work

By changing the hyperparameters, the Inception_v3 and Xception performance may be significantly improved, outperforming competing models. Given its effective use of computing resources and recent advancements in embedded hardware processing units, it can have a wide range of applications. I trying to do GUI for this project so In future I will do that as well .

7. Reference list / Bibliography

1. Hsu, H. and Siwiec, R.M., 2018. Knee osteoarthritis
2. Lawson, T., Morrison, A., Blaxland, S., Wenman, M., Schmidt, C.G. and Hunt, M.A., 2015. Laboratory-based measurement of standing balance in individuals with knee osteoarthritis: A systematic review. Clinical Biomechanics, 30(4), pp.330-342.
3. Hatfield, G.L., Morrison, A., Wenman, M., Hammond, C.A. and Hunt, M.A., 2016. Clinical tests of standing balance in the knee osteoarthritis population: systematic review and meta-analysis. Physical therapy, 96(3), pp.324-337.
4. Liu, C., Wan, Q., Zhou, W., Feng, X. and Shang, S., 2017. Factors associated with balance function in patients with knee osteoarthritis: An integrative review. *International journal of nursing sciences*, *4*(4), pp.402-409.
5. Toivanen, A.T., Heliövaara, M., Impivaara, O., Arokoski, J.P., Knekt, P., Lauren, H. and Kröger, H., 2010. Obesity, physically demanding work and traumatic knee injury are major risk factors for knee osteoarthritis—a population-based study with a follow-up of 22 years. *Rheumatology*, *49*(2), pp.308-314.
6. Gupta, S., Hawker, G.A., Laporte, A., Croxford, R. and Coyte, P.C., 2005. The economic burden of disabling hip and knee osteoarthritis (OA) from the perspective of individuals living with this condition. *Rheumatology*, *44*(12), pp.1531-1537.
7. "Computer-aided Diagnosis: The Tipping Point for Digital Pathology". Digital Pathology Association. 27 April 2017.
8. Silverwood, V., Blagojevic-Bucknall, M., Jinks, C., Jordan, J.L., Protheroe, J. and Jordan, K.P., 2015. Current evidence on risk factors for knee osteoarthritis in older adults: a systematic review and meta-analysis. *Osteoarthritis and cartilage*, *23*(4), pp.507-515.
9. Ackerman, I.N., Kemp, J.L., Crossley, K.M., Culvenor, A.G. and Hinman, R.S., 2017. Hip and knee osteoarthritis affects younger people, too. *journal of orthopaedic & sports physical therapy*, *47*(2), pp.67-79.
10. Antony, B., Jones, G., Jin, X. and Ding, C., 2016. Do early life factors affect the development of knee osteoarthritis in later life: a narrative review. *Arthritis research & therapy*, *18*(1), pp.1-8.
11. Lespasio, M.J., Piuzzi, N.S., Husni, M.E., Muschler, G.F., Guarino, A.J. and Mont, M.A., 2017. Knee osteoarthritis: a primer. *The Permanente Journal*, *21*.
12. Farr II, J., Miller, L.E. and Block, J.E., 2013. Quality of life in patients with knee osteoarthritis: a commentary on nonsurgical and surgical treatments. *The open orthopaedics journal*, *7*, p.619.

