Knee Arthritis Classification Using Attention Mechanism

Chalie Lijalem Yirsaw

A Thesis Submitted to the Department of Computer Science and Engineering,

School of Electrical Engineering and Computing

Presented in Partial Fulfilment of the Requirement for the Degree of Master's in Computer Science and Engineering

Office of Graduate Studies

Adama Science and Technology University

June 2024

Adama, Ethiopia

Knee Arthritis Classification Using Attention Mechanism

Chalie Lijalem Yirsaw

Advisor: Getinet Yilma (Ph.D.)

A Thesis Submitted to the Department of Computer Science and Engineering,

School of Electrical Engineering and Computing

Presented in Partial Fulfilment of the Requirement for the Degree of Master's in Computer Science and Engineering

Office of Graduate Studies

Adama Science and Technology University

June 2024

Adama, Ethiopia

Declaration

I hereby declare that this Master Thesis entitled "**Knee Arthritis Classification Using Attention Mechanism**" is my original work. That is, it has not been submitted for the award of any academic degree, diploma, or certificate in any other university. All sources of materials that are used for this thesis have been duly acknowledged through citation.

Chalie Lijalem	____________________	__________________
Name of student	Signature	Date

Recommendation

I, the major advisor of this thesis, hereby certify that I have read the revised version of the thesis entitled "**Knee Arthritis Classification Using Attention Mechanism**" prepared under my guidance by **Chalie Lijalem Yirsaw** submitted in partial fulfillment of the requirements for the degree of Master of Science in **Computer Science and Engineering**. Therefore, I recommend the submission of revised version of the thesis to the department following the applicable procedures.

Getinet Yilma (Ph.D.)	____________________	__________________
Major Advisor	Signature	Date

Approval Sheet

I, the advisor of the thesis entitled "**Knee Arthritis Classification Using Attention Mechanism**" and developed by **Chalie Lijalem Yirsaw**, hereby certify that the recommendation and suggestions made by the board of examiners are appropriately incorporated into the final version of the thesis.

Dr. Getinet Yilma	____________________	_______________
Major Advisor	Signature	Date

We, the undersigned, members of the Board of Examiners of the thesis by **Chalie Lijalem Yirsaw** have read and evaluated the thesis entitled "**Knee Arthritis Classification Using Attention Mechanism**" and examined the candidate during open defense. This is, therefore, to certify that the thesis is accepted for partial fulfillment of the requirement of the degree of Master of Science in **Computer Science and Engineering**.

__________________________	____________________	__________________
Chairperson	Signature	Date
__________________________	____________________	__________________
Internal Examiner	Signature	Date
__________________________	____________________	__________________
External Examiner	Signature	Date

Final approval and acceptance of the thesis is contingent upon submission of its final copy to the Office of Postgraduate Studies (OPGS) through the Department Graduate Council (DGC) and School Graduate Committee (SGC).

__________________________	____________________	__________________
Department Head	Signature	Date
__________________________	____________________	__________________
School Dean	Signature	Date
__________________________	____________________	__________________
Office of Postgraduate Studies, Dean	Signature	Data

ACKNOWLEDGEMENT

I extend my heartfelt gratitude to my advisor, Dr. Getinet Yilma, for his unwavering guidance, understanding, and patience throughout this thesis journey. His positive encouragement has been instrumental in bringing this work to fruition. It has been an honor and a pleasure to have him as my mentor.

I am deeply thankful for the exceptional support from the staff members of All Girum Hospital, whose conducive environment facilitated the execution of my research and the writing of this thesis. A special acknowledgment goes to Ephrem Zeleke (CPD Coordinator) at Girum Hospital, whose invaluable assistance was indispensable in gathering data and conducting research using local datasets.

My sincere appreciation also goes to my dear friends, Mebatsion Sahle and Elias Mekuanent, who have been like siblings to me, for their unwavering support, insightful suggestions, and motivating presence throughout this research endeavor.

Furthermore, I would like to express my gratitude to my brother, Dr. Fasil Lijalem, for his invaluable contributions to this research. His expertise in medical areas and his theoretical and practical experience from Ethiopian hospitals have provided invaluable insights and support, enriching the depth and scope of this work.

Lastly, I am profoundly grateful to my family for their unwavering support and encouragement throughout my life. Their belief in me has been a guiding light, empowering me to navigate through life's challenges and opportunities.

TABLE OF CONTENTS

LISTS OF TABLES

LISTS OF FIGURES

LIST OF EQUATIONS

LIST OF ACRONYMS

ACR	American College of Rheumatology
CAD	Computer Aided Diagnosis
CBAM	Convolutional Block Attention Module
CLAHE	Contrast Limited Adaptive Histogram Equalization
CNN	Convolutional Neural Network
CORN	Consistent Ordinal Regression-based Network
CPD	Continuing Professional. Development
DCNN	Deep Convolutional Neural Network
GPU	Graphics Processing Unit
IRAIC	International Rheumatoid Arthritis Image Collaboration
KL	Kellgren-Lawrence
MSE	Mean Squared Error
OA	Osteoarthritis
OAI	Osteoarthritis Initiative
OPD	Outpatient Data
RA	Rheumatoid Arthritis
RAM	Random Access Memory
YOLO	You Only Look Once

ABSTRACT

Arthritis is the generic term for a group of inflammatory diseases that cause pain, stiffness, and swelling in the bones, muscles, and joints. Especially in the major joints like the knees, arthritis can be highly dangerous. Rheumatoid arthritis (RA), Osteoarthritis (OA), and Other types of arthritis can considerably affect people's way of daily life. Because of symptom similarities and uncertainties in diagnosis, detecting and classifying these types using X-ray images is a challenging task. There is a significant lack of knowledge regarding the detection and classification of many types of arthritis because prior research has primarily concentrated on detecting individual arthritis diseases and they faces challenges in capturing fine-grained disease features and multiscale discriminative features. A lack of qualified radiologists complicates the problem, especially in developing countries like Ethiopia, where it burdens healthcare providers and delays diagnosis. To overcome these challenges, in this study, proposed a multiscale attention deep learning approach that incorporates attention mechanisms and multiscale feature extraction to enhance arthritis detection and classification in X-ray images. With a dataset of X-ray images from an Ethiopian local hospital from 2018 to 2024, evaluate the proposed model's effectiveness in comparison to other pretrained models. The proposed model's remarkable 0.995 accuracy was attained along with metrics for precision, recall, and F1-score that were all similarly high. The results demonstrate that this approach outperforms from other models in terms of arthritis detection and classification accuracy. By incorporating attention mechanisms, this proposed method effectively captures multiscale fine-grained disease features present in X-ray images. This improvement in arthritis detection and classification can significantly contribute to diagnosis and appropriate treatment decisions. These findings highlight the model's effectiveness in correctly classifying multiple types of arthritis, addressing the urgent demand for advanced imaging techniques in areas where radiologists with the necessary training are scarce. This study eventually showed that automated detection and classification of arthritis was improved by integrating methods of attention and multi-scale feature extraction, ultimately leading to better patient outcomes through prompt and accurate diagnosis.

Keywords*: Arthritis, knee arthritis, deep learning, medical, X-ray images, Osteoarthritis (OA), Rheumatoid Arthritis (RA), attention mechanisms, multi-scale feature extraction*

CHAPTER ONE

1. INTRODUCTION

1.1. Background of the Study

The word "arthritis" refers to a group of inflammatory diseases that impact the body's joints, bones, and muscles. Our joints become painful, swollen, and stiff as a result. The strongest and largest joints in our body may be impacted. In the knees, it is common. Knee arthritis can be a serious, debilitating condition. Many types of arthritis, including psoriatic arthritis, gouty arthritis, osteoarthritis, rheumatoid arthritis, and juvenile arthritis, can cause swelling, redness, stiffness, and pain in the joints (Roques et al., 2014).

The most prevalent kind of arthritis is osteoarthritis. Our knee joint's cartilage, which serves as an extra layer between the three bones, is eroded by it. Our bones rub against one another in the absence of that shield. Pain, stiffness, and restricted movement may result from this. Additionally, it may result in the formation of bone spurs. Osteoarthritis decreases with time. Osteoarthritis affected 595 million individuals worldwide in 2020, accounting for 7.6% of the world's population and a 132.2% rise in cases since 1990. In comparison to 2020, it will be expected that by 2050, cases of osteoarthritis will rise by 74.9% in the knee, 48.6% in the hand, 78.6% in the hip, and 95.1% in different types of arthritis (Collaborators, 2023).

One type of autoimmune illness is rheumatoid arthritis. When our immune systems are functioning properly, they produce inflammation, either internal or external, to defend us against pathogens, injuries, toxins, or other outside invaders. Our body defends itself in part by inducing an inflammatory reaction. Rheumatoid arthritis is a condition in which the immune system malfunctions and causes inflammation in the joints despite the absence of an external invader. Our cartilage may also be worn down by the inflammation, which also makes the synovial membrane painful, stiff, and swollen. Rheumatoid arthritis affected 17.6 million individuals globally in 2020, according to estimates. The global age-standardized prevalence rate increased by 14.1% from 1990 to 208.8 instances per 100,000 people. According to their predictions, 31.7 million people globally would have rheumatoid arthritis by the year 2050 (Collaborators, 2023).

Computer Aided Diagnosis (CAD) technologies are becoming more and more popular in the field of medical imaging. The tasks carried out by skilled physicians are automated by these technologies when analyzing medical images. The time efficiency (a machine responds in seconds compared to days or weeks for a doctor) and repeatability of such assessments are the clear benefits. Enhancing the dependability and precision of these techniques is essential for optimizing clinical trials and offering an efficient and affordable radiography picture evaluation. Early detection of OA can help prevent further damage and lessen symptoms by advising lifestyle modifications and other preventative measures. Early diagnosis of OA can help slow down disease development and reduce symptoms by suggesting changes to people's lifestyles and adopting other precautions. In the CAD method, the physician first makes a routine evaluation of the image, re-evaluates his interpretation with the help of the CAD system, and makes the final decision. Thus, the physician receives a second objective interpretation aid. Some evidence suggests that the inclusion of the CAD system in the diagnostic process provides quantitative support for clinical decisions by reducing inter-observer variations (Singh et al., 2011).

We chose X-ray imaging for the detection and classification of arthritis disease because it is the most widely accessible and used tool for the diagnosis of knee arthritis, as it is a non-invasive method. It is comparatively inexpensive, rapid, and easy to assess imaging techniques to monitor disease progression(Üreten et al., 2020). In addition to this, X-ray imaging can reflect the variations in the structure of bones at the early stage(Liu et al., 2020).

Plain radiographs serve as crucial instruments for diagnosing, differentiating, and monitoring disorders among different types of arthritis although these disorders are very inexpensive and easily accessible, radiological alterations in them can appear gradually, necessitating a skilled evaluation by qualified experts. Obstacles like hard work, exhaustion, negligence, and time restraints might make it easier to miss significant discoveries. The extensive use of plain radiography may be restricted in certain healthcare facilities, particularly in Ethiopia and Africa, due to the lack of experts such as radiologists, rheumatologists, or physical therapists. Furthermore, skilled medical professionals, including radiologists and rheumatologists, can use Computer-Aided Diagnosis (CAD) techniques to improve decision-making and accurate diagnosis.

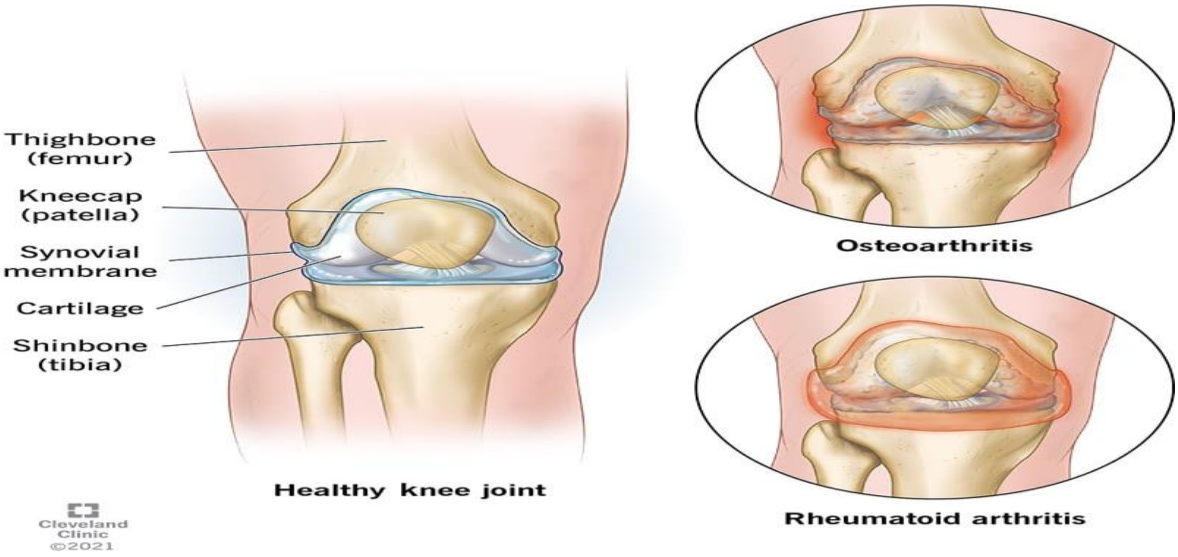


Figure 1 Knee Arthritis Disease Visualization

1.2. Motivation of the Study

Arthritis is a common and disabling disease that is often difficult to diagnose and classify early, significantly impacting patient outcomes and management. Traditional methods for detection arthritis frequently rely on subjective human interpretation, leading to unpredictability and delays in diagnosis. This problem is made worse by a shortage of radiologists in many primary healthcare settings and the similar symptoms across different types of arthritis, which further delay the timely examination of radiographic images. The urgent need to bridge this diagnostic gap by applying attention deep learning methods to knee X-ray images motivates our research. The limitations of manual analysis and the lack of specialized expertise underscore the need for more reliable and efficient diagnostic solutions. Therefore, we aim to develop a trustworthy and effective method for classifying and detecting arthritis. This research intends to enhance early intervention and the diagnostic process by filling gaps in the literature and leveraging attention-based deep learning, ultimately improving patient care and outcomes on a broader scale.

1.3. Statement of the Problem

The detection and classification of different arthritis types remains a complex challenge in the field of medical imaging, particularly with knee X-ray images. The complexity arises from the similarities in symptoms among different arthritis diseases and the ongoing uncertainty surrounding the exact causes of most types. Previous research has predominantly focused on detecting individual arthritis diseases, leaving a considerable gap in techniques to detect and classify different arthritic conditions.

While significant efforts by Üreten and Maraş (2022) and Ma et al. (2023) have addressed multiple types of arthritis detection and classification, their findings underscore the need for further research. These studies highlight the critical need for more techniques to the detection and classification of different arthritis diseases.

This issue is made worse by the widespread scarcity of certified radiologists, which is particularly acute in poor countries such as Ethiopia. Many institutions in these regions lack the expertise required for accurate arthritis diagnosis from radiographs. This shortage not only delays arthritis detection but also burdens healthcare providers, who struggle to diagnose and categorize arthritis accurately without specialist assistance. This issue is particularly acute in Ethiopian hospitals, where the lack of radiologists presents a significant barrier to accurate arthritis diagnosis. Consequently, physicians face substantial challenges in initiating timely treatments and providing appropriate care to patients with various types of arthritis.

This research aims to address these gaps by developing an attention-base deep learning model specifically designed for knee arthritis classification and detection. By doing so, it seeks to empower medical professionals in resource-limited settings to overcome diagnostic challenges and improve the care provided to patients with arthritis.

1.4. Research Questions

In solving the problems stated above, attempts are made to answer the following questions:

RQ1. What impact does multi-scale feature extraction have on the classification of arthritis?

RQ2. To what extent does the integration of attention mechanisms improve the detection and classification of arthritis diseases?

RQ3. Which deep learning model performs best in the detection and classification of arthritis diseases?

1.5. Objectives of the Study

1.5.1. General Objective of the Study

The general objective of this study is to design and implement an Arthritis disease detection and classification model in X-ray using attention deep learning.

1.5.2. Specific Objectives of the Study

The following specific objectives are addressed to achieve the general objective.

- To conduct a literature review on existing radiograph-based arthritis classification methods.
- To develop and implement the proposed deep learning model integrating attention mechanisms and multi-scale feature extraction.
- To evaluate the effects of attention mechanisms on the model's performance in arthritis disease detection.
- To examine the effect of multi scale feature extraction on model performance in arthritis detection.
- To measure and evaluate the performance of the proposed model.

1.6. Scope and Limitations of the Study

1.6.1 Scope of the Study

This study focuses on the detection and classification of various arthritic diseases using knee X-ray images, addressing the challenges posed by symptom similarities and the complex nature of arthritis diagnosis. It integrates advanced deep learning techniques, specifically multi-scale feature extraction and attention mechanisms, to enhance model accuracy and reliability. The research encompasses a thorough literature review, evaluate the effects of attention mechanisms and multi-scale feature extraction on model performance, and compare the proposed model to existing state-of-the-art methods. By targeting environments with limited healthcare resources, such as those in Ethiopia, the study aims to provide a reliable, automated diagnostic tool that can alleviate the shortage of qualified radiologists and improve patient care. As a pioneering effort, this research not only fills significant gaps in current literature but also lays the groundwork for future advancements in medical image analysis and disease diagnosis.

1.6.2 Limitation of the Study

There are some restrictions on this study. It excludes many additional types of arthritis that could benefit from similar diagnostic procedures by concentrating on detecting only four categories of arthritis: Normal, Osteoarthritis (OA), Rheumatoid Arthritis (RA), and Other. Additionally, because this research focuses exclusively on knee arthritis, it is unable to detect or classify arthritis in other body areas. Furthermore, the study's robustness and generalizability of the proposed model are limited by the small number of datasets included. The lack of digital systems and infrastructure in many hospitals makes it difficult to obtain digital knee X-ray images, which further reduces the amount of data that can be used for training and validation. Most hospitals' lack of digital support highlights a significant obstacle to gathering the large number of datasets required for developing and evaluating efficient deep learning models for the diagnosis of arthritis.

1.7. Significance of the Study

This study has significant implications for the fields of medical imaging and deep learning, particularly in detecting arthritis using knee X-ray images. By shifting from a traditional focus on individual arthritis diseases to a more approach, the study aims for more accurate and insightful diagnoses through the integration of attention mechanisms and multi-scale feature extraction in deep learning models. This can lead to immediate action and better patient outcomes through early diagnosis, facilitating timely interventions and preventing disease progression. Additionally, as one of the first study that detect and classify different arthritic disorders using knee X-ray images, the research addresses a significant gap in the literature by developing a technique that distinguishes between different types of arthritis. Furthermore, by developing and validating a deep learning model for arthritis classification, this study lays the foundation for future research, opening new avenues for exploring advanced deep learning techniques in medical diagnostics. The evaluation of the proposed model against existing state-of-the-art models provides insights into the strengths and limitations of various approaches, guiding future improvements and innovations.

1.8. Organization of the Thesis

This thesis is divided into six chapters, a conclusion, and further research. In the first chapter, the scope, limitations, significance, research questions, objectives, background of the study, problem statement, and research questions are covered. The algorithms used to classify arthritis diseases, the state of the art at the time, gaps in the literature, and related studies are all covered in detail in the second chapter. The research's methods, instruments, and strategies are covered in more detail in the third chapter. The proposed solution's architecture, procedure, and design are presented in the fourth chapter. The fifth chapter describes how the proposed solution is put into practice using real code. The sixth chapter presents the experiment outcomes and shows how the study issues have been addressed by contrasting them with earlier studies. The thesis ends with a summary of the results and recommendations for more research.

CHAPTER TWO

2. LITERATURE REVIEW AND RELATED WORKS

2.1. Introduction

An enormous global health concern is arthritis, which is typified by joint damage and inflammation. Due to their subjectivity and variability, traditional diagnostic methods which sometimes depend on manual interpretation have drawbacks. A promising path to improving the precision and effectiveness of arthritis diagnosis is provided by the emergence of automated techniques, especially those that make use of deep learning and machine learning. By addressing the drawbacks of traditional techniques, automated systems may be able to produce more consistent and objective outcomes.

2.2. Arthritis and Its Types

A wide range of diseases, each with distinct causes and symptoms, are together referred to as arthritis. It is a joint degenerative condition that can be debilitating. A few types of arthritis include gout, psoriatic, juvenile, osteoarthritis, and rheumatoid. Following is a quick discussion of rheumatoid arthritis, osteoarthritis, and psoriatic arthritis.

2.2.1. Rheumatoid Arthritis

An inflammatory autoimmune condition, rheumatoid arthritis (RA) affects many organs as well as one or more joints (Li & Zhao, 2022). It is a disease that has an unknown cause and has been brought on by a confluence of environmental and genetic variables. The complicated connections between these parts affect the development and progress of disease (Kourilovitch et al., 2014). In general, morning stiffness and joint inflammation are used to categorize RA, therefore accurate diagnosis of the illness requires training and experience. Although this was not adequate for early disease analysis, the American College of Rheumatology (ACR) created a criterion for the diagnosis of rheumatoid arthritis in 1987 based on morning stiffness and joint swelling. Later that year, the ACR/EULAR established a new criterion for predicting rheumatic patients (Kourilovitch et al., 2014), as early detection and treatment of rheumatoid arthritis can decrease its progression.

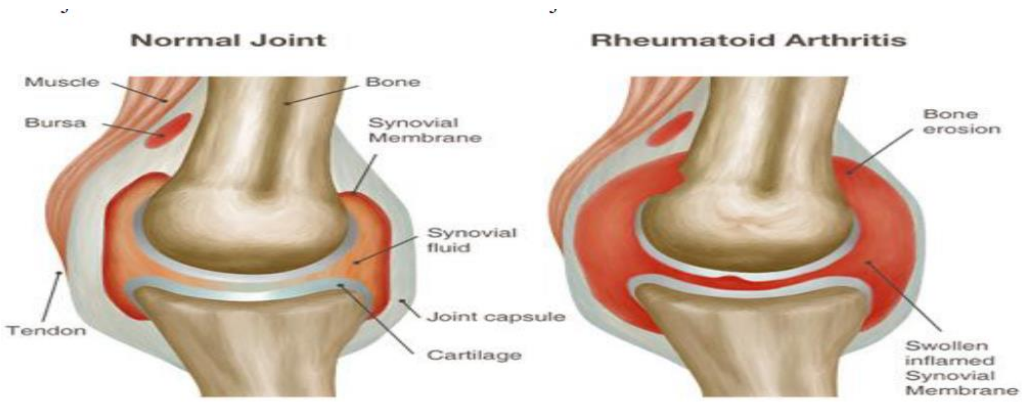


Figure 2 Normal joint vs Rheumatoid Arthritis

2.2.2 Osteoarthritis

Osteoarthritis (OA) is a common musculoskeletal disease that can cause severe disability in patients. Knee OA is the 11th leading cause of disability worldwide (Tiulpin & Saarakkala, 2020). OA is a condition that causes articular cartilage degradation. Cartilage is a smooth, stable layer that allows knee joints to move freely. In OA, cartilage explodes, loses flexibility, and weakens(S. Gornale et al., 2017). In general, it affects the knee, hip, spine, and foot joints. The basic symptoms of OA include joint discomfort and difficulty moving joints, as well as joint stiffness in the morning or after a lengthy rest. Because of the uncertain origin, OA is usually not recognized until it is too late for effective treatment, and occasionally expensive and invasive joint replacement surgery is required (Tiulpin & Saarakkala, 2020). Early detection of disease, on the other hand, can delay its progression. Aside from the many proposed OA diagnosis approaches, the Kellgren-Lawrence (KL) grading system is a gold clinical standard for classifying specific joints into five OA severity classifications(S. Gornale et al., 2017).

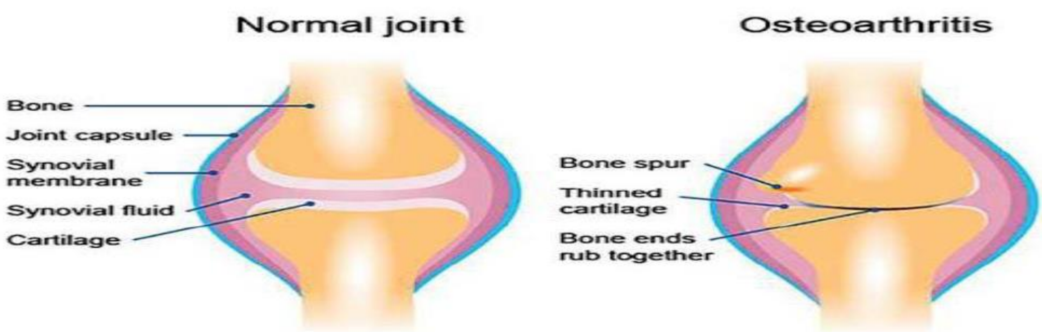


Figure 3 Normal joint vs Osteoarthritis.

2.2.3. Gout Arthritis

Gout is a common type of inflammatory arthritis that causes intense pain in a single joint, most commonly the big toe joint. The illness develops as episodic flares with intensified symptoms and of remission with no symptoms. Repeated gout attacks can lead to gouty arthritis, a more severe and degenerative form of arthritis. Gout is caused by an excess of uric acid in the body, which periods are a result of purine breakdown from both natural and nutritional sources. The inflammatory symptoms of gout are caused by the accumulation of uric acid crystals in joints, fluids, and tissues. While there is no cure for gout, good management through drugs and self-care measures is critical for symptom relief and disease prevention. Importantly, not everyone with hyperuricemia develops gout, and those who do not have symptoms usually do not require medical attention.

2.2.4. Psoriatic Arthritis

Psoriatic arthritis is a separate disorder defined by concurrent inflammation of the skin and joints, which is typically seen in people with psoriasis. Psoriasis appears as patchy, elevated, and red areas of irritated skin covered in scales, most commonly on the elbows, knees, scalp, navel, and vaginal or anal regions. While psoriasis is the first manifestation, only a fraction, ranging from 10% to 30% of psoriasis patients, progresses to psoriatic arthritis. Onset usually occurs between the ages of 30 and 50, however, it can begin in childhood. Men and women are equally affected by the illness. Psoriatic arthritis can cause swelling in the fingers and toes, as well as pitted or discolored fingernails. The joints affected can vary, with some people reporting involvement of a single joint, such as a knee, while others may experience involvement of many joints, including the spine, fingers, or toes. This complex interaction between skin inflammation and joint involvement emphasizes the diverse nature of psoriatic arthritis.

2.3. Related Work

Some researchers have researched the detection and classification of arthritis disease. Among them, we have reviewed the papers that are believed to have a strong relationship with this thesis.

(Üreten & Maraş, 2022) this paper aims to automate the classification of hand radiographs, specifically distinguishing between rheumatoid arthritis, osteoarthritis, and normal hand radiographs using deep learning approaches. The study employs well-known approaches such as the You Only Look Once (YOLO) algorithm for object detection and transfer learning with pre-trained networks, specifically VGG-16, to improve the model's classification performance. They used The YOLO method to aid in the exact diagnosis of arthritis-related anomalies in hand radiographs, and they applied transfer learning to improve the model's proficiency by using knowledge from large datasets. Furthermore, during training, the research employs data augmentation techniques such as rotation, translation, and flipping to augment the dataset and improve the model's robustness. The investigation yields good results, displaying high accuracy, sensitivity, specificity, precision, and AUC values. In the classification of rheumatoid arthritis and normal hand radiographs, accuracy, sensitivity, specificity, precision, and AUC results were 90.7%, 92.6%, 88.7%, 89.3%, and 0.97, respectively. In the classification of osteoarthritis and normal hand radiographs, accuracy, sensitivity, specificity, precision, and AUC results were 90.8%, 91.4%, 90.2%, 91.4%, and 0.96, respectively. The overall accuracy for the classification of rheumatoid arthritis, osteoarthritis, and normal hand radiographs was 80.6%. Future potential for collaboration with numerous centers, dataset diversification, and development of various deep learning models and methodologies to advance automated hand radiograph categorization are suggested by the study.

(Ma et al., 2023) this work uses hand radiography to conduct a complete study with the primary goal of creating and assessing a deep learning model for differentiating between rheumatoid arthritis (RA), osteoarthritis (OA), and the absence of arthritis. The researchers use a retrospective training strategy with a convolutional neural network (CNN) and analyze its discriminatory power on a large dataset of 9714 hand radiograph images. They carefully analyze numerous pretraining and training characteristics, such as imaging resolution, musculoskeletal radiograph pretraining, and consideration of diverse viewpoints, to determine their impact on model performance. The deep learning model achieves high AUC values of 0.975 for distinguishing between no arthritis and osteoarthritis/rheumatoid arthritis, and 0.955 for distinguishing between rheumatoid arthritis and no arthritis/osteoarthritis. The model achieves a kappa of 0.806 and an accuracy of 87.2% on the test set in three-way classification (no arthritis versus osteoarthritis versus rheumatoid arthritis), demonstrating its efficacy in discriminating among these arthritic diseases. Furthermore, the in-depth failure analysis of the paper sheds light on typical causes of inaccurate predictions, providing significant insights for prospective model modification.

(Berihun Molla, Teklu Urgessa (Associate Professor), & Worku B. (2021)) this paper aims to automate the classification of knee radiographs, specifically distinguishing between rheumatoid arthritis, osteoarthritis, gouty arthritis, and normal knee radiographs using CNN. The study employs custom model of CNN. Furthermore, during training, the research employs data augmentation techniques such as rotation, translation, and flipping to augment the dataset and improve the model's robustness. The investigation yields good results, displaying high accuracy, sensitivity, specificity and precision.

(Tariq et al., 2023) The study intends to further the field of knee osteoarthritis (KOA) detection by constructing a model using Convolutional Neural Networks (CNNs). To improve training, the study uses image adjustments like rotation and flipping during frontal image processing. Notably, a rank-consistent ordinal regression-based framework (CORN) is used for loss computation and grade prediction, and an ensemble model is used to combine outputs via a fully connected layer. The fact that they used an ImageNet pre-trained deep learning network as a feature extractor demonstrates the model's sophistication. The study, which used a single dataset from the Osteoarthritis Initiative (OAI) for training, testing, and validation, results in an ensemble model with an overall accuracy of 0.98, an overall precision of 0.98, and an overall F1-score of 0.97. The study suggests the limitation of dataset diversity and advises the use of other datasets for future research, giving significant insights and paving the way for further advancements in KOA detection using X-rays.

(Zhang et al., 2020) The study aims to improve knee osteoarthritis (OA) diagnosis utilizing a deep learning-based system, with a specific focus on automated Kellgren-Lawrence (KL) grade classification using plain radiographs. A two-step process is used with the Osteoarthritis Initiative (OAI) dataset: first, a ResNet-18 model is adjusted for knee joint localization, using mean squared error (MSE) loss optimization and model selection based on Intersection over Union (IoU). Following that, for KL-grade classification, a modified ResNet-34 model with a Convolutional Block Attention Module (CBAM) is used, improving accuracy by concentrated localization of key regions. The research exceeds previous techniques in terms of classification accuracy and performance metrics. Despite constraints such as a single dataset and an inconsistent KL-grade distribution, the study serves as a model for future research. Future studies should look at model generalization across different datasets and the use of weighted loss functions for improved classification in imbalanced class circumstances. Overall, the study provides a promising path for precise and automated KL-grade classification in knee osteoarthritis diagnosis using deep learning algorithms.

(Üreten et al., 2020) The purpose of this paper is to build an automated diagnostic approach for rheumatoid arthritis (RA) using hand radiographs and convolutional neural networks (CNNs). Based on a dataset of 135 right-hand radiographs (61 normal, 74 RA), the study demonstrates the CNN model's promising diagnostic ability, with an accuracy of 73.33% and a low error rate of 0.0167. CNN’s performance is further supported by data such as a sensitivity of 0.6818, specificity of 0.7826, and accuracy of 0.7500. These results highlight CNN’s capacity to correctly detect RA cases, indicating its potential as a useful tool for RA diagnosis. They utilize online data augmentation techniques, such as random horizontal and vertical translation and rotation, which contribute to the robustness of the CNN. The work not only addresses CNNs' diagnostic abilities for RA detection but also emphasizes current attempts to refine datasets and the future to differentiate RA and osteoarthritis (OA) from hand radiographs.

(Olsson et al., 2021) The study's goal is to improve the classification of knee osteoarthritis severity. The careful gathering of 6103 knee radiographic tests from Danderyd University Hospital between 2002 and 2016 is the first step in the strategy. These photos are manually classified using the Kellgren & Lawrence grading scale (KL), laying the groundwork for training and evaluation. The neural network training uses a ResNet architecture convolutional neural network (CNN) built in PyTorch, showcasing the study's dedication to cutting-edge deep learning techniques. CNN’s performance is thoroughly assessed against a test set of 300 exams, which is independently examined by top orthopedic surgeons during the evaluation phase. Interobserver consensus sessions provide a credible ground truth for network evaluation. Notably, the study uses customized output categories such as medial/lateral osteoarthritis (OA) to address complex features of osteoarthritis manifestation.

Table 1 Summary of Related Work

Authors and Year	**Class**	**Methodology**	**Limitation**
Üreten & Maraş, 2022	RA, OA, Normal (hand)	YOLO algorithm, Transfer Learning (Pretrained VGG16)	➢ It is only for hand x-ray images. ➢ Low performance. ➢ They suggested further research for multiple arthritis disease detection.

Ma et al., 2023	RA, OA, Normal (hand)	CNN	➢ It is only for hand x-ray images. ➢ Low performance.
Tariq et al., 2023	Knee Osteoarthritis (KOA)	Transfer learning and fine-tuned ResNet-34, VGG-19, DenseNet 121, and DenseNet 161 joined them in an ensemble	➢ It is only for Knee Osteoarthritis.
Zhang et al., 2020	Knee Osteoarthritis (OA)	ResNet-18, ResNet-34 with CBAM	➢ It is only for Knee Osteoarthritis.
Üreten et al., 2020	RA, Normal (hand)	CNN	➢ It is only for single class. ➢ Low performance (73.33% accuracy).
Olsson et al., 2021	Knee Osteoarthritis Severity	ResNet Architecture of CNN	➢ It is only for Knee Osteoarthritis.

The existing literature on automated arthritis detection and classification in radiographic images has primarily focused on individual diseases, such as rheumatoid arthritis (RA) or osteoarthritis (OA), with little attention paid to the simultaneous detection of multiple arthritis types in hand radiographs. (Üreten & Maraş, 2022) study shows advances in hand radiograph categorization using the You Only Look Once (YOLO) algorithm and transfer learning for RA, OA, and normal hand radiographs. Similarly, (Ma et al., 2023)explore the distinction between RA, OA, and the absence of arthritis in hand radiographs, reaching good accuracy. The key study gap, however, is a lack of studies addressing the simultaneous detection of multiple arthritis types utilizing knee radiographs. While both findings promote future collaboration, dataset diversification, and model development, more study is needed to explicitly target automated detection and classification of multiple arthritis types using knee arthritis. These studies show the untapped potential in this specific domain, suggesting additional exploration and innovation to improve the capabilities of automated systems in detecting many arthritic diseases at the same time.