13. Ntakolia, C., Kokkotis, C., Moustakidis, S. and Tsaopoulos, D., 2020, October. A machine learning pipeline for predicting joint space narrowing in knee osteoarthritis patients. In *2020 IEEE 20th International Conference on Bioinformatics and Bioengineering (BIBE)* (pp. 934-941). IEEE.
14. Moustakidis, S., Papandrianos, N.I., Christodolou, E., Papageorgiou, E. and Tsaopoulos, D., 2020. Dense neural networks in knee osteoarthritis classification: a study on accuracy and fairness. *Neural Computing and Applications*, pp.1-13.
15. Moustakidis, S., Christodoulou, E., Papageorgiou, E., Kokkotis, C., Papandrianos, N. and Tsaopoulos, D., 2019. Application of machine intelligence for osteoarthritis classification: a classical implementation and a quantum perspective. *Quantum Machine Intelligence*, *1*(3), pp.73-86.
16. Kokkotis, C., Moustakidis, S., Giakas, G. and Tsaopoulos, D., 2020. Identification of risk factors and machine learning-based prediction models for knee osteoarthritis patients. *Applied Sciences*, *10*(19), p.6797.
17. Alexos, A., Kokkotis, C., Moustakidis, S., Papageorgiou, E. and Tsaopoulos, D., 2020, July. Prediction of pain in knee osteoarthritis patients using machine learning: Data from Osteoarthritis Initiative. In *2020 11th International Conference on Information, Intelligence, Systems and Applications (IISA* (pp. 1-7). IEEE.
18. Alexos, A., Moustakidis, S., Kokkotis, C. and Tsaopoulos, D., 2020, May. Physical activity as a risk factor in the progression of osteoarthritis: a machine learning perspective. In *International Conference on Learning and Intelligent Optimization* (pp. 16-26). Springer, Cham.
19. Kokkotis, C., Moustakidis, S., Papageorgiou, E., Giakas, G. and Tsaopoulos, D.E., 2020. Machine learning in knee osteoarthritis: a review. *Osteoarthritis and Cartilage Open*, *2*(3), p.100069.
20. Jamshidi, A., Pelletier, J.P. and Martel-Pelletier, J., 2019. Machine-learning-based patient-specific prediction models for knee osteoarthritis. *Nature Reviews Rheumatology*, *15*(1), pp.49-60.
21. Conaghan, P.G., Porcheret, M., Kingsbury, S.R., Gammon, A., Soni, A., Hurley, M., Rayman, M.P., Barlow, J., Hull, R.G., Cumming, J. and Llewelyn, K., 2015. Impact and therapy of osteoarthritis: the Arthritis Care OA Nation 2012 survey. *Clinical rheumatology*, *34*(9), pp.1581-1588.
22. Primorac, D., Molnar, V., Rod, E., Jeleč, Ž., Čukelj, F., Matišić, V., Vrdoljak, T., Hudetz, D., Hajsok, H. and Borić, I., 2020. Knee osteoarthritis: a review of pathogenesis and state-of-the-art non-operative therapeutic considerations. *Genes*, *11*(8), p.854.

23. Janvier, T., Jennane, R., Toumi, H. and Lespessailles, E., 2017. Subchondral tibial bone texture predicts the incidence of radiographic knee osteoarthritis: data from the osteoarthritis initiative. *Osteoarthritis and cartilage*, *25*(12), pp.2047-2054.
24. Shamir, L., Ling, S.M., Scott, W., Hochberg, M., Ferrucci, L. and Goldberg, I.G., 2009. Early detection of radiographic knee osteoarthritis using computer-aided analysis. *Osteoarthritis and Cartilage*, *17*(10), pp.1307-1312.
25. Antony, J., McGuinness, K., O'Connor, N.E. and Moran, K., 2016, December. Quantifying radiographic knee osteoarthritis severity using deep convolutional neural networks. In *2016 23rd International Conference on Pattern Recognition (ICPR)* (pp. 1195-1200). IEEE.
26. Guermazi, A., Burstein, D., Conaghan, P., Eckstein, F., Le Graverand-Gastineau, M.P.H., Keen, H. and Roemer, F.W., 2008. Imaging in osteoarthritis. *Rheumatic Disease Clinics of North America*, *34*(3), pp.645-687.
27. Guermazi, A., Hayashi, D., Roemer, F.W. and Felson, D.T., 2013. Osteoarthritis: a review of strengths and weaknesses of different imaging options. *Rheumatic Disease Clinics*, *39*(3), pp.567-591.
28. Zhang, W., Doherty, M., Peat, G., Bierma-Zeinstra, M.A., Arden, N.K., Bresnihan, B., Herrero-Beaumont, G., Kirschner, S., Leeb, B.F., Lohmander, L.S. and Mazières, B., 2010. EULAR evidence-based recommendations for the diagnosis of knee osteoarthritis. *Annals of the rheumatic diseases*, *69*(3), pp.483-489.
29. Kwon, S.B., Ku, Y., Han, H.S., Lee, M.C., Kim, H.C. and Ro, D.H., 2020. A machine learning-based diagnostic model associated with knee osteoarthritis severity. *Scientific reports*, *10*(1), pp.1-8.
30. Antony J, et al (2017) Automatic detection of knee joints and quantification of knee osteoarthritis severity using convolutional neural networks, in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp 376–390.
31. Tiulpin A et al (2018) Automatic knee osteoarthritis diagnosis from plain radiographs: a deep learning-based approach. Sci Rep 8(1):1727
32. Pedoia V, Lee J, Norman B et al (2019) Diagnosing osteoarthritis from T2 maps using deep learning: an analysis of the entire Osteoarthritis Initiative baseline cohort. Osteoarthr Cartil 27:1002–1010. https://doi.org/10.1016/j.joca.2019.02.800
33. Schwartz, A.J., Clarke, H.D., Spangehl, M.J., Bingham, J.S., Etzioni, D.A. and Neville, M.R., 2020. Can a convolutional neural network classify knee osteoarthritis on plain radiographs as accurately as fellowship-trained knee arthroplasty surgeons? *The Journal of Arthroplasty*, *35*(9), pp.2423-2428.

34. Swiecicki, A., Li, N., O'Donnell, J., Said, N., Yang, J., Mather, R.C., Jiranek, W.A. and Mazurowski, M.A., 2021. Deep learning-based algorithm for assessment of knee osteoarthritis severity in radiographs matches performance of radiologists. *Computers in biology and medicine*, *133*, p.104334.
35. Norman, B., Pedoia, V., Noworolski, A., Link, T.M. and Majumdar, S., 2019. Applying densely connected convolutional neural networks for staging osteoarthritis severity from plain radiographs. *Journal of digital imaging*, *32*(3), pp.471-477.
36. Kim, D.H., Lee, K.J., Choi, D., Lee, J.I., Choi, H.G. and Lee, Y.S., 2020. Can additional patient information improve the diagnostic performance of deep learning for the interpretation of knee osteoarthritis severity. *Journal of Clinical Medicine*, *9*(10), p.3341.
37. Tiulpin, A. and Saarakkala, S., 2020. Automatic grading of individual knee osteoarthritis features in plain radiographs using deep convolutional neural networks. *Diagnostics*, *10*(11), p.932.
38. Ahmed, S.M. and Mstafa, R.J., 2022. A Comprehensive Survey on Bone Segmentation Techniques in Knee Osteoarthritis Research: From Conventional Methods to Deep Learning. *Diagnostics*, *12*(3), p.611.
39. Guan, B., Liu, F., Mizaian, A.H., Demehri, S., Samsonov, A., Guermazi, A. and Kijowski, R., 2022. Deep learning approach to predict pain progression in knee osteoarthritis. *Skeletal Radiology*, *51*(2), pp.363-373.
40. Yong, C.W., Teo, K., Murphy, B.P., Hum, Y.C., Tee, Y.K., Xia, K. and Lai, K.W., 2021. Knee osteoarthritis severity classification with ordinal regression module. *Multimedia Tools and Applications*, pp.1-13.
41. Chen, P., Gao, L., Shi, X., Allen, K. and Yang, L., 2019. Fully automatic knee osteoarthritis severity grading using deep neural networks with a novel ordinal loss. *Computerized Medical Imaging and Graphics*, *75*, pp.84-92.
42. Thomas, K.A., Kidziński, Ł., Halilaj, E., Fleming, S.L., Venkataraman, G.R., Oei, E.H., Gold, G.E. and Delp, S.L., 2020. Automated classification of radiographic knee osteoarthritis severity using deep neural networks. *Radiology. Artificial intelligence*, *2*(2).
43. Jain, R.K., Sharma, P.K., Gaj, S., Sur, A. and Ghosh, P., 2021. Knee osteoarthritis severity prediction using an attentive multi-scale deep convolutional neural network. *arXiv preprint arXiv:2106.14292*.
44. Yeoh, P.S.Q., Lai, K.W., Goh, S.L., Hasikin, K., Hum, Y.C., Tee, Y.K. and Dhanalakshmi, S., 2021. Emergence of deep learning in knee osteoarthritis diagnosis. *Computational intelligence and neuroscience*, *2021*.