CHAPTER THREE

3. RESEARCH METHODOLOGY

3.1 Overview of the Methodology

This section describes the study's methodology. The proposed methodology, data sources, and preprocessing stages are all provided. In addition, the model evaluation methods are described. So, to achieve the objective of this study the following procedures are employed.

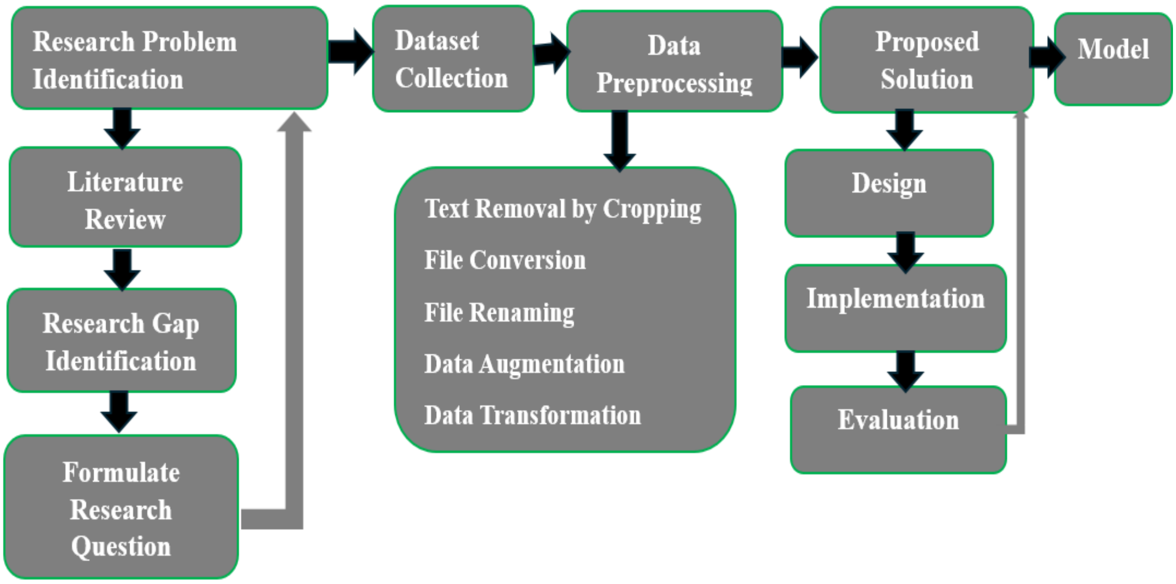


Figure 4 Research Overall Process

3.2. Datasets Source and Collection

Effective learning of deep learning models is strongly dependent on data. Strong secondary datasets are usually created by academics, or they can be obtained from publicly accessible datasets, which are essential for training these models. In this study, we concentrate on the analysis of arthritis using knee x-ray images. The process starts with image acquisition, which is the process of obtaining images from local hospital to make additional processing and analysis easier.

3.2.1 Datasets Collection Procedures:

Step-1. Patient Data Collection from OPD:

First, patient card numbers were gathered and filtered from Girum Hospital's outpatient data records. Various information is contained in these records, such as the diagnosis and demographics of the patients. An emphasis was made on searching through patient records that contained diagnoses for arthritis and its variations. Including relevant instances in the study was secured by taking this step.

Step-2. Retrieval from MedAxs Hospital Management System:

The next step required gaining access to Girum Hospital's MedAxs patient management system by using the filtered patient card numbers that were acquired from the OPD. More patient data was obtained by entering the recognized patient card numbers into MedAxs. This data included detailed medical records and imaging results.

Step-3. Examination of Knee X-ray Images and Reports:

After gaining entry to patient records in MedAxs, focus was placed on obtaining knee x-ray images and the reports that came with them from radiologists. Every knee x-ray image was carefully inspected, and the radiologist's report was studied to determine whether arthritis-related disorders were present, and how severe they were.

Step-4. Image Sorting and Organization:

The last stage was to systematically organize and sort the obtained images after reviewing the reports and knee x-ray images. Knee x-ray images were classified into four classes: Normal, Osteoarthritis (OA), Rheumatoid Arthritis (RA), and Other Arthritis, based on the diagnosis given in the radiologist reports. After that, the images were placed into the appropriate folders, guaranteeing a well-organized dataset for further examination.

Figure 6 Image Sorting and Organization

The above-described dataset collection processes were carefully planned to guarantee the capture the dataset for x-ray images of arthritis. This study provides a strong basis for future research and analysis in the field of medical imaging and deep learning by utilizing patient data from the hospital management system and the outpatient department, in addition to analysis of knee x-ray images and reports.

3.2.2 Datasets Description

The study's dataset, which includes x-ray images of arthritis obtained from Girum Hospital, covers 2018 to 2024. Dr. Girum Berhane, a board-certified American physician, and his family founded Girum Hospital in 2007 G.C. with the goal of bringing cutting-edge medical procedures to Ethiopia. The hospital has established itself as a national leader in the provision of high-quality healthcare services by investing in cutting-edge medical technologies and offering a wide range of medical specialties. In addition, Girum Hospital hopes to reduce the number of patients who need to be referred overseas by becoming into a hub for Ethiopian medical tourism. Table 3.1 lists all the images that were used in this study.

Table 2 Arthritis x-ray images taken from Girum Hospital

No	Arthritis Types	Total Number of images
1.	OA	152
2.	RA	103
3.	Other Arthritis	119
4.	Normal	180
	Total	**554**

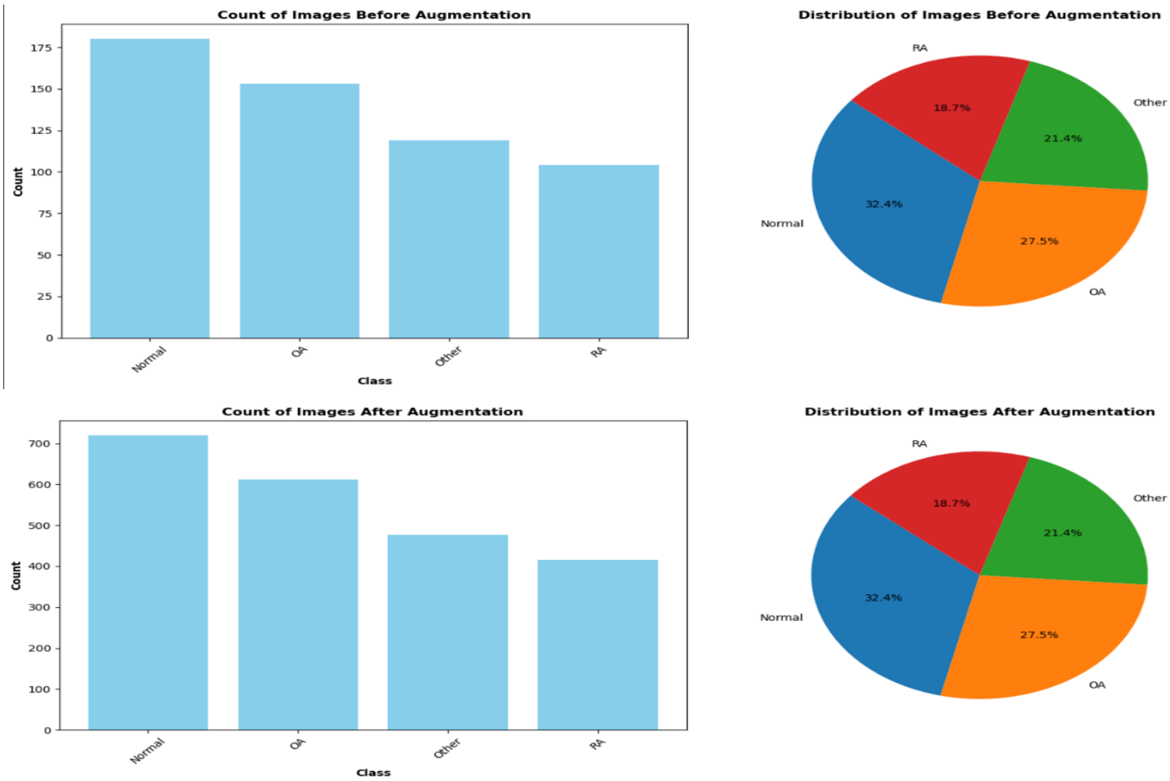


Figure 5 Number of Datasets and Its Distribution

Figure 6 Sample knee x-ray Arthritis Images, OA (A), RA (B), Other Arthritis (C)

3.3. Data Preprocessing Techniques

The medical images are pre-processed to make them appropriate for use with the Deep Learning model. To ensure that the images are Cropping, resizing, and standardization is required to achieve uniform size and format. To improve the model's performance, the preprocessing step may also include removing any unnecessary data or background noise from the images.

1. **Text Removal by Cropping**

   The initial preprocessing step for the paper was cropping the images to get remove of any writing. Eliminating recognizable text overlays and other sensitive data, such as patient card numbers and hospital names, was a crucial step. This method improved the model's ability to accurately detect the important components of the photos while protecting data privacy by focusing solely on pertinent visual content. Text was removed from the dataset by manually cropping the photos using image processing techniques, which improved the dataset's potential for additional analysis and model training.

2. **Data Transformation**

   Data transformation techniques were used to normalize the images for analysis once the text was removed. The images received a series of adjustments, such as uniform resizing, cropping to extract central regions of interest, tensor format conversion, and pixel value normalization. To guarantee uniformity in feature representation throughout the dataset and compliance with deep learning methods, this preprocessing step was essential. During implementation, libraries like PyTorch and OpenCV were used to carry out the transformations quickly and methodically.

3. **File Conversion**

   The format of the images in the dataset was then standardized by performing file conversion. During this procedure, images from different formats like JPEG and PNG were converted to a standard format like PNG. The standardization of image formats facilitated consistency and interoperability with tools for analysis and subsequent processing stages. Python libraries like Pillow or OpenCV were used to implement file conversion, guaranteeing a smooth connection with the research pipeline.

4. **File Renaming**

   To improve the organization and management of the dataset, file renaming was also used. Contextual information, like number identifiers or descriptive labels, was added to filenames to make it easier to track and retrieve specific images during analysis. This preprocessing stage accelerated the research workflow and enhanced dataset management. Python scripts were used to implement file renaming programmatically, automating the procedure, and minimizing human labor.

5. **Data Augmentation**

   Lastly, to improve model resilience and dataset diversity, data augmentation approaches were used. These methods added additional training data to the dataset by introducing variables such random brightness and contrast adjustment, gaussian blur and random rotation. Data augmentation enhanced the model's performance on unseen data and allowed it to be exposed to a greater variety of variances in the input data. Using libraries like OpenCV or TensorFlow, augmentation methods were implemented into the research pipeline to improve the quality of the dataset and the performance of the model in subsequent tasks.

Through the systematic implementation of these preprocessing techniques, the research paper ensured that the dataset was well-prepared for subsequent analysis and model training, ultimately leading to more accurate and reliable research outcomes.

3.3.1. Fine-Grained Feature Extraction

Using deep learning, fine-grid feature extraction emerges as a useful method for arthritis diagnosis and classification in knee radiographs. This method goes beyond previous methods by collecting detailed information from picture regions, which could lead to more accurate and robust diagnoses. Let us see the below methodologies and their implications for our investigation.

1. **Multi-Scale Feature Extraction:** In knee radiograph analysis, the practice of multi-scale feature extraction requires splitting the image into overlapping smaller patches and extracting features from each patch at several scales or sizes. This procedure is like zooming in on various parts of a map to discover hidden details. (Anastasis Alexopoulos1, Jukka Hirvasniemi2,*, 2021) demonstrated the effectiveness of this method. The advantages of using multi-scale are substantial because they allow the radiograph to capture both fine-grained features and large structural information. This examination may increase the identification of minor changes, such as joint space narrowing or changes in bone texture, which are important indications of osteoarthritis (OA). Chen et al. demonstrated high accuracy in OA classification by utilizing multi-scale feature extraction, emphasizing the practical benefits of this strategy in improving the sensitivity and specificity of knee radiograph analysis for osteoarthritis detection.
2. **Attention Mechanisms:** In the area of arthritis detection implementing knee radiographs, attention mechanisms approximate human visual attention by assigning weights to distinct image regions based on their relevance to the disease, like focusing on select sections of a painting that vividly communicate the story. This method is advantageous because it draws the model's attention to critical places, such as joint edges, where arthritic symptoms are most visible. Attention processes, as shown by (Zhang et al., 2020) have been shown to greatly increase the discrimination between

various forms of arthritis. Their findings highlight the effectiveness of attention mechanisms in improving the model's ability to recognize specific traits associated with rheumatoid arthritis, resulting in more accurate and detailed disease detection. The model's attention mechanism enables it to dynamically prioritize information, perhaps leading to a more complicated understanding of disease-related patterns in radiographic images.

3.4. Algorithms for Arthritis disease classification

3.4.1 Res2Net (Residual Attention Network)

Res2Net, or Residual Attention Network, is an extension of the ResNet architecture that adds hierarchical connections within each residual block that resemble residuals. The model's ability to efficiently capture multi-scale information thanks to its design improves performance. Res2Net also includes attention methods, which enable the network to concentrate on informative areas of the input image. Res2Net's attention processes and hierarchical connections can improve the model's capacity to recognize minute patterns suggestive of arthritis in the classification of arthritis diseases.

3.4.2 VGG (Visual Geometry Group)

This convolutional neural network (CNN) architecture is renowned for its efficiency and simplicity. It was created by the University of Oxford's Visual Geometry Group. It is composed of many convolutional layers, max-pooling layers, and fully linked layers on top. Various VGG variations, like VGG16 and VGG19, have been extensively employed in image classification applications, encompassing medical image analysis. VGG is a practical contender for arthritic disease classification because of its capacity to extract complex patterns and features from X-ray images.

3.4.3 ResNet (Residual Networks)

Microsoft Research introduced ResNet (Residual Networks), a revolutionary framework for residual learning that completely changed the field of deep learning. To solve the disappearing gradient issue, ResNet adds skip connections, which facilitate more direct gradient flow during training. With this design, performance may be maintained when training incredibly deep networks with hundreds of layers. ResNet can learn complex features from X-ray images, which improves the model's accuracy and generalization when it comes to arthritis classification.

3.4.4 DenseNet (Densely Connected Convolutional Networks)

Facebook AI Research researchers proposed DenseNet (Densely Connected Convolutional Networks), which incorporates dense connectivity patterns across layers so that every layer receives input from every layer that comes before it. Improved parameter efficiency and feature propagation result from this architecture, which promotes feature reuse and makes gradient flow throughout the network easier. Because of its dense connection and feature reuse, DenseNet is a good fit for jobs involving medical image analysis, such as arthritis classification, where precise diagnosis depends on minute features in X-ray images.

3.4.5 InceptionV3

Its inception modules use several parallel convolutional procedures of different kernel sizes to collect features at different scales. The network's capacity to effectively extract a variety of multi-scale properties is improved by this architecture. Factorized convolutions and batch normalization are two more features that InceptionV3, a variation of Inception, adds to enhance training effectiveness and performance. Fine-grained details from X-ray images can be captured with the help of Inception's multi-scale feature extraction capabilities in arthritis classification.

3.5 Development Tools

This research is developed using a variety of design and development tools. A design tool is used to create flowcharts and diagrams in various portions of this study. Different programming environments and libraries are also used to construct this research. This section provides a summary of these design and development tools.

3.5.1 Design Tools

Canva.com makes it simple for anybody to create great creative assets. Canva is an amazing tool for helping us level up our design skills, even if we're just getting started, as it allows us to build flowcharts and custom-made diagrams, multimedia presentations, and much more. It assists us in storing the diagram on the cloud or on a local device.

3.5.2 Hardware Tools

The following hardware tools are used in this research implementation.

Table 3 Hardware Tools

No.	Tools	Used for
1.	RAM	Working with the GPU to accelerate the training process
2.	GPU	To increase computation and speed up the training process
3.	Hard Disk	To store the datasets

3.5.3 Software Tools

To implement the study via coding, various writing software, and coding tools are employed, and these are given below with descriptions.

Table 4 Software Tools

No.	Software Tools	Description
1.	Python	Python is a high-level, general-purpose programming language with an extensive number of modules and frameworks that make writing easier and reduce development time.
2.	Anaconda	Anaconda is a program that is used to install the most recent version of Python, along with its various modules and IDEs, and to carry out the recommended solution.
3.	PyTorch	Its versatility and user-friendliness make it a popular choice for research and production applications in computer vision, natural language processing, and other fields.
4.	Jupyter Notebook	It is a free and open-source web-based application with coding and real-time visualization.
5.	Keras	Keras is an open-source package that serves as a bridge between Python and PyTorch by providing a Python interface for neural networks.
6.	Scikit-learn	It is a Python open-source software library containing machine learning methods.

3.6 Evaluation Methods

It is essential to utilize scientific assessment techniques that offer valuable perspectives on the developed model's capabilities, shortcomings, and overall performance while assessing its effectiveness. In this research, the dataset is divided into 60% for training, 20% for validation, and 20% for testing. This approach ensures that the model is trained on a majority portion of the data, validated to tune its parameters, and tested on a separate, unseen dataset to evaluate its performance.