45. Cheung, J.C.W., Tam, A.Y.C., Chan, L.C., Chan, P.K. and Wen, C., 2021. Superiority of multiple-joint space width over minimum-joint space width approach in the machine learning for radiographic severity and knee osteoarthritis progression. *Biology*, *10*(11), p.1107.
46. Brahim, A., Jennane, R., Riad, R., Janvier, T., Khedher, L., Toumi, H. and Lespessailles, E., 2019. A decision support tool for early detection of knee OsteoArthritis using X-ray imaging and machine learning: Data from the OsteoArthritis Initiative. *Computerized Medical Imaging and Graphics*, *73*, pp.11-18.
47. Christodoulou, E., Moustakidis, S., Papandrianos, N., Tsaopoulos, D. and Papageorgiou, E., 2019, July. Exploring deep learning capabilities in knee osteoarthritis case study for classification. In *2019 10th International Conference on Information, Intelligence, Systems and Applications (IISA)* (pp. 1-6). IEEE.
48. Lim, J., Kim, J. and Cheon, S., 2019. A deep neural network-based method for early detection of osteoarthritis using statistical data. *International journal of environmental research and public health*, *16*(7), p.1281.
49. Abdullah, S.S. and Rajasekaran, M.P., 2022. Automatic detection and classification of knee osteoarthritis using deep learning approach. *La radiologia medica*, *127*(4), pp.398-406.
50. Hall, M., Bennell, K.L., Wrigley, T.V., Metcalf, B.R., Campbell, P.K., Kasza, J., Paterson, K.L., Hunter, D.J. and Hinman, R.S., 2017. The knee adduction moment and knee osteoarthritis symptoms: relationships according to radiographic disease severity. *Osteoarthritis and cartilage*, *25*(1), pp.34-41.
51. Saygili, A. and Albayrak, S., 2020. Knee meniscus segmentation and tear detection from mri: A review. *Current Medical Imaging*, *16*(1), pp.2-15.
52. Ahn, C., Bui, T.D., Lee, Y.W., Shin, J. and Park, H., 2016. Fully automated, level set-based segmentation for knee MRIs using an adaptive force function and template: data from the osteoarthritis initiative. *Biomedical engineering online*, *15*(1), pp.1-14.
53. Nguyen, H.H., Saarakkala, S., Blaschko, M.B. and Tiulpin, A., 2020. Semixup: in-and out-of-manifold regularization for deep semi-supervised knee osteoarthritis severity grading from plain radiographs. *IEEE Transactions on Medical Imaging*, *39*(12), pp.4346-4356.
54. Wahyuningrum, R.T., Anifah, L., Purnama, I.K.E. and Purnomo, M.H., 2019, October. A new approach to classify knee osteoarthritis severity from radiographic images based on CNN-LSTM method. In *2019 IEEE 10th International Conference on Awareness Science and Technology (iCAST)* (pp. 1-6). IEEE.
55. Tolpadi, A.A., Lee, J.J., Pedoia, V. and Majumdar, S., 2020. Deep learning predicts total knee replacement from magnetic resonance images. *Scientific reports*, *10*(1), pp.1-12.
56. von Schacky, C.E., Sohn, J.H., Liu, F., Ozhinsky, E., Jungmann, P.M., Nardo, L., Posadzy, M., Foreman, S.C., Nevitt, M.C., Link, T.M. and Pedoia, V., 2020. Development and

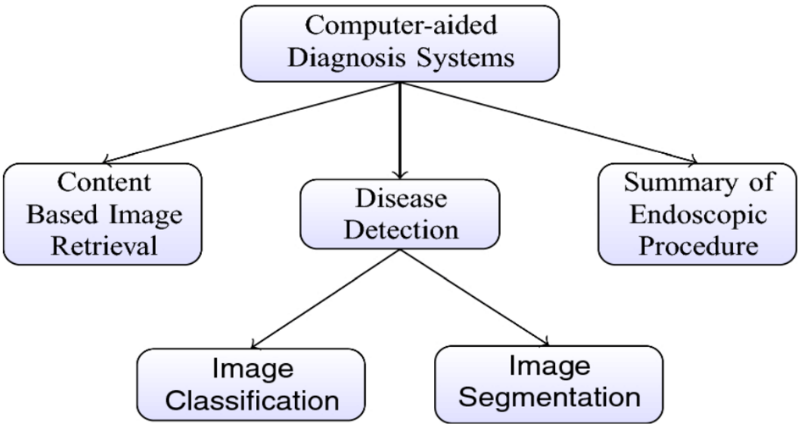
validation of a multitask deep learning model for severity grading of hip osteoarthritis features on radiographs. *Radiology*, *295*(1), p.136.

57. Liu, B., Luo, J. and Huang, H., 2020. Toward automatic quantification of knee osteoarthritis severity using improved Faster R-CNN. *International journal of computer assisted radiology and surgery*, *15*(3), pp.457-466.
58. Olsson, S., Akbarian, E., Lind, A., Razavian, A.S. and Gordon, M., 2021. Automating classification of osteoarthritis according to Kellgren-Lawrence in the knee using deep learning in an unfiltered adult population. *BMC Musculoskeletal Disorders*, *22*(1), pp.1-8.
59. Raju, S.S., Niranjan, T., Pandiyan, P. and Snehitha, M.S., 2021, February. A review of an early detection and quantification of osteoarthritis severity in knee using machine learning techniques. In *IOP Conference Series: Materials Science and Engineering* (Vol. 1057, No. 1, p. 012095). IOP Publishing.
60. Schutte, K., Moindrot, O., Hérent, P., Schiratti, J.B. and Jégou, S., 2021. Using stylegan for visual interpretability of deep learning models on medical images. *arXiv preprint arXiv:2101.07563*.
61. Saini, D., Chand, T., Chouhan, D.K. and Prakash, M., 2021. A comparative analysis of automatic classification and grading methods for knee osteoarthritis focusing on X-Ray images. *Biocybernetics and Biomedical Engineering*, *41*(2), pp.419-444.
62. Mahum, R., Rehman, S.U., Meraj, T., Rauf, H.T., Irtaza, A., El-Sherbeeny, A.M. and El-Meligy, M.A., 2021. A novel hybrid approach based on deep cnn features to detect knee osteoarthritis. *Sensors*, *21*(18), p.6189.
63. Suresha, S., Kidziński, L., Halilaj, E., Gold, G.E. and Delp, S.L., 2018. Automated staging of knee osteoarthritis severity using deep neural networks. *Osteoarthritis and Cartilage*, *26*, p.S441.
64. Karsdal, M.A., Michaelis, M., Ladel, C., Siebuhr, A.S., Bihlet, A.R., Andersen, J.R., Guehring, H., Christiansen, C., Bay-Jensen, A.C. and Kraus, V.B., 2016. Disease-modifying treatments for osteoarthritis (DMOADs) of the knee and hip: lessons learned from failures and opportunities for the future. *Osteoarthritis and cartilage*, *24*(12), pp.2013-2021.
65. Saleem, M., Farid, M.S., Saleem, S. and Khan, M.H., 2020. X-ray image analysis for automated knee osteoarthritis detection. Signal, Image and Video Processing, 14(6), pp.1079-1087.
66. https://www.kaggle.com/datasets/tommyngx/kneeoa
67. https://iq.opengen from google
68. https://us.org/Xception-model/ from google
69. Muntasir Al-Afsoor, Deep Learning Approach for COVID-19 Diagnosis Using X-Ray Images, May 26th, 2020, https://doi.org/10.21203/rs.3.rs-31278/v1

70. Woolf AD. The bone and joint decade. strategies to reduce the burden of disease: the Bone and Joint Monitor Project. J Rheumatol Suppl. 2003;67:6–9. [PubMed] [Google Scholar]
71. Deshpande BR, Katz JN, Solomon DH, Yelin EH, Hunter DJ, Messier SP, et al. Number of persons with symptomatic knee osteoarthritis in the US: impact of race and ethnicity, age, sex, and obesity. Arthritis Care Res (Hoboken) 2016;68:1743–1750. doi: 10.1002/acr.22897. [PMC free article] [PubMed] [CrossRef] [Google Scholar]
72. Vina ER, Kwoh CK. Epidemiology of osteoarthritis: literature update. Curr OpinRheumatol. 2018;30:160167.doi: 10.1097/BOR.0000000000000479. [PMC free article] [PubMed] [CrossRef] [Google Scholar]
73. Hermans J, Koopmanschap MA, Bierma-Zeinstra SMA, van Linge JH, Verhaar JAN, Reijman M, et al. Productivity costs and medical costs among working patients with knee osteoarthritis. Arthritis Care Res
74. (Hoboken) 2012;64:853–861.doi: 10.1002/acr.21617. [PubMed] [CrossRef] [Google Scholar]