The most common issue in deep learning is overfitting, which happens when a model works well on training data but not on unseen data. To mitigate overfitting, the model needs to be tested on data that wasn't included in the training set. While train-test split and other conventional evaluation methods have been used, they can still result in overfitting if not managed correctly. The 60-20-20 split helps in providing a robust evaluation by ensuring that the model is exposed to unseen data during the validation and testing phases.

Accuracy is a basic performance metric that is obtained from the confusion matrix. On the other hand, depending only on accuracy might not offer a thorough understanding of the model's functionality. As a result, extra assessment metrics like F1-score, recall, and precision are used.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}.$$

Equation 1 Accuracy Performance Measure

Precision shows how well the model avoids false positives by quantifying the percentage of true positive findings among all positive results. Conversely, **recall** highlights the model's ability to prevent false negatives by measuring its accuracy in identifying all pertinent events.

$$Precision = \frac{TP}{TP + FP}$$

Equation 2 Precision Performance Measure

$$Recall = \frac{TP}{TP + FN}$$

Equation 3 Recall Performance Measure

F1-score: also known as the F-measure, serves as a harmonic mean of precision and recall, offering a balanced assessment of the model's performance. A higher F1-score indicates better overall performance.

$$F = 2 * \frac{Precision * Recall}{Precision + Recall}$$

Equation 4 F-score Performance Measure

True positives (TP) in this sense indicate images that were accurately detected as positive, while false positives (FP) indicate images that were incorrectly labeled as positive. False negative (FN) denotes positive images that were mistakenly classified as negative, whereas true negative (TN) denotes negative images that were accurately identified as negative.

The research attempts to thoroughly evaluate the developed model's performance, pinpoint areas for enhancement, and guarantee its ability to endure in practical applications by utilizing these assessment metrics and methodologies.

CHAPTER FOUR

4. PROPOSED SOLUTION

4.1 Chapter Overview

In this chapter present our proposed deep learning strategy and pretrained models with attention multi scale for the classification of arthritis diseases. We describe the architecture and design choices of proposed model, which integrates attention and multi scale and other recent designs of convolutional neural networks, such as Inception, VGG, ResNet, Res2Net and DenseNet. as well as diagrams and descriptions of X-ray images data preprocessing approaches.

4.2 Proposed Solution Architecture

Using deep learning models to analyze medical imaging, the proposed solution architecture for arthritis disease classification is intended to maximize its capabilities. X-ray images are the first input, as they are a frequently employed modality in the diagnosis of arthritis. A thorough preprocessing workflow is applied to these images to guarantee data consistency and integrity. This involves text removal through cropping to get rid of any unnecessary comments or markings that might interfere with the model's analysis. To ensure a smooth incorporation into the classification pipeline, the data is next subjected to transformation, conversion, and renaming procedures that standardize the format and metadata. Additionally, data augmentation techniques are used to add changes to the dataset, improving the model's capacity to generalize under various circumstances.

After preprocessing, the core of the proposed architecture is based on a fine-grained, scaled classification model implemented using Res2Net as the foundation. Res2Net is well known for its capacity to capture complex visual features at many scale, which helps the model identify small patterns and irregularities that could be signs of arthritis. This intricate architecture improves the model's sensitivity to tiny variations in X-ray images, enabling dependable and accurate classification. Moreover, a fully connected layer receives the output from the Res2Net backbone, where the retrieved features are combined and mapped to the appropriate arthritis classifications. The propoed solution architecture offers physicians insights into arthritis conditions by classifying images into four groups: Osteoarthritis (OA), Rheumatoid Arthritis (RA), Other arthritis, and Normal. This allows for immediate action and individualized treatment plans.

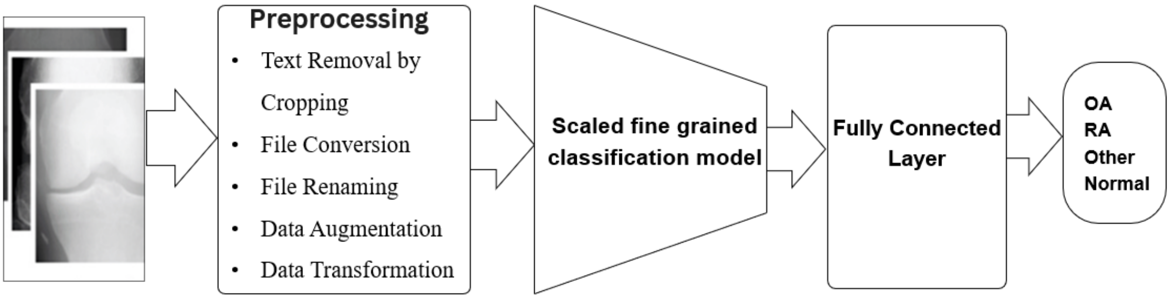


Figure 7 Proposed approach for Arthritis classification.

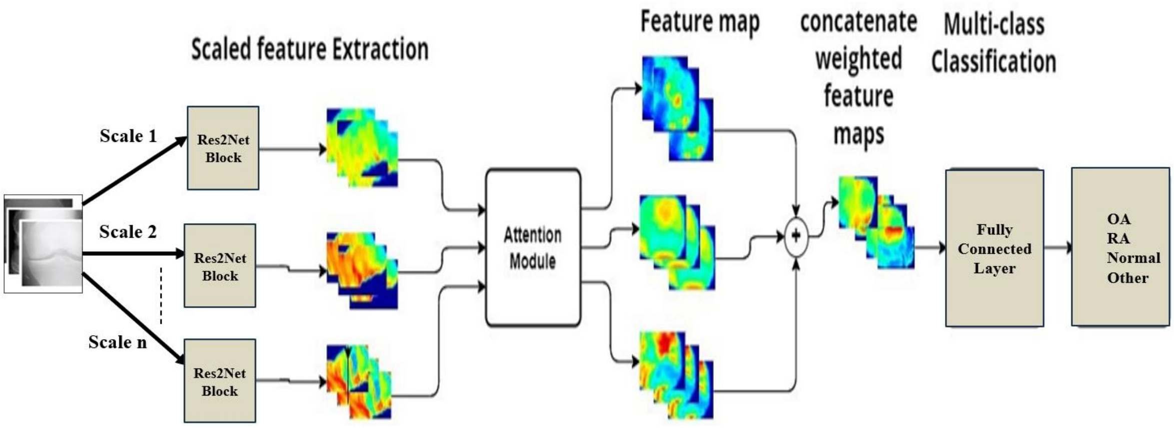


Figure 8 Scaled Fine Grained Classification Model.

Classification of arthritis disease using images from X-rays is made possible by the proposed solution architecture, which combines thorough data preprocessing and cutting-edge deep learning algorithms. The architecture guarantees precise and detailed diagnosis by utilizing the powers of Res2Net and a hierarchical classification scheme. This provides healthcare practitioners with important information for patient management and care.

4.3. Knee X-ray Images Pre-Processing

Preprocessing is essential to this research on knee X-ray classification since it helps to refine the input data prior to analysis. Several essential phases are involved in the process: cutting textual comments, transferring information to appropriate formats, renaming files for organization, altering images for standardization, and augmenting data for greater diversity. By taking these procedures together, the dataset's quality and consistency are improved, resulting in more reliable and accurate classification models. Preprocessing ensures that subsequent analysis is based on clean and accurate data by addressing artifacts and undesired text included in raw knee X-ray images, as seen by the example below.

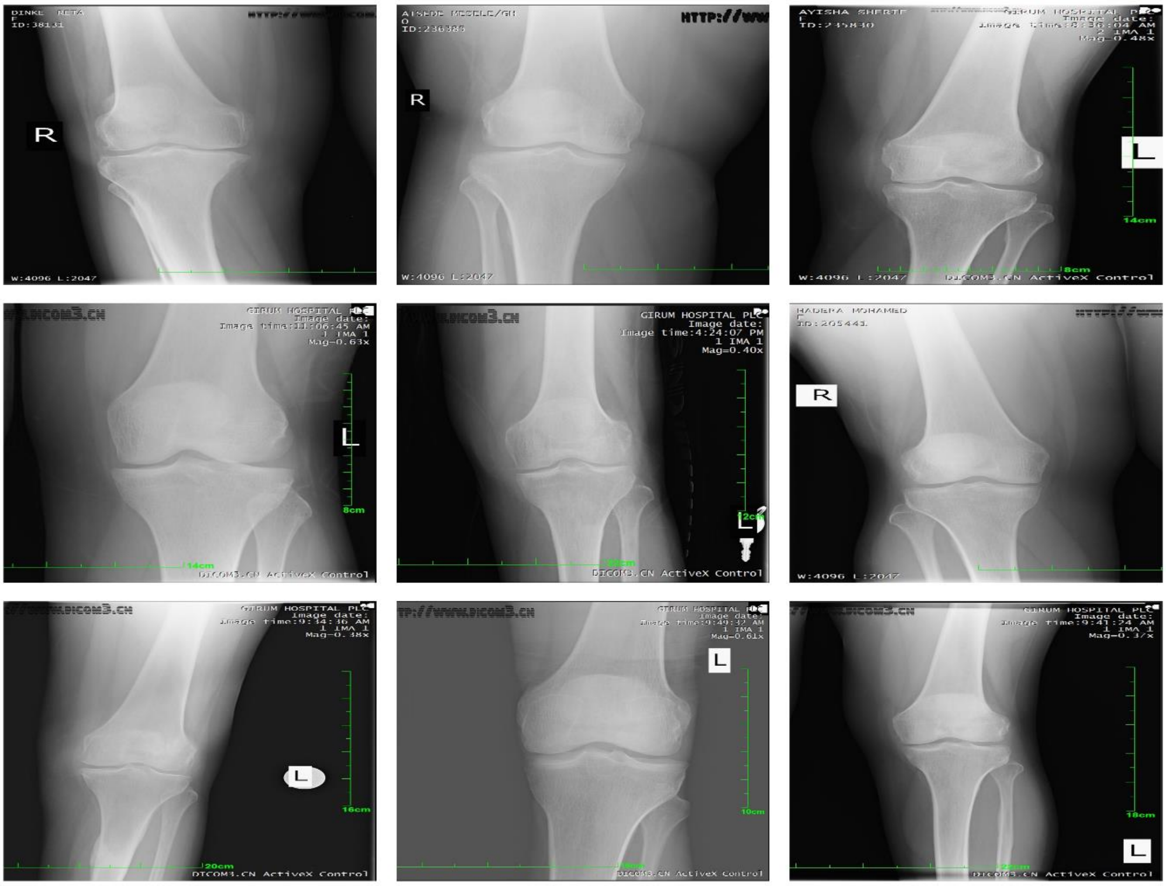


Figure 9 Sample Local Knee X-ray Images before Preprocessing.

4.3.1 Text Removal by Cropping

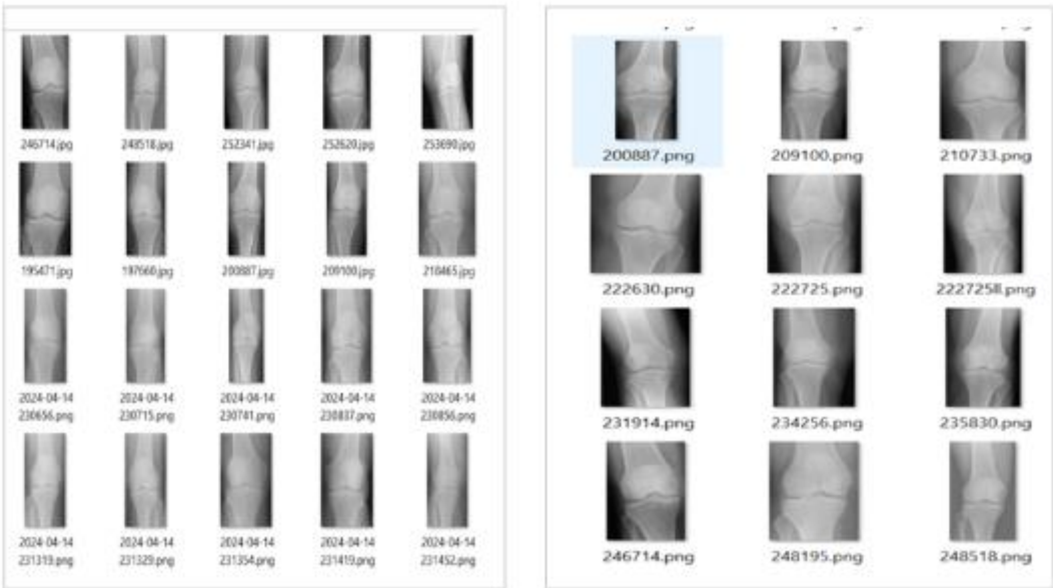
Textual annotations, such as patient data or radiologist remarks, were included on many knee X-ray images, which impeded image processing. By locating and isolating the region of interest that only included the knee joint and its surrounding structures, cropping aimed to eliminate these annotations. By doing this step, it was guaranteed that the model would only focus on the anatomical traits that were essential for classification.

(A) (B)

Figure 10 Text Removal By Cropping (A) Before Removal (B) After Removal

4.3.2. File Format Conversion

Images from the knee X-ray collection were kept in different file formats, such as PNG and JPG. The preparation pipeline was streamlined by converting all files to a single format of PNG in this case. This conversion improved data management and streamlined the workflow by guaranteeing consistency throughout the dataset and enabling smooth incorporation into later processing stages.

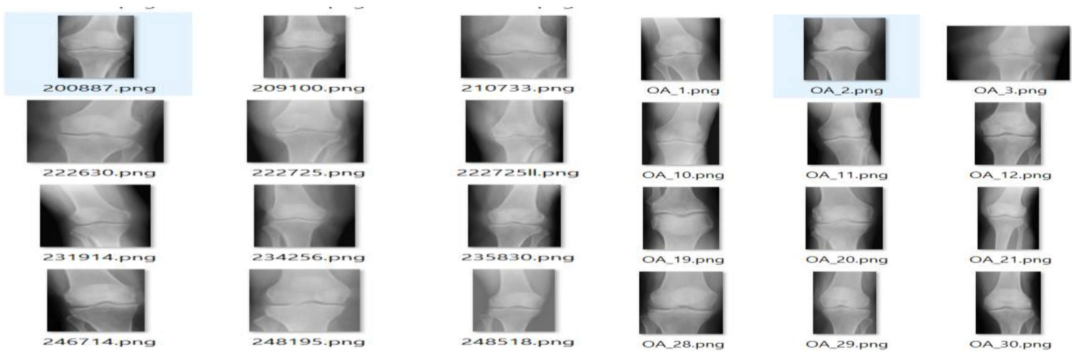


(A) (B)

Figure 11 File Format Conversion (A) Before Change (B) After Change.

4.3.3. File Renaming

To improve the organization and management of the dataset, files were renamed with clear and understandable names. Assigning distinct identifiers or patient data to every image file was usually the goal of this stage, which made the data easier to monitor and refer to during analysis and model training.



(A) (B)

Figure 12 File Renaming (A) Before Rename (B) After Rename.

4.3.4. Data Augmentation

As a crucial preprocessing step in the study, data augmentation techniques were applied with the goal of improving the model's robustness and capacity for generalization as well as enriching the dataset. Variations like random brightness and contrast adjustment, gaussian blur and random rotation modifications greatly increased the diversity of the sample. This process of augmentation exposed the model to a wider variety of data representations, which not only diversified the dataset but also decreased the likelihood of overfitting. The study aimed to simulate real-world variances and difficulties found in knee X-ray images by integrating random brightness and contrast adjustment, gaussian blur and random rotation augmentation techniques. As such, the model's robustness and performance against unknown data were enhanced, providing a solid basis for precise and trustworthy disease classification.

4.3.5. Data Transformation

Standardization approaches were utilized during the data transformation step to guarantee consistency in the knee X-ray image dimensions, orientation, and pixel values. A few transformation operations, including resizing, center cropping, and normalizing, were carried in this study. To enable uniform interpretation and comparison, these changes were essential in getting the images ready for input into the classification model. The images were converted into a standardized format appropriate for model ingestion by center cropping them to 224 by 224 pixels and scaling each image to a consistent size of 256 x 256 pixels. To further ensure that the pixel values fell inside a predetermined range, normalizing was applied. The dataset was optimized during this phase for use in later processing and classification activities.

The knee X-ray images were improved by each preprocessing stage, which made sure they were formatted, annotated, and enhanced suitably for efficient training and testing of classification models.

4.4 Algorithms for Proposed Model Development

The architecture demonstrated here improves knee X-ray classification of images performance by combining the Res2Net and SEModule modules. By adding a multi-scale feature integration mechanism that enables the aggregation of features from several scales, the Res2Net module expands the capabilities of the ResNet architecture and enhances its representational capacity. The network can capture deeper hierarchical characteristics thanks to the Bottle2neck structure, which consists of many convolutional layers with different receptive fields. In contrast, the SEModule uses an adaptive squeeze-and-excitation method to adjust channel-wise feature responses. The SEModule increases the discriminative capacity of the network by learning feature dependencies, which enables it to suppress irrelevant features and concentrate on more informative ones.

The Bottle2neck module is composed of several convolutional layers with varying sizes. Each layer combines feature maps from the preceding layer to produce representations that are more detailed. Each layer of the convolutional layers, which are arranged hierarchically, captures features that are more and more abstract. By adaptively recalibrating these characteristics' relevance based on channel-wise dependencies, the SEModule further improves them. An adaptive average pooling operation, two convolutional layers, and a sigmoid activation function are used in this recalibration procedure to provide a collection of attention weights that regulate the input data.