8. Appendices

Appendix A: Types of COD



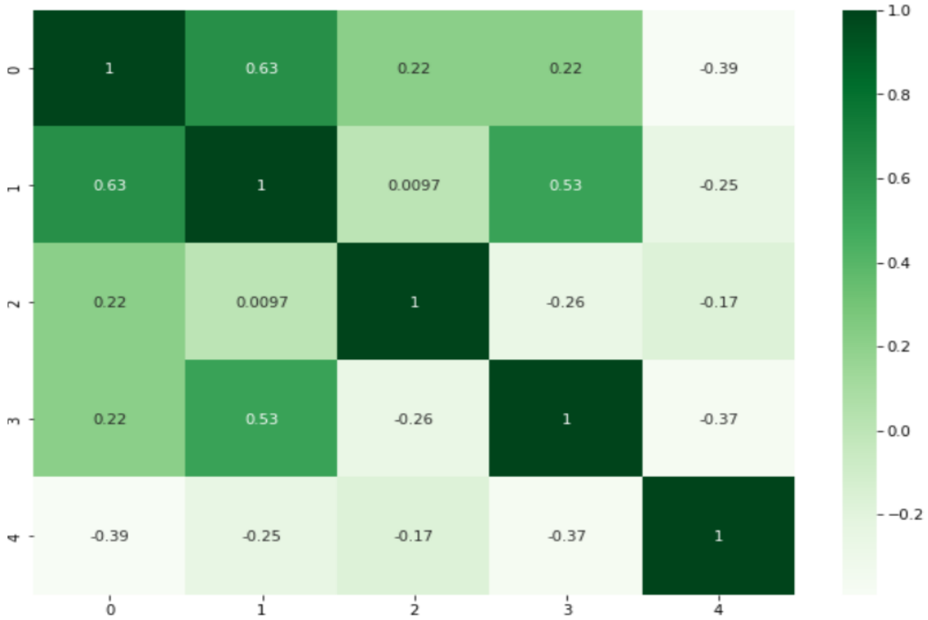
Appendix B : correlation matrix

```
#check data correlation
import pandas as pd
import numpy as np

rs = np.random.RandomState(0)
df = pd.DataFrame(rs.rand(5, 5))
corr = df.corr()
corr.style.background_gradient(cmap='coolwarm')
```

	0	1	2	3	4
0	1.000000	0.861882	-0.687578	0.220408	-0.965567
1	0.861882	1.000000	-0.770134	-0.283359	-0.807373
2	-0.687578	-0.770134	1.000000	0.341363	0.701713
3	0.220408	-0.283359	0.341363	1.000000	-0.262302
4	-0.965567	-0.807373	0.701713	-0.262302	1.000000

Appendix C : Heatmap



Appendix D : Training loss and accuracy of inception_v3

```python
from matplotlib import pyplot as plt
# plot the training loss and accuracy
N = 100 #number of epochs
plt.style.use("ggplot")
plt.figure()
plt.plot(np.arange(0, N), history.history["loss"], label="train_loss")
plt.plot(np.arange(0, N), history.history["val_loss"], label="val_loss")
plt.plot(np.arange(0, N), history.history["accuracy"], label="train_acc")
plt.plot(np.arange(0, N), history.history["val_accuracy"], label="val_acc")
plt.title("Training Loss and Accuracy")
plt.xlabel("Epoch #")
plt.ylabel("Loss/Accuracy")
plt.legend(loc="center right")
plt.savefig("Inception_v3_Model")
```

Appendix E: Training loss and accuracy of Xception model

```
from matplotlib import pyplot as plt
# plot the training loss and accuracy
N = 100 #number of epochs
plt.style.use("ggplot")
plt.figure()
plt.plot(np.arange(0, N), history.history["loss"], label="train_loss")
plt.plot(np.arange(0, N), history.history["val_loss"], label="val_loss")
plt.plot(np.arange(0, N), history.history["accuracy"], label="train_acc")
plt.plot(np.arange(0, N), history.history["val_accuracy"], label="val_acc")
plt.title("Training Loss and Accuracy")
plt.xlabel("Epoch #")
plt.ylabel("Loss/Accuracy")
plt.legend(loc="center right")
plt.savefig("Xceotion_Model")
```

Appendix F: Importing necessary library

```
import numpy as np
import pandas as pd
import os
import random
import matplotlib.pyplot as plt
import matplotlib.image as mpimg
from collections import Counter
import cv2
import glob
import tensorflow as tf


from tensorflow.keras.applications.inception_v3 import InceptionV3, preprocess_input


from tensorflow.keras.preprocessing.image import ImageDataGenerator
from tensorflow.keras.callbacks import EarlyStopping, ReduceLROnPlateau
from tensorflow.keras.optimizers import Adam, SGD

from tensorflow.keras.models import Model
from tensorflow.keras.layers import Dense, Flatten, Dropout, Input
```

Appendix G: Giving path code

```
base_dir = "/content/drive/My Drive/Kneeoa"

train_dir = os.path.join(base_dir,'train')
val_dir = os.path.join(base_dir,'val')
test_dir = os.path.join(base_dir, 'test')
```

Appendix H: Plotting some sample images from each of the classes

```
fig, ax = plt.subplots(5,5, figsize=(18,18))

for class_id in range(5):
    folder = os.path.join(train_dir,str(class_id))
    os.chdir(folder)
    samples = random.sample(os.listdir(folder), 5)

    if class_id == 0:
      mystr1 = 'minimal'
    elif class_id == 1:
      mystr1 = 'healthy'
    elif class_id == 2:
      mystr1 = 'moderate'
    elif class_id == 3:
      mystr1 = 'doubtful'
    else :
      mystr1 = 'severe'
    for col in range(5):
        image = cv2.imread(samples[col])
        ax[class_id, col].imshow(image)
        ax[class_id, col].set_title("class_" + str(class_id)+mystr1)
        ax[class_id, col].set_axis_off()
```

Appendix I: Disbalance class to balance the class code

```
train_datagen = ImageDataGenerator()
train_generator = train_datagen.flow_from_directory(train_dir)
counter = Counter(train_generator.classes)
max_val = float(max(counter.values()))
class_weights = {class_id : max_val/num_images for class_id, num_images in counter.items()}
print("\nThe class weights are : \n\n", class_weights)
```

```
Found 5778 images belonging to 5 classes.

The class weights are :

 {0: 1.0, 1: 2.18546845124283, 2: 1.507915567282322, 3: 3.0198150594451785, 4: 13.213872832369942}
```

Appendix J : Image size distribution

```
folder = os.path.join(train_dir,'0')
os.chdir(folder)
samples = random.sample(os.listdir(folder), 5)

for filename in samples:
    image = cv2.imread(filename)
    print(image.shape)
```

```
(224, 224, 3)
(224, 224, 3)
(224, 224, 3)
(224, 224, 3)
(224, 224, 3)
```

Appendix K: Data pre-processing code

```
NUM_CLASSES = 5
IMAGE_SIZE=[224, 224]
BATCH_SIZE=64
```

```
# Create the Generators
train_val_generator = ImageDataGenerator(
                                        preprocessing_function=preprocess_input,
                                        rotation_range=10,
                                        width_shift_range=0.2,
                                        height_shift_range=0.2,
                                        shear_range=0.1,
                                        zoom_range=0.2,
                                        horizontal_flip=True,
                                        vertical_flip=False,
                                        fill_mode='nearest'
                                        )


# Train data generator
train_data = train_val_generator.flow_from_directory(train_dir,
                                                     target_size=IMAGE_SIZE,
                                                     batch_size=BATCH_SIZE,
                                                     class_mode='categorical',
                                                     color_mode='rgb',
                                                     shuffle=True
                                                     )

# Validation data generator
val_data = train_val_generator.flow_from_directory(val_dir,
                                                   target_size=IMAGE_SIZE,
                                                   batch_size=BATCH_SIZE,
                                                   class_mode='categorical',
```