4.4.1 Res2Net Module

Res2Net is the original ResNet with the red portion on the right (which is at least directly connected without passing through 3 × 3 convolutions) in place of the center 3 × 3 convolution. The following Figures 17 illustrate this. Divide the feature mapping uniformly into subsets of feature mappings, represented by xi , where i ∈ {1, 2,..., s}, following a 1 ×

1 convolution. Each feature submap Xi has the same spatial size as the input features, but with one-third fewer channels (excluding $x1$), and each xi has a corresponding 3 × 3 convolutional transformation, represented by $Ki()$.The feature subgraphs $Ki()$ and Xi outputs are added together and supplied to $Ki-1()$. So, yi can be expressed as follows:

$$y_i = \begin{cases} x_i & i = 1; \\ K_i\,(x_i) & i = 2; \\ K_i\,(x_i + y_{i-1}) & 2 < i \leq s. \end{cases}$$

Equation 5 Each Scale Output

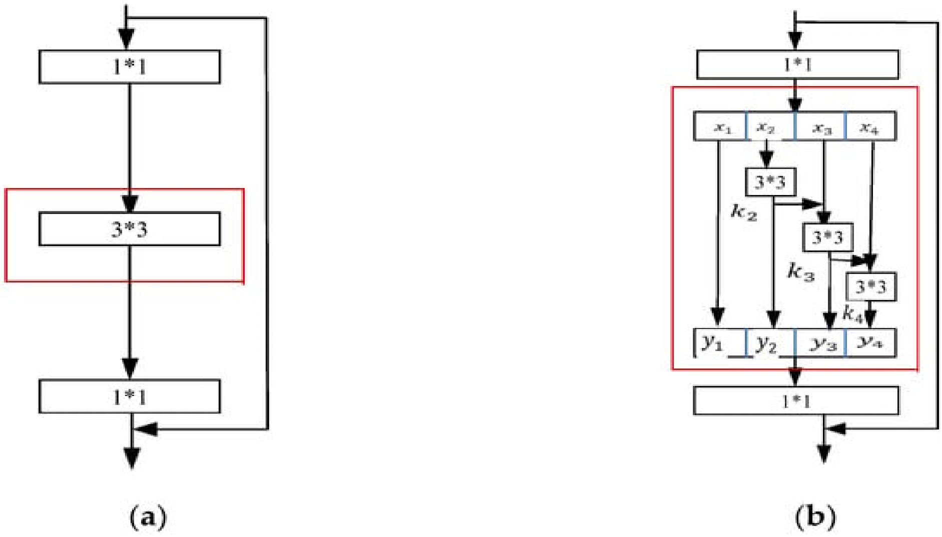


Figure 13 ResNet and Res2net diagram. (a) ResNet diagram; (b) Res2Net diagram.

(source: https://www.mdpi.com/2076-3417/13/7/4235)

It is important to remember that all feature partitions $\{xj,\ j \leq i\}$ may provide feature information to any 3 × 3 convolution operator $Ki()$. Every time a 3 × 3 convolution operator breaks down a feature, the resultant output could have a receptive field that is more than xj. The output of the Res2Net module contains a variety of numbers and combinations of receptive field sizes/scales because of the combinatorial explosion effect. The multi-scale processing of the breakdown in the Res2Net module makes it easier to extract both local and global information. To effectively fuse data at various scales, aggregate all splits and subject them to a 1 x 1 convolution. To improve processing, the splitting and cascading technique enables more effective forced convolution. Remove the first split's convolution to lower the number of parameters; this is also a type of feature reuse. S was employed in this work as the scale size control parameter (S. -H. Gao et al., 2021).

By using multi-scale feature fusion and integrating many branches with varying scales, the Res2Net module improves feature extraction and increases classification accuracy. Multiple accessible receptive fields with finer granularity are referred to as multi-scale.

4.4.2. SE-Res2Net Module

The squeeze-and-excitation network (SENet) was introduced by J. Hu et al. in 2018. Its purpose is to highlight significant characteristics and disregard unimportant ones by obtaining varying weights on the channel dimension of the feature map.

The SE module's network structure appears in the above diagram. Through a series of general transformations, a feature with a feature channel count of C can be obtained given an input X. The three operations listed below are also used to rescale the previously acquired features: (1) The squeeze procedure compresses the characteristics spatially, reducing them to a single integer per channel. This provides a picture of every channel globally. The dimensions of the input channels and the output are the same. (2) Similar to recurrent neural networks, the excitation operation is a gating mechanism. Each feature channel's weights are determined by a parameter W, which is learned to explicitly describe the feature channel correlation. (3) The output of the weight by excitation is considered as the significance of every feature channel following feature selection in the scaling operation. After the previous features have been multiplied one by one with the weights of each feature channel, the original features are completed and reconstructed in the channel dimension.

To maintain inter-channel correlation during channel grouping, the output y from the Res2Net module is fed into the SE module. In Figure 6, the network structure is displayed. The characteristics of this module are first compressed using global average pooling into $y'\in R1\times1\times1\times c$. After that, fully linked layers are used to fit the correlation between the channels, and a sigmoid activation function is used to normalize the results. Therefore, $fc=\sigma(FC(\delta(FC(y')))$ is the weight vector of the channels, where FC indicates the fully connected layer, σ stands for the ReLU function, and δ for the sigmoid function. The SE module's output is $f=f'+x$. To finish the feature adjustment, the original features in the channel dimension are rescaled inside the residual unit. Ultimately, a jump connection is used to link the residual unit's input (x) to its output (f') to derive the SE-Res2Net module's output, which is $f'=fc\cdot y$ (Aysa Z et al., 2023). The advantages of the SE module are increased by reassigning alternative weights to the channel features, removing invalid features, and maximizing the use of the single-layer features when the SE module is fused after the 1 × 1 convolution. The SE-Res2Net module developed in this work can enhance the model's stability, emphasize residual mapping, and encourage network convergence.

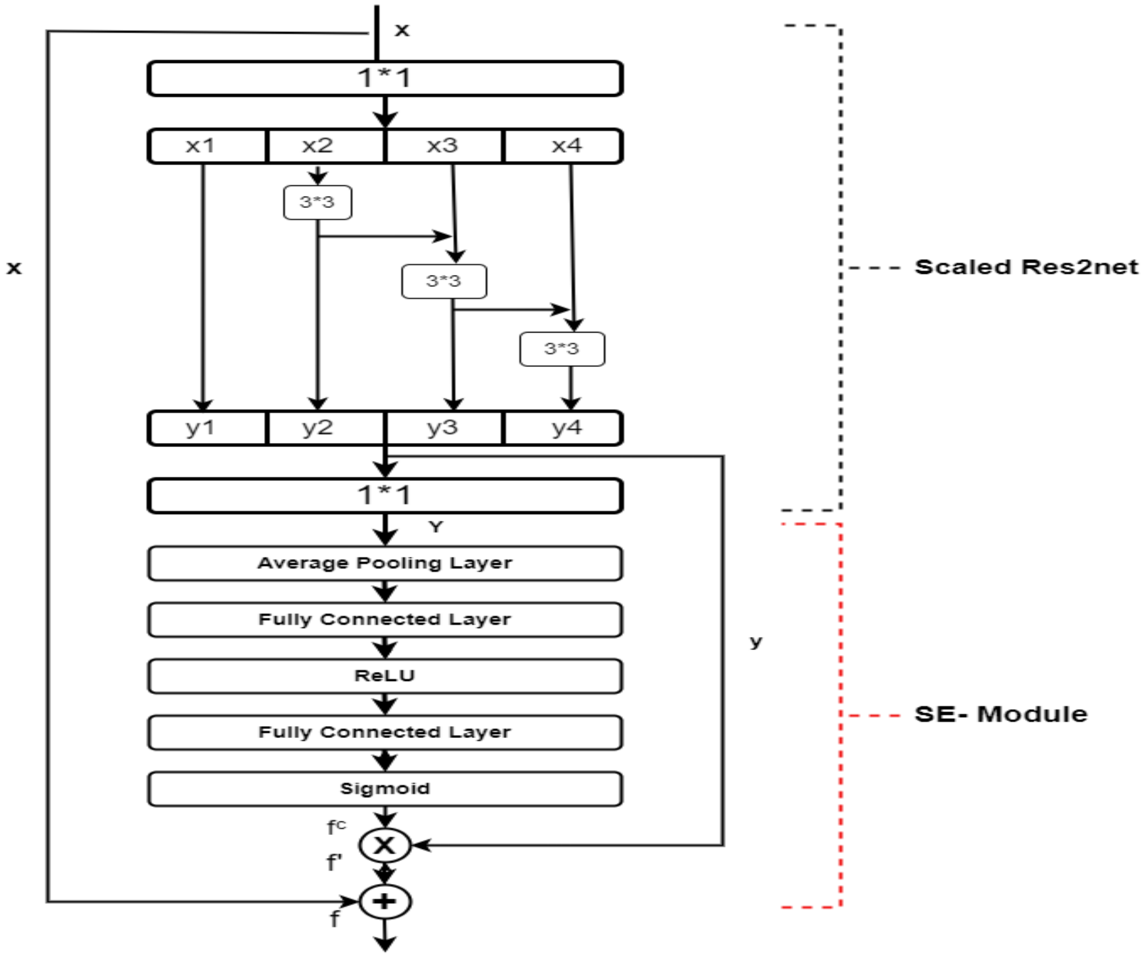


Figure 14 SE-Res2Net Network Structure

The proposed architecture offers better performance in knee X-ray image classification tasks by integrating the Res2Net and SEModule modules. The SEModule improves the discriminative power of these features by channel-wise recalibration, while the Res2Net module allows the network to efficiently capture multi-scale features. This all-inclusive method yields a highly expressive and discriminative model that performs exceptionally well in knee X-ray image classification across a wide range of pathological situations.

For this research, we focus on several variants of the Res2Net architecture, each configured with different combinations of base width and scale parameters. These variants include the

SE-Res2net50_14w_8s, **SE-Res2net101_26w_4s**, **SE-Res2net50_26w_6s**, **SE-Res2net50_26w_8s**, **SE-Res2net50_48w_2s**, and **SE-Res2net50_26w_4s** models.

The SE-Res2net50_14w_8s model represents a variant with a lower base width of 14 and a higher scale of 8. This configuration aims to investigate the effects of reduced width and increased scale on feature representation capacity. On the other hand, the SE-Res2net101_26w_4s model is a more complex variant featuring 101 layers while maintaining a base width of 26 and a scale of 4. Leveraging a deeper architecture, this model aims to capture intricate patterns and hierarchical representations more effectively.

Additionally, variants such as the SE-Res2net50_26w_6s and SE-Res2net50_26w_8s models aim to increase the scale of feature maps within each block, potentially capturing more detailed and contextually rich information for arthritis classification. Conversely, models like the SE-Res2net50_48w_2s and SE-Res2net50_26w_4s explore the impact of increased width on feature representation while maintaining a simpler scale factor, striking a balance between computational efficiency and feature representation capacity.

By evaluating and comparing these Res2Net variants, aim to gain insights into their effectiveness for arthritis detection and classification tasks. Through rigorous experimentation and analysis, seek to identify the most suitable model architecture that can accurately classify arthritis subtypes while minimizing computational overhead. Ultimately, this research endeavors to contribute to the advancement of medical image analysis techniques, with potential applications in clinical diagnosis and treatment planning for arthritis patients.

CHAPTER FIVE

5. PROPOSED SOLUTION IMPLEMENTATION

5.1 Chapter Overview

This chapter, the focus is on presenting and discussing the implementation of the proposed solution, starting with the preparation of the environment to provide a suitable framework for model building. The initial processing of knee X-ray images then explained in depth, covering methods such data transformation, augmentation, and text removal. The chapter also shows model implementations that are specifically designed to handle this disease classification. To give a thorough grasp of the solution's implementation process and its underlying complexities, each step is carefully documented.

5.2 Environment Setup

In the implementation of this thesis, three laptop computers were utilized for environment setup. The first laptop is equipped with 28GB of RAM, running Windows 11, and powered by an Intel(R) Core i5 processor clocked at 2.60GHz. The second and third laptops each feature 8GB of RAM, operating on Windows 11 Enterprise, and powered by 11th Gen Intel(R) Core (TM) i5-1145G7 processors running at 2.60GHz with a boost clock of 2.61 GHz.

Python was selected as the programming language for the development of the Arthritis disease classification model because of its adaptability to deep learning settings and ease of use.

For developing the model, many libraries and frameworks were setup for my environment:

Table 5 Libraries and frameworks setup for my environment

No	Name	Version
1.	Jupyter Lab	4.1.8
2.	Matplotlib	3.8.4
3.	NumPy	1.26.4
4.	Pandas	2.2.2
5.	Python	3.12.3
6.	PyTorch	2.3.0

7.	Tqdm	4.66.2
8.	Scikit-learn	1.3.0

5.3. Implementation of Preprocessing for Knee X-ray Images

5.3.1. File Format Conversion

Images from the knee X-ray collection were kept in different file formats, such as PNG and JPG. The preparation pipeline was streamlined by converting all files to a single format—PNG in this case.

5.3.2. File Renaming

To improve the organization and management of the dataset, files were renamed with clear and understandable names. Assigning distinct identifiers or patient data to every image file was usually the goal of this stage, which made the data easier to monitor and refer to during analysis and model training.

5.3.3. Data Augmentation

As a crucial preprocessing step in the study, data augmentation techniques were applied with the goal of improving the model's robustness and capacity for generalization as well as enriching the dataset. Variations like flipping, rotation, and arbitrary brightness and contrast modifications greatly increased the diversity of the sample.

5.3.4. Data Transformation

Standardization approaches were utilized during the data transformation step to guarantee consistency in the knee X-ray image dimensions, orientation, and pixel values. A few transformation operations, including resizing, center cropping, and normalizing, were carried out in this study.

5.4. Implement the Proposed Model

5.4.1 Define Res2Net model.

By integrating the Res2Net module, the Res2Net class provides improvements over conventional ResNet structures by encapsulating an advanced convolutional neural network architecture designed for image classification applications. The ability of the network to extract complex patterns from images is improved by this module, which presents a revolutionary multi-scale feature processing technique. Key elements in the class are defined via the initialization method, which also includes several residual layers including convolutional, batch normalization, activation, and pooling layers. The _make_layer method is used to create these residual layers, which include blocks of the desired type and amount. This allows for feature extraction at various scales. In the forward pass, a series of operations are applied to input tensors as they move through several layers and residual blocks to extract key features from the image. The generated features are then subjected to adaptive average pooling and subsequent classification using a fully linked layer. Renowned for its efficacy in recognizing complicated image patterns, this architecture offers a substantial improvement in image classification paradigms, particularly in settings needing complicated multi-scale feature analysis.

5.4.2 Define SEModule

The Squeeze-and-Excitation (SE) module is an essential part of convolutional neural network architectures, and it is represented by the SEModule class. By adding channel-wise attention techniques to the network, this module improves the network's ability to concentrate on important aspects when extracting features. The class specifies activation functions, convolution, and adaptive average pooling—operations necessary for channel-wise attention computation—in its initialization method. To be more precise, the module uses a series of convolutional and activation layers to compute channel-wise significance scores, which are then activated by a sigmoid to provide attention weights. Input tensors flow through these layers during the forward pass, when they undergo changes that highlight useful channels and suppress less relevant ones. By amplifying significant characteristics selectively, the resulting attention-weighted features are then multiplied element-wise with the original input, therefore increasing its representational power. The SEModule is a widely recognized tool for strengthening network discriminative capacity. It is particularly useful for improving feature representation and classification accuracy on a variety of image datasets.

5.4.3 Define Different Res2Net Models

These functions construct different Res2Net designs, each with unique parameters to meet distinct needs. For example, res2net101_26w_4s builds a 101-layer Res2Net model with a base Width of 26 and a scale of 4. Res2Net models with various configurations are produced by res2net50_26w_6s, res2net50_26w_8s, res2net50_48w_2s, and res2net50_14w_8s. These models make use of the Res2Net framework, which improves feature representation through a hierarchical decomposition technique. Additionally, they facilitate pre-training on ImageNet for transfer learning objectives, allowing features to be learnt for use in subsequent tasks.

5.4.4. Model Training

The training loop for a deep learning model with PyTorch is represented by this little bit of code. The model is set to training mode (model.train()) at the beginning of each epoch, and the parameters are updated by computing and backpropagating the loss. Many metrics, including loss, accuracy, precision, recall, and F1 score, are calculated and recorded during training to assess performance. A progress bar is provided by the tqdm library to monitor the training progress. The model's performance is tracked by printing the training metrics at the conclusion of each epoch. These measurements are also stored in lists for a cross-temporal analysis and display.

5.4.5. Model Testing and Evaluation

This evaluate model function is designed to evaluate the performance of a trained model on a given dataset. It sets the model to evaluation mode (model.eval()) and computes various evaluation metrics such as loss, accuracy, precision, recall, and F1 score. These metrics are calculated over the entire dataset using a loop over the provided data loader. The torch.no_grad() context manager ensures that no gradients are computed during evaluation to conserve memory. The function returns the average loss and the calculated evaluation metrics. Additionally, it handles cases where the dataset is imbalanced by setting zero_division=0 in the precision, recall, and F1 score calculations to avoid division by zero errors.