Appendix L: Reducer and early stopping code

```
reduce_lr = ReduceLROnPlateau(monitor='val_loss', factor=0.2,
                              patience=5, min_lr=0.001)
early_stopping = EarlyStopping(monitor = 'loss', patience = 5)
```

Appendix M: Inception_v3 model code

```
rom tensorflow.keras.applications.inception_v3 import InceptionV3
rom tensorflow.keras.preprocessing import image
rom tensorflow.keras.models import Model
rom tensorflow.keras.layers import Dense, GlobalAveragePooling2D
rom tensorflow.keras.optimizers import Adam

 create the base pre-trained model
ase_model = InceptionV3(weights='imagenet', include_top=False)

 add a global spatial average pooling layer
 = base_model.output
 = GlobalAveragePooling2D()(x)
 let's add a fully-connected layer
 = Dense(1024, activation='relu')(x)
 and a logistic layer -- let's say we have 5 classes
redictions = Dense(5, activation='softmax')(x)

 this is the model we will train
odel = Model(inputs=base_model.input, outputs=predictions)

 first: train only the top layers (which were randomly initialized)
 i.e. freeze all convolutional InceptionV3 layers
or layer in base_model.layers:
   layer.trainable = False

 compile the model (should be done *after* setting layers to non-trainable)
odel.compile(optimizer='adam', loss='categorical_crossentropy', metrics = ['accuracy'])

 train the model on the new data for a few epochs
istory = model.fit(train_data,
        validation_data = val_data,
         epochs = 100,
         callbacks = [reduce_lr, early_stopping]
         )

 at this point, the top layers are well trained and we can start fine-tuning
 convolutional layers from inception V3. We will freeze the bottom N layers
 and train the remaining top layers.

 let's visualize layer names and layer indices to see how many layers
 we should freeze:
or i, layer in enumerate(base_model.layers):
  print(i, layer.name)

 we chose to train the top 2 inception blocks, i.e. we will freeze
 the first 249 layers and unfreeze the rest:
or layer in model.layers[:249]:
  layer.trainable = False
or layer in model.layers[249:]:
  layer.trainable = True
```

Appendix N: Xception model code

```
from tensorflow.keras.applications.xception import Xception
from tensorflow.keras.preprocessing import image
from tensorflow.keras.models import Model
from tensorflow.keras.layers import Dense, GlobalAveragePooling2D
from tensorflow.keras.optimizers import Adam

# create the base pre-trained model
ase_model = Xception(weights='imagenet', include_top=False)

# add a global spatial average pooling layer
 = base_model.output
 = GlobalAveragePooling2D()(x)
# let's add a fully-connected layer
 = Dense(1024, activation='relu')(x)
# and a logistic layer -- let's say we have 5 classes
redictions = Dense(5, activation='softmax')(x)

# this is the model we will train
odel = Model(inputs=base_model.input, outputs=predictions)

# first: train only the top layers (which were randomly initialized)
# i.e. freeze all convolutional Xception layers
for layer in base_model.layers:
   layer.trainable = False

# compile the model (should be done *after* setting layers to non-trainable)
odel.compile(optimizer='adam', loss='categorical_crossentropy', metrics = ['accuracy'])

# train the model on the new data for a few epochs
istory = model.fit(train_data,
        validation_data = val_data,
         epochs = 100,
         #callbacks = [reduce_lr, early_stopping]
         )

# at this point, the top layers are well trained and we can start fine-tuning
# convolutional layers from Xception. We will freeze the bottom N layers
# and train the remaining top layers.

# let's visualize layer names and layer indices to see how many layers
# we should freeze:
for i, layer in enumerate(base_model.layers):
  print(i, layer.name)

# we chose to train the top 2 Xception blocks, i.e. we will freeze
# the first 249 layers and unfreeze the rest:
for layer in model.layers[:249]:
  layer.trainable = False
for layer in model.layers[249:]:
  layer.trainable = True
```

Appendix O : Ethics Form

Ethics.pdf

Project status

Status

Approved

Actions

Date	Who	Action	Comments
12:40:00 05 September 2022	Femi Isiaq	Supervisor approved	
17:39:00 02 September 2022	Viralben Raval	Principal investigator submitted	