CHAPTER SIX

6. RESULT AND DISCUSION

6.1 Chapter Overview

In this chapter, the focus is on presenting and discussing the outcomes of the proposed architecture's performance in arthritis classification, alongside comparative analyses with other transfer learning models including VGG, DenseNet, ResNet, and Inception. The chapter begins with a detailed exposition of the results obtained from the proposed architecture, including its performance in accurately classifying arthritis cases. Subsequently, a discussion is provided in the strengths and limitations of the proposed approach in contrast to the other models examined.

6.2 Proposed Architecture Model Result

Several Res2Net architecture configurations have been investigated in attempts to optimize arthritis classification models. This choice includes Res2Net models with varying base widths and scales to take advantage of the network's capacity to efficiently capture characteristics at several scales. The versions known as **SE_Res2Net50_14w_8s**, **SE_Res2Net50_26w_4s**,**SE_Res2Net101_26w_4s**,**SE_Res2Net50_26w_6s**,**SE_Res2Net 50_26w_8s**, and **SE_Res2Net50_48w_2s** are unique architectural configurations, each created to take certain factors like computational efficiency and feature representation capacity into account. These models have the potential to provide insights into the relationship between classification performance and architectural design decisions in the field of arthritis diagnosis. The comparative effectiveness of each Res2Net variation is then thoroughly examined, offering insightful direction for next model selection and improvement initiatives.

Table 6 Summary of varying base width and scale of the proposed model

No	Model	Depth	Base Width	Scale
1.	SE_Res2Net50_14w_8s	50	14	8
2.	SE_Res2Net50_26w_4s	50	26	4
3.	SE_Res2Net101_26w_4s	101	26	4
4.	SE_Res2Net50_26w_6s	50	26	6
5.	SE_Res2Net50_26w_8s	50	26	8

6.	SE_Res2Net50_48w_2s	50	48	2

Each model variant was trained and evaluated on a dataset consisting of X-ray images categorized into four classes: Normal, Osteoarthritis (OA), Other, and Rheumatoid Arthritis (RA). The performance of each model was assessed using confusion matrix analysis, providing insights into its ability to accurately classify arthritis subtypes.

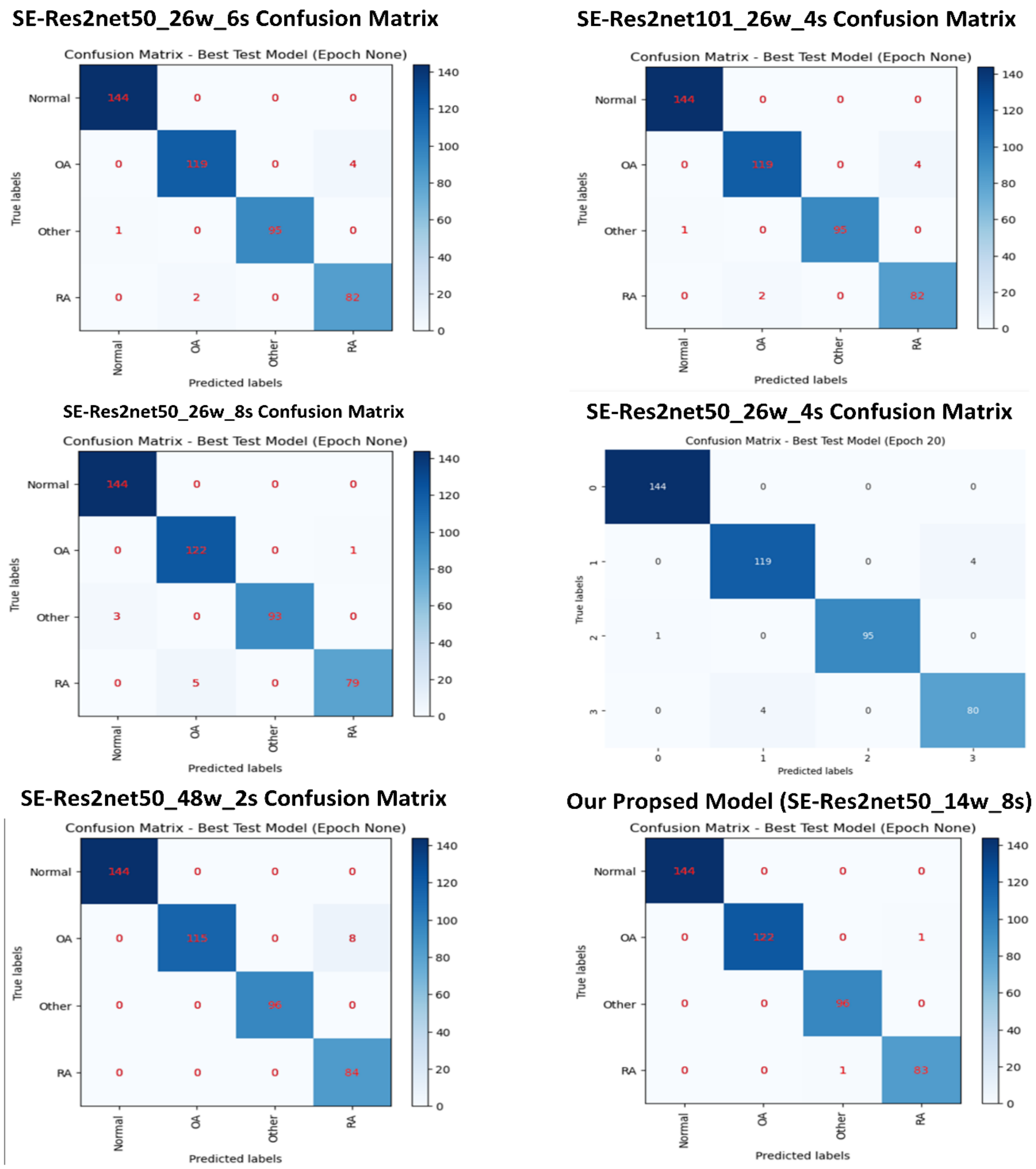


Figure 15 Res2Net models confusion matrix result.

In analyzing the confusion matrices presented in Figure 20, Conclude that.

SE-Res2net50_26w_6s Model

The SE-Res2net50_26w_6s model demonstrated robust performance across all arthritis subtypes, with minor difficulties observed in the OA and RA classes. Specifically, the model encountered around 4 false negatives in the OA class, indicating challenges in correctly detecting instances of osteoarthritis. Similarly, approximately 2 false negatives were recorded in the RA class, suggesting difficulties in accurately detecting cases of rheumatoid arthritis.

SE-Res2net101_26w_4s Model

Like the SE-Res2net50_26w_6s model, the SE-Res2net101_26w_4s model exhibited excellent performance in arthritis classification, with minimal difficulties observed across all classes. The model achieved near-perfect precision and recall scores for OA, Other, and RA classes, while accurately detecting normal cases without any misclassifications.

SE-Res2net50_26w_8s Model

The SE-Res2net50_26w_8s model demonstrated commendable performance in distinguishing between arthritis subtypes. However, the model encountered slight challenges in correctly detecting normal cases, resulting in a few false negatives. Specifically, around 1 false negative was observed in the OA class, and 5 false negatives were recorded in the RA class.

SE-Res2net50_26w_4s Model

The SE-Res2net50_26w_4s model performed admirably in arthritis classification, achieving high precision and recall scores across all classes. Minor difficulties were observed in the OA and RA classes, with approximately 4 false negatives recorded in the OA class and none in the RA class.

SE-Res2net50_48w_2s Model

The SE-Res2net50_48w_2s model demonstrated superior performance in arthritis classification, particularly in distinguishing between OA and RA cases. However, the model encountered difficulties in correctly detecting normal cases, resulting in a few false negatives. Specifically, around 8 false negatives were recorded in the OA class, and none were observed in the RA class.

Proposed Model (SE-Res2net50_14w_8s)

The proposed model, SE-Res2net50_14w_8s, exhibited outstanding performance in arthritis classification, surpassing the performance of other Res2Net variants. While it achieved near-zero false negatives across all classes, minor difficulties were observed in the OA and RA classes. Specifically, 1 false negative was recorded in each of these classes, indicating some challenges in accurately detecting instances of osteoarthritis and rheumatoid arthritis.

Overall, all Res2Net variants demonstrated strong performance in arthritis classification, with each model exhibiting unique strengths and encountering varying degrees of difficulties. However, the proposed model (SE-Res2net50_14w_8s) emerged as the top performer, achieving the highest precision, recall, and accuracy scores. These results underscore the efficacy of the proposed model in arthritis classification and its potential for enhancing clinical diagnosis and treatment planning.

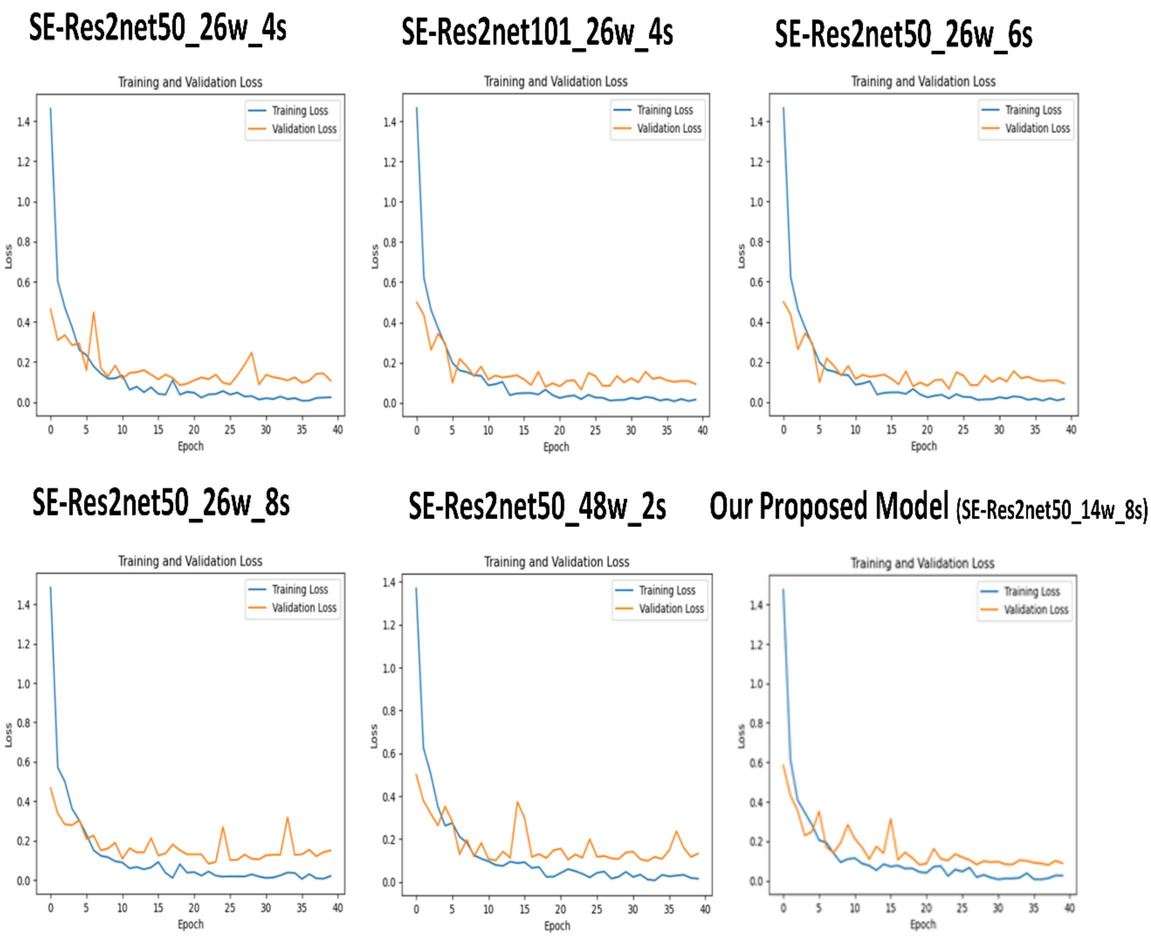


Figure 16 Res2Net models training and validation loss graph.

Figure 21 displays the training and validation loss graphs for each of the Res2Net models that we used in this research. These graphs demonstrate each model variant's strong training capabilities by showing the loss decreasing over time in successive epochs. The training and validation losses' smooth convergence suggests that the models can efficiently learn from the training set while making good generalizations to previously unseen validation data. In addition, the loss curves' lack of notable fluctuations or plateaus emphasizes how stable and reliable the training procedure is.

Table 7 Res2Net models result.

No	Res2Net Model Type	Accuracy	Precision	Recall	F1score
1.	SE-Res2Net50_48w_2s	0.9821	0.9836	0.982	0.982
2.	SE-Res2Net50_26w_8s	0.979	0.980	0.979	0.979
3.	SE-Res2Net50_26w_6s	0.984	0.984	0.984	0.984
4.	Res2Net101_26w_4s	0.984	0.984	0.984	0.984
5.	Res2Net50_26w_4s	0.979	0.979	0.979	0.979
6.	**Proposed Model (SE-Res2Net50_14w_8s)**	**0.995**	**0.995**	**0.995**	**0.995**

The results of the proposed Res2Net architecture showcase its remarkable performance across various configurations. Among the models evaluated, res2net50_14w_8s stands out with an impressive accuracy of 99.5%, demonstrating its robustness in accurately classifying arthritis cases. Additionally, res2net50_48w_2s exhibits exceptional accuracy at 98.21%, highlighting the efficacy of the wider base width and lower scale configuration. Notably, all models consistently achieve high precision, recall, and F1 scores, indicating their ability to effectively classify arthritis cases with high confidence and minimal misclassifications. These results underscore the potential of the proposed Res2Net architecture as a powerful tool for arthritis classification, paving the way for enhanced diagnostic accuracy and patient care in the medical domain.

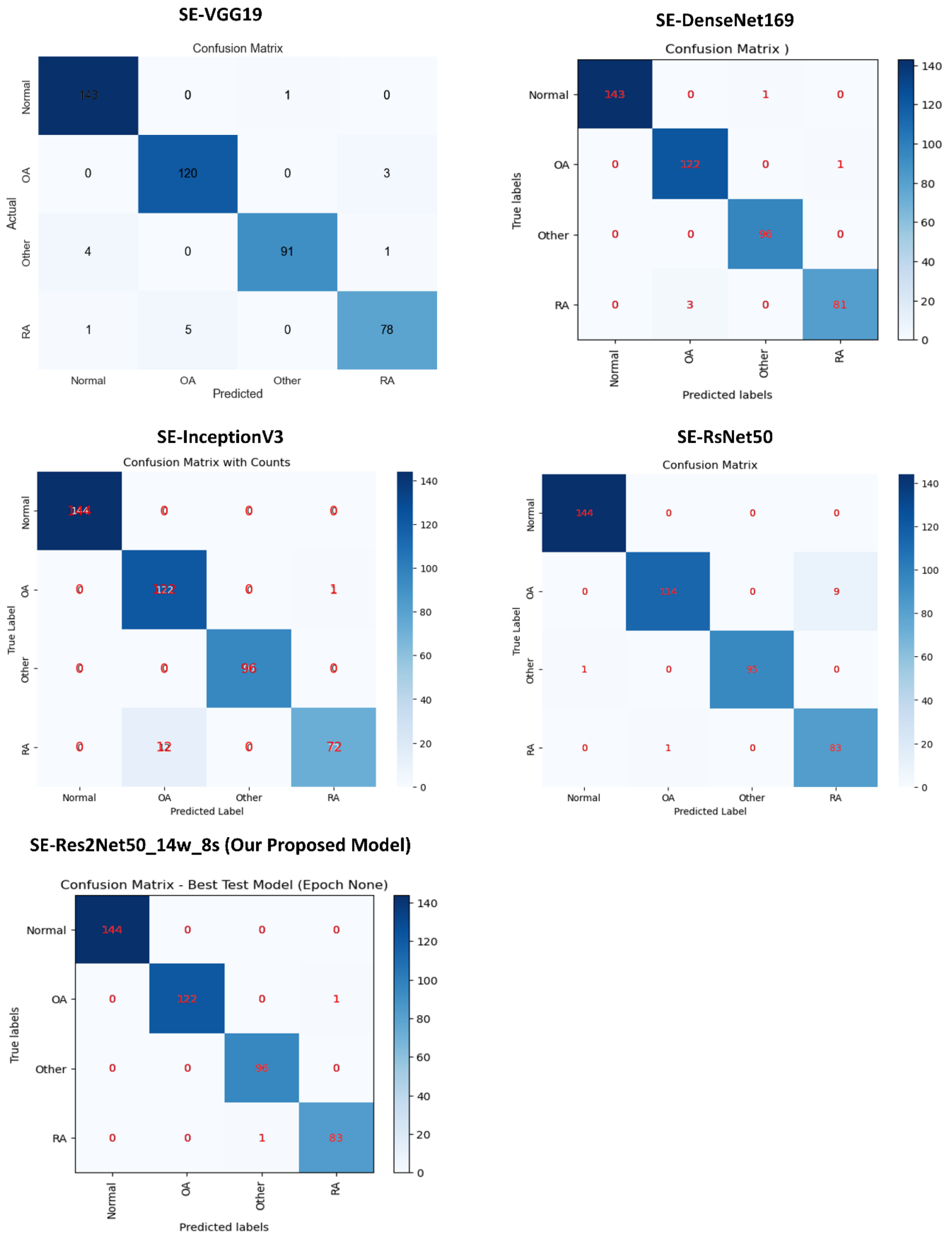


Figure 17 Confusion Matrix of Other Models with Proposed Model

Figure 22 shows the confusion matrices of multiple pretrained models in addition to the proposed model (SE-Res2net50_14w_8s), offering an understanding of how well each model performed in the classification of arthritis.

SE-VGG19 Model

The SE-VGG19 model performed wonderfully in classifying cases of arthritis, attaining high recall and precision scores in most of the classes. Four false negatives resulted from the model's inability to correctly detect instances of "Other" cases, though. When compared to other pretrained models, the SE-VGG19 model performed competitively overall.

SE-DenseNet169 Model

As with the SE-VGG19 model, the SE-DenseNet169 model demonstrated strong performance in classifying cases of arthritis, attaining excellent recall and precision scores in all classes. Three false negatives resulted from the model's inability to correctly detect instances of "Rheumatoid Arthritis" cases, proving that it could correctly detect different subtypes of arthritis.

SE-InceptionV3 Model

For most classes, the SE-InceptionV3 model showed outstanding precision and recall scores, demonstrating its effectiveness in classifying arthritis. But the model had trouble correctly detecting cases of rheumatoid arthritis, which led to 12 false negatives. Despite this, the SE InceptionV3 model performed competitively all around.

Model SE-ResNet50

In terms of arthritis classification, the SE-ResNet50 model performed satisfactorily, attaining good precision and recall scores for most classes. But the model had trouble distinguishing between osteoarthritis patients and rheumatoid arthritis cases; as a result, it produced nine false negatives. However, when compared to other pretrained models, the SE-ResNet50 model performed competitively.

Proposed Model (SE-Res2net50_14w_8s)

When it came to classifying cases of arthritis, the proposed model, SE-Res2net50_14w_8s, outperformed other pretrained models. Although almost all classes experienced almost no false negatives, a small number of "Other" cases were incorrectly identified, leading to one false negative. The proposed model performed the best overall, recording few false negatives and attaining excellent scores for accuracy, recall, and precision.

As a whole, all pretrained models performed well in classifying cases of arthritis; nevertheless, different strengths and levels of difficulty were encountered by each model. But compared to other pretrained models, the proposed model (SE-Res2net50_14w_8s) performed better, scoring the highest in accuracy, recall, and precision while generating the fewest false negatives. These outcomes highlight the proposed model's effectiveness in classifying arthritis.

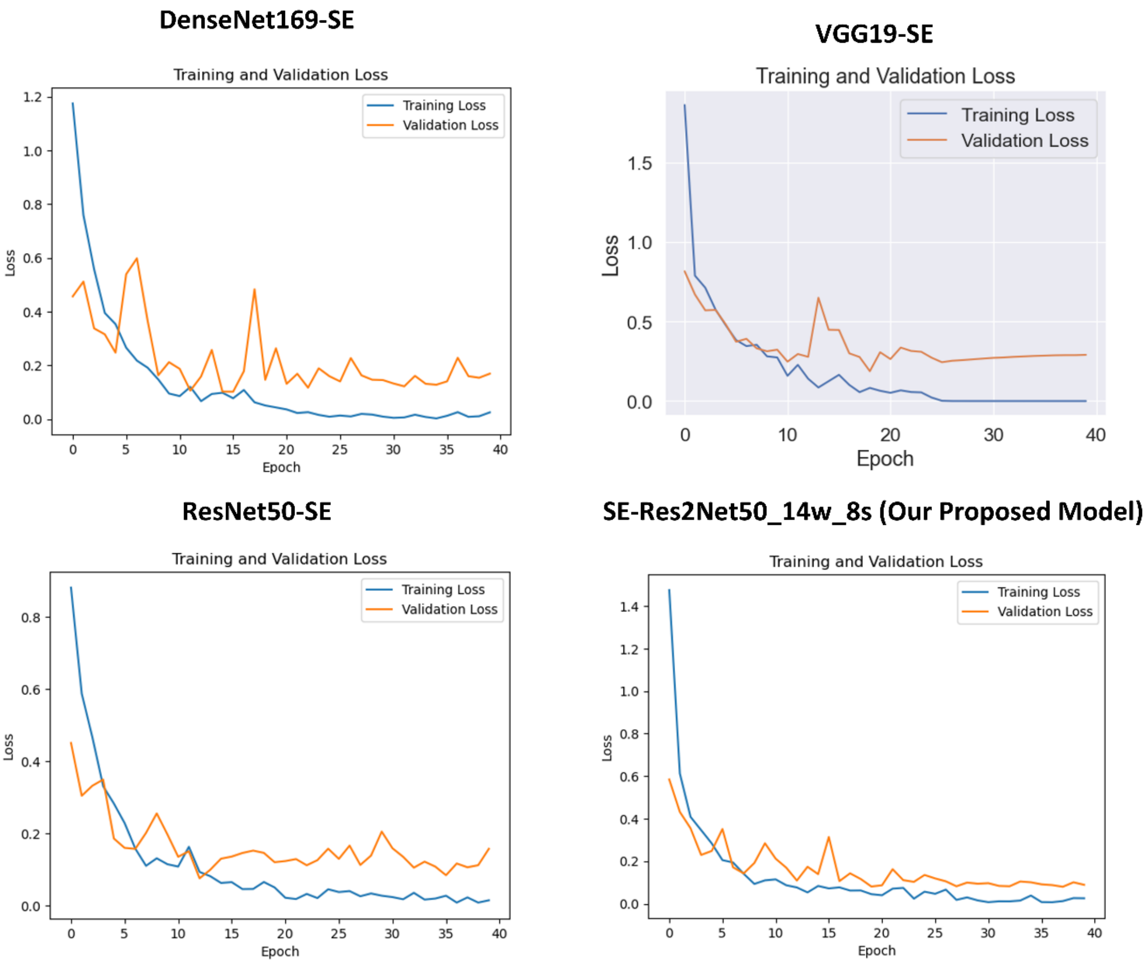


Figure 18 Training and Validation Loss Graph of Other Models with Proposed Model

Figure 23 shows the SE-VGG19, SE-DenseNet169, SE-ResNet50, and the proposed model's (SE-Res2net50_14w_8s) training and validation loss curves. When compare the proposed model to the other models, it is clear from analysis that it has better loss behavior. The proposed model's loss curves show a balanced convergence, characterized by minimum gaps between validation and training losses and negligible fluctuation. Strong training stability and efficient generalization to new data are indicated by this. In contrast, there are more obvious gaps between training and validation loss as well as noticeable oscillations in the loss curves for SE-VGG19, SE-DenseNet169, and SE-ResNet50, which may indicate overfitting or underfitting problems. The constant and smooth loss behavior of the proposed model, thus, highlights its improved training efficiency and generalization capacity and confirms its dominance in the context of arthritis classification.

Table 8 Other Pretrained Models with our Proposed Model Result

No	Pretrained Models	Accuracy	Precision	Recall	F1score
1.	SE-VGG19	0.966	0.966	0.966	0.966
2.	SE-Densenet169	0.988	0.988	0.988	0.988
3.	SE-InceptionV3	0.970	0.974	0.962	0.966
4.	SE-Resnet50	0.975	0.977	0.975	0.975
5.	**Proposed Model (SE-Res2Net50_14w_8s)**	**0.995**	**0.995**	**0.995**	**0.995**

The findings displayed in Table 8 demonstrate that all pretrained models—SE-VGG19, SE-Densenet169, SE-InceptionV3, and SE-Resnet50—perform admirably in classification of arthritis, with accuracy scores ranging from 0.966 to 0.975. These models likewise show excellent recall, F1-score, and precision metrics; the ranges of their respective values are 0.966 to 0.977 for precision, 0.962 to 0.988 for recall, and 0.966 to 0.988 for F1-score.

Compared to other pretrained models, the proposed model SE-Res2Net50_14w_8s has the highest accuracy score of 0.995. Furthermore, it has superior recall, F1-score metrics, and precision, all at 0.995. This demonstrates the remarkable ability of the proposed approach to correctly categorize subtypes of arthritis.

The results highlight the usefulness of pretrained models in arthritis classification tasks overall, and the proposed model (SE-Res2Net50_14w_8s) is the best performer, outperforming all other assessed metrics. These results highlight how the proposed strategy can improve clinical diagnosis and treatment planning for patients with arthritis.

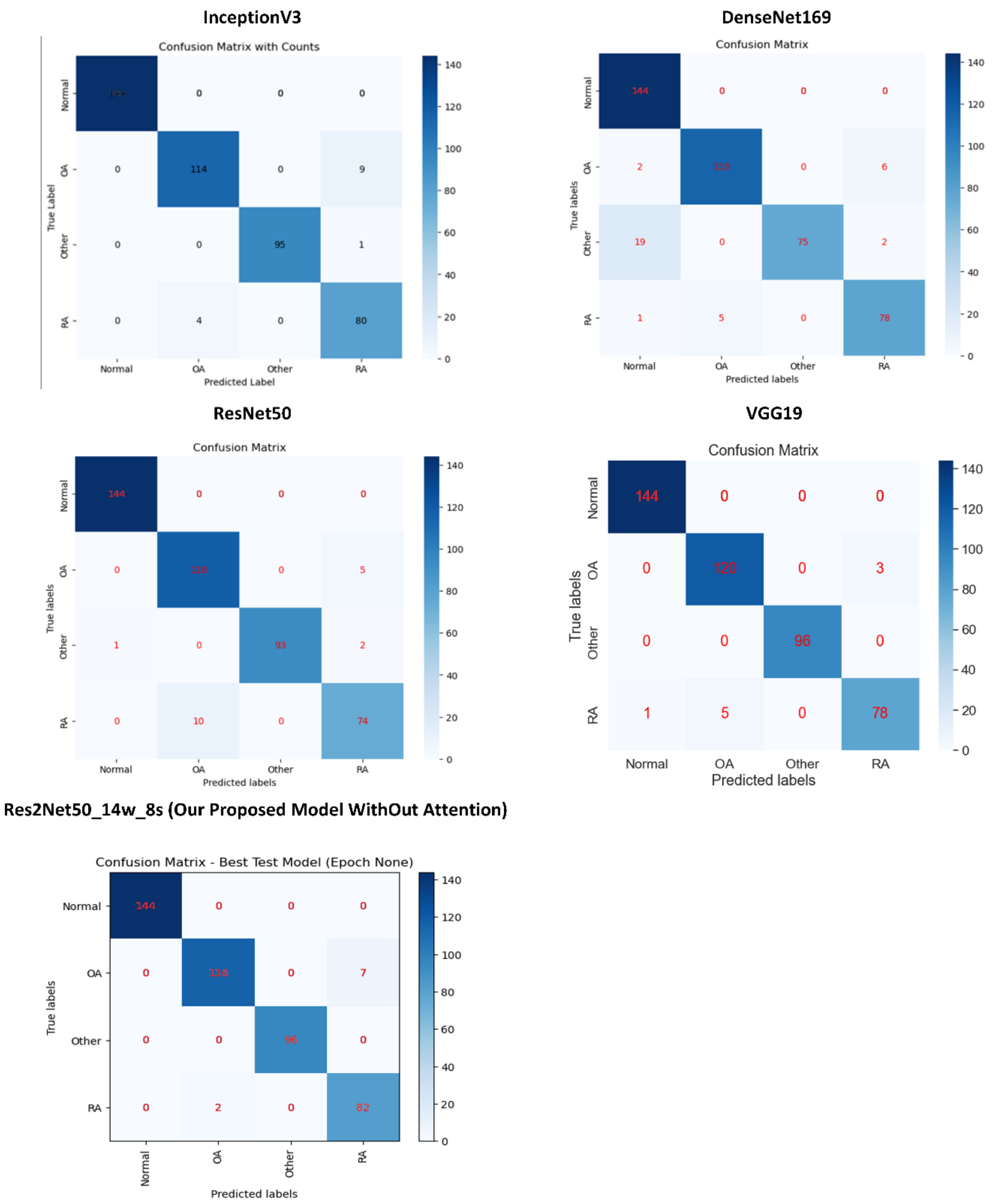


Figure 19 Confusion Matrix of Other Models with Proposed Model without SE

In Figure 24, present the confusion matrices for several other pretrained models, including InceptionV3, DenseNet169, ResNet50, and VGG19, alongside the proposed model (SE-Res2net50_14w_8s), without attention mechanisms. These matrices offer insights into the performance of each model variant in arthritis classification.

InceptionV3 Model

The InceptionV3 model demonstrates robust performance across most arthritis subtypes, achieving high precision and recall scores. However, the model encounters challenges in correctly detecting instances of osteoarthritis (OA), with 9 false negatives recorded in this class. Additionally, 4 false negatives are observed in the rheumatoid arthritis (RA) class.

Model DenseNet169

Like InceptionV3, the DenseNet169 model performs admirably when it comes to classifying arthritis. But the model has trouble correctly distinguishing "Other" case situations, which leads to 19 false negatives. Six false negatives are also documented in the OA class.

Model ResNet50

In terms of arthritis classification, the ResNet50 model performs satisfactorily, attaining good precision and recall scores in most classes. Nevertheless, the model has trouble accurately detecting RA cases; in this class, 10 false negatives were noted.

Model VGG19

With respect to arthritis classification, the VGG19 model performs competitively, attaining good precision and recall scores. Nevertheless, there are three false negatives in this class, indicating that the model has trouble accurately recognizing cases with OA. In the RA class, five false negatives are also noted.

Proposed Model (SE-Res2net50_14w_8s)

When compared to other pretrained models, the proposed model (SE-Res2net50_14w_8s) performs better because it does not have attention mechanisms. In every class, it earns high recall and precision scores with very few incorrect negatives. Particularly, there are only 7 false negatives in the OA class and 2 in the RA class. This demonstrates how well our proposed model performs in reliably detecting different forms of arthritis.

All things considered, the comparison highlights the advantages and disadvantages of each model variant in the classification tasks related to arthritis, with our proposed model.

Table 9 Other Models with Proposed Model without Attention (SE) Result

No	Pretrained Models	Accuracy	Precision	Recall	F1score
1.	VGG19	**0.97986**	0.97901	0.97694	0.97743

2.	Densenet169	0.9217	0.928	0.921	0.920
3.	InceptionV3	0.968	0.969	0.968	0.968
4.	Resnet50	0.923	0.923	0.923	0.923
5.	**Proposed Model Without Attention (Res2Net50_14w_8s)**	**0.97986**	**0.98055**	**0.97986**	**0.97994**

Based on the results presented in Table 9, it's evident that all pretrained models, including VGG19, DenseNet169, InceptionV3, and ResNet50, demonstrate strong performance in arthritis classification tasks, with accuracy scores ranging from 0.9217 to 0.97986. These models also exhibit commendable precision, recall, and F1-score metrics, reflecting their effectiveness in accurately classifying arthritis subtypes.

Notably, the proposed model without attention (Res2Net50_14w_8s) achieves an accuracy of 0.97986, precision of 0.98055, recall of 0.97986, and F1-score of 0.97994, surpassing the performance of other pretrained models. This highlights the effectiveness of the proposed model in accurately classifying arthritis subtypes, with high precision and recall while maintaining strong overall accuracy.

Overall, the results underscore the competitive performance of pretrained models in arthritis classification tasks, with the proposed model without attention emerging as a top performer. These findings suggest the potential of the proposed model for enhancing clinical diagnosis and treatment planning in arthritis patients.

6.3 Research Questions Discussion

RQ1. What impact does multi-scale feature extraction have on the classification of arthritis?

By gathering data at different granularities, multi-scale feature extraction greatly improves arthritis classification. Models utilizing multi-scale feature extraction perform better at correctly detecting and classifying arthritic subtypes, as shown by the confusion matrices shown in Figures 20 and 22. As seen by the near-zero false negatives achieved by the proposed multi-scale feature model SE-Res2net50_14w_8s in Figure 20, it is clear that this model is capable of capturing both fine-grained local details and global contextual information that is pertinent to the classification of arthritis. Furthermore, Figure 20 illustrates how multi-scale feature extraction can reduce misclassifications and increase the accuracy of arthritis classification by showing that the model with attention and multi-scale features has fewer false negatives than models without these features.

RQ2. To what extent does attention mechanism integration improve arthritis disease classification?

Integration of the attention mechanism greatly improves the detection of arthritis diseases by allowing the model to concentrate on specific regions within the input data. Comparing the confusion matrices in Figure 22 and Figure 24, the attention-based model shows less false negatives, especially in important classes like RA and OA in osteoarthritis. This decrease in false negatives shows that the model can focus more attention on arthritic feature discrimination. Additionally, the training and validation loss curves in Figure 21 and Figure 23 show that the model with attention mechanisms converges more steadily than the model without attention. This stability suggests that attention mechanisms aid the model in more accurately detecting arthritis-related traits, which enhances the model's capacity to detect the disease.

RQ3. Which model performs best for arthritis detection and classification?

Out of all the models analyzed for the detection and classification of arthritis, the proposed model with attention (SE-Res2Net50_14w_8s) shows the best results. As shown in Table 8, when compared to other pretrained models such as SE-VGG19, SE-DenseNet169, SE-InceptionV3, and SE-ResNet50, this model achieves the top metrics for accuracy, precision, recall, and F1-score. Moreover, as Figures 20 and 22 show, the proposed model performs better in terms of precisely detecting and classifying arthritic subtypes, with few false negatives and strong feature representation abilities. As a result, the proposed model sticks out as the best option for classifying and detecting arthritis, demonstrating the significance of multi-scale feature extraction and attention mechanisms in raising the accuracy of disease detection.

CHAPTER SEVEN

7. CONCLUSION AND FUTURE WORK

7.1 Conclusion

In this study, investigated the efficacy of different deep learning models for arthritis detection and classification. Through experimentation and analysis, addressed three key research questions: the impact of attention mechanism integration, the significance of multi-scale feature extraction, and the performance comparison of different model architectures.

The findings highlight the critical role of attention mechanisms in enhancing arthritis disease detection, as evidenced by the improved stability in training convergence and the reduction in false negatives observed in models with attention. Furthermore, demonstrate the importance of multi-scale feature extraction in arthritis classification, with models leveraging this technique achieving superior performance in capturing both global contextual information and fine-grained local details relevant to arthritis subtypes.

Among the evaluated models, the proposed model with attention, SE-Res2Net50_14w_8s, emerges as the top performer, showcasing the highest accuracy, precision, recall, and F1-score metrics. This model demonstrates exceptional capabilities in accurately detecting and classifying arthritis subtypes, underscoring the significance of attention mechanisms and multi-scale feature extraction in improving disease detection accuracy.

Overall, this research underscores the potential of deep learning models in enhancing arthritis diagnosis and treatment planning. By leveraging attention mechanisms and multi-scale feature extraction techniques, we can develop more robust and effective solutions for arthritis detection, ultimately contributing to improved patient care and outcomes in clinical practice.

7.2 Contributions of the Study

- ✓ To improve arthritis detection and classification, an attention deep learning model (SE-Res2Net50_14w_8s) was developed that integrates attention mechanisms and multiscale feature extraction.
- ✓ Outperforming other pretrained models such as SE-VGG19, SE-DenseNet169, SE-InceptionV3, and SE-ResNet50, with good accuracy (0.995), precision, recall, and F1-score as well.

- ✓ Address the issue of low availability of radiologists in developing nations by utilizing an X-ray image collection from a local hospital in Ethiopia (2018–2024).
- ✓ Indicating that fine-grained disease features can be captured by multiscale feature extraction to improve arthritis classification.
- ✓ Demonstrating how the integration of attention mechanisms enhances feature representation and reduces false negatives by concentrating attention on relevant regions inside X-ray images.

7.3 Future Work

In future research endeavors, it is imperative to expand the scope of datasets by acquiring larger and more varied collections, encompassing a broader range of arthritis diseases and subtypes. This expansion would provide a more foundation for evaluating the robustness and generalization capabilities of deep learning models. Moreover, integrating clinical data, such as patient demographics, medical history, and laboratory findings, holds promise in enhancing the contextual understanding of arthritis cases and improving classification accuracy. By developing hybrid models that integrates image-based features with clinical data, researchers can unlock new avenues for refining disease detection and classification. Furthermore, efforts towards enhancing the interpretability of deep learning models would foster trust and understanding among clinicians, thereby facilitating their adoption in clinical practice. Ultimately, conducting rigorous clinical validation studies and collaborating with healthcare institutions and clinicians would be pivotal in translating these advancements into meaningful improvements in arthritis diagnosis and patient care.

Special Acknowledgements

This study was funded by Adama Science and Technology University with grant number:

ASTU/SM-R/1057/24

Adama, Ethiopia

REFERENCES

Alexopoulos, A., Hirvasniemi, J., & Tümer, N. (2022). Early detection of knee osteoarthritis using deep learning on knee magnetic resonance images. *arXiv preprint arXiv:2209.01192*.

Aysa Z, Ablimit M, Hamdulla A. Multi-Scale Feature Learning for Language Identification of Overlapped Speech. *Applied Sciences*. 2023; 13(7):4235. https://doi.org/10.3390/app13074235

Berihun Molla, Teklu Urgessa (Associate Professor), & Worku B. "Developing Knee Arthritis Classification Model from Knee X-Ray Image Using Deep Learning Approach." 2021.

Collaborators, R. A. (2023). *Global , regional , and national burden of rheumatoid arthritis , 1990 – 2020 , and projections to 2050 : a systematic analysis of the Global Burden of Disease Study 2021*. *5*(October).

Gornale, S. S., Patravali, P. U., & Hiremath, P. S. (2020). Automatic Detection and Classification of Knee Osteoarthritis Using Hu's Invariant Moments. *Frontiers in Robotics and AI*, *7*. https://doi.org/10.3389/frobt.2020.591827

J. Hu, L. Shen and G. Sun, "Squeeze-and-Excitation Networks," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 2018, pp. 7132-7141, doi: 10.1109/CVPR.2018.00745.

Kourilovitch, M., Galarza-maldonado, C., & Ortiz-prado, E. (2014). *Author ' s personal copy Diagnosis and classi fi cation of rheumatoid arthritis*. 1–5.

Li, Y., & Zhao, L. (2022). Application of Machine Learning in Rheumatic Immune Diseases. *Journal of Healthcare Engineering*, *2022*, 708–722. https://doi.org/10.1155/2022/9273641

Liu, B., Luo, J., & 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), 457–466. https://doi.org/10.1007/s11548-019-02096-9

Ma, Y., Pan, I., Kim, S. Y., Wieschhoff, G. G., Andriole, K. P., & Mandell, J. C. (2023). Deep learning discrimination of rheumatoid arthritis from osteoarthritis on hand radiography. *Skeletal Radiology*, *53*(2), 377–383. https://doi.org/10.1007/s00256-023-

04408-2

Murakami, S., Hatano, K., Tan, J., Kim, H., & Aoki, T. (2018). Automatic identification of bone erosions in rheumatoid arthritis from hand radiographs based on deep convolutional neural network. *Multimedia Tools and Applications*, *77*(9), 10921–10937. https://doi.org/10.1007/s11042-017-5449-4

Olsson, S., Akbarian, E., Lind, A., Razavian, A. S., & 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). https://doi.org/10.1186/s12891-021-04722-7

Roques, M., Tanchoux, F., Calvière, L., Cuinat, L., Lubrano, V., Uro-Coste, E., Cognard, C., Larrue, V., & Bonneville, F. (2014). MRI with DWI helps in depicting rheumatoid meningitis. In *Journal of Neuroradiology* (Vol. 41, Issue 4, pp. 275–277). Elsevier Masson SAS. https://doi.org/10.1016/j.neurad.2013.10.005

S. Gornale, S., U. Patravali, P., S. Marathe, K., & S. Hiremath, P. (2017). Determination of Osteoarthritis Using Histogram of Oriented Gradients and Multiclass SVM. *International Journal of Image, Graphics and Signal Processing*, *9*(12), 41–49. https://doi.org/10.5815/ijigsp.2017.12.05

Singh, S., Maxwell, J., Baker, J. A., Nicholas, J. L., & Lo, J. Y. (2011). Computer-aided classification of breast masses: Performance and interobserver variability of expert radiologists versus residents. *Radiology*, *258*(1), 73–80. https://doi.org/10.1148/radiol.10081308

S. -H. Gao, M. -M. Cheng, K. Zhao, X. -Y. Zhang, M. -H. Yang and P. Torr, "Res2Net: A New Multi-Scale Backbone Architecture," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 43, no. 2, pp. 652-662, 1 Feb. 2021, doi: 10.1109/TPAMI.2019.2938758.

Swiecicki, A., Li, N., O'Donnell, J., Said, N., Yang, J., Mather, R. C., Jiranek, W. A., & 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*. https://doi.org/10.1016/j.compbiomed.2021.104334

Tariq, T., Suhail, Z., & Nawaz, Z. (2023). Knee Osteoarthritis Detection and Classification Using X-Rays. *IEEE Access*, *11*, 48292–48303. https://doi.org/10.1109/ACCESS.2023.3276810

Thomson, J., O'Neill, T., Felson, D., & Cootes, T. (2015). Automated shape and texture analysis for detection of osteoarthritis from radiographs of the knee. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *9350*, 127–134. https://doi.org/10.1007/978-3-319-24571-3_16

Tiulpin, A., & Saarakkala, S. (2020). Automatic grading of individual knee osteoarthritis features in plain radiographs using deep convolutional neural networks. *Diagnostics*, *10*(11). https://doi.org/10.3390/diagnostics10110932

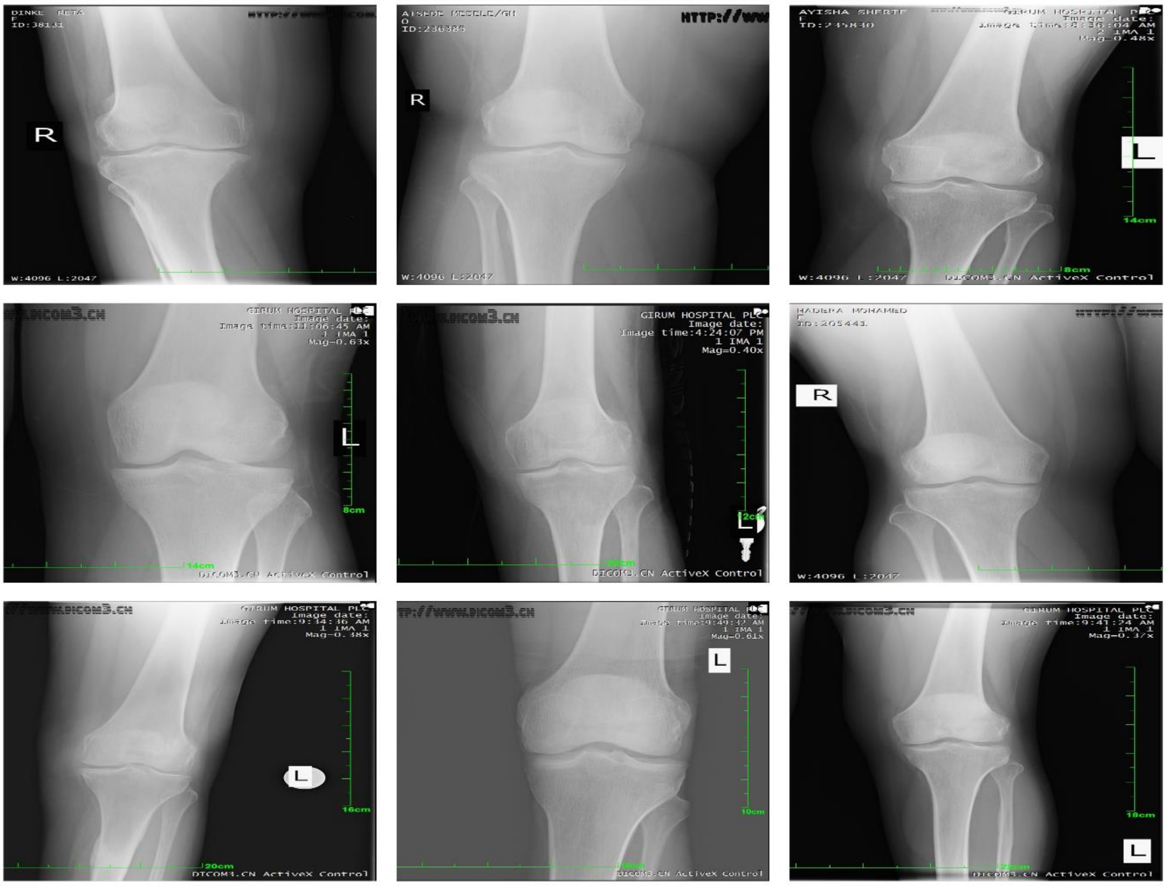
Üreten, K., Erbay, H., & Maraş, H. H. (2020). Detection of rheumatoid arthritis from hand radiographs using a convolutional neural network. *Clinical Rheumatology*, *39*(4), 969–974. https://doi.org/10.1007/s10067-019-04487-4

Üreten, K., & Maraş, H. H. (2022). Automated Classification of Rheumatoid Arthritis, Osteoarthritis, and Normal Hand Radiographs with Deep Learning Methods. *Journal of Digital Imaging*, *35*(2), 193–199. https://doi.org/10.1007/s10278-021-00564-w

Zhang, B., Tan, J., Cho, K., Chang, G., & Deniz, C. M. (2020). Attention-based CNN for KL Grade Classification: Data from the Osteoarthritis Initiative. *Proceedings - International Symposium on Biomedical Imaging*, *2020-April*, 731–735. https://doi.org/10.1109/ISBI45749.2020.9098456

APENDIX

Appendix A: Sample Knee X-ray Images

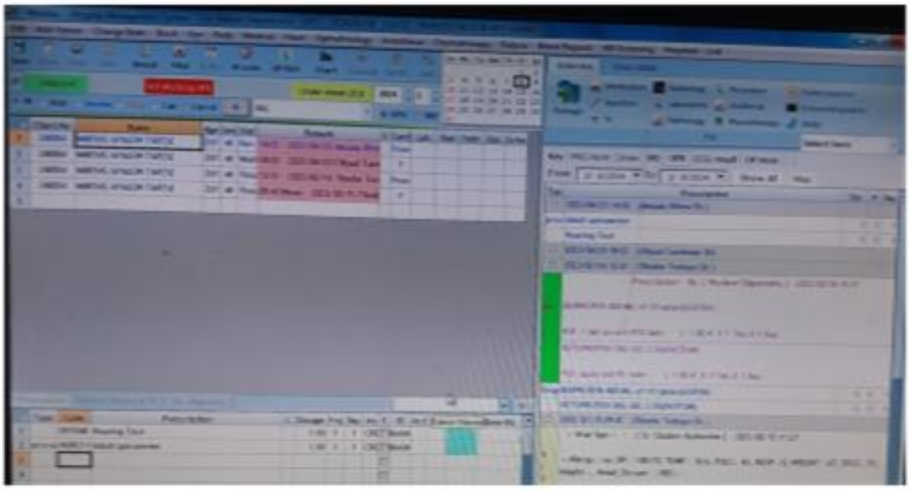


Appendix B: Last Epochs Progress Report for the Proposed Model

```
Epoch 37/40: 100%|████████████████████████████████████████| 334/334 [14:53<00:00,  2.67s/it]
Epoch 37/40, Training Loss: 0.012229774775504382, Accuracy: 0.9969992498124531, Precision: 0.9970144386729574, Rec
all: 0.9969992498124531, F1 Score: 0.9969948369377922
Evaluation: 100%|█████████████████████████████████████████| 111/111 [01:35<00:00,  1.16it/s]
Epoch 37/40, Validation Loss: 0.0796111168607027, Accuracy: 0.9819819819819819, Precision: 0.9822997416020671, Rec
all: 0.9819819819819819, F1 Score: 0.982062150064879
Saving checkpoint: C:/Users/Chalie.lijalem/Research/checkpoints/res2net50_14w_8s-6527dddc/epoch_37.pt
Epoch 38/40: 100%|████████████████████████████████████████| 334/334 [14:02<00:00,  2.52s/it]
Epoch 38/40, Training Loss: 0.026166465978815157, Accuracy: 0.9924981245311327, Precision: 0.9926093884635463, Rec
all: 0.9924981245311327, F1 Score: 0.9925013301529686
Evaluation: 100%|█████████████████████████████████████████| 111/111 [01:28<00:00,  1.25it/s]
Epoch 38/40, Validation Loss: 0.10048639347797353, Accuracy: 0.9752252252252253, Precision: 0.9756309361687269, Re
call: 0.9752252252252253, F1 Score: 0.9753326220759312
Saving checkpoint: C:/Users/Chalie.lijalem/Research/checkpoints/res2net50_14w_8s-6527dddc/epoch_38.pt
Epoch 39/40: 100%|████████████████████████████████████████| 334/334 [12:33<00:00,  2.25s/it]
Epoch 39/40, Training Loss: 0.025389697782091644, Accuracy: 0.9954988747186797, Precision: 0.995518941990316, Reca
ll: 0.9954988747186797, F1 Score: 0.9954994164720409
Evaluation: 100%|█████████████████████████████████████████| 111/111 [01:24<00:00,  1.32it/s]
Epoch 39/40, Validation Loss: 0.08874720364190843, Accuracy: 0.9752252252252253, Precision: 0.9762474228766362, Re
call: 0.9752252252252253, F1 Score: 0.9754021040201761
Saving checkpoint: C:/Users/Chalie.lijalem/Research/checkpoints/res2net50_14w_8s-6527dddc/epoch_39.pt
Evaluation: 100%|█████████████████████████████████████████| 112/112 [01:30<00:00,  1.24it/s]
Test Loss: 0.023116238376025132, Accuracy: 0.9955257270693513, Precision: 0.9955487903318804, Recall: 0.9955257270
693513, F1 Score: 0.9955269571750528
```

Appendix C: Patient Data Collection from OPD

Appendix D: Retrieval from MedAxs Hospital Management System



Appendix E: Sample Codes for Data Preprocessing

```
import os

import cv2

import numpy as np


# Function to perform data augmentation

def augment_data(input_folder, output_folder):

    # Create output folder if it doesn't exist

    if not os.path.exists(output_folder):

        os.makedirs(output_folder)
```

```
    # Loop through each image in the input folder
    for filename in os.listdir(input_folder):
        if filename.endswith('.png') or filename.endswith('.jpg'):
            # Read the image
            image = cv2.imread(os.path.join(input_folder, filename))
            # Save original image
            cv2.imwrite(os.path.join(output_folder, filename), image)
            # Apply augmentation and save augmented images
            augment_and_save(image, output_folder, filename)

# Function to apply augmentation and save augmented images
def augment_and_save(image, output_folder, filename):
    # Augmentation 1: Random Brightness and Contrast Adjustment
    brightness_contrast_adjusted_image = adjust_brightness_contrast(image)
    brightness_contrast_filename = f"{os.path.splitext(filename)[0]}_brightness_contrast.png"
    cv2.imwrite(os.path.join(output_folder, brightness_contrast_filename),
brightness_contrast_adjusted_image)

    # Augmentation 2: Gaussian Blur
    blurred_image = apply_gaussian_blur(image)
    blurred_filename = f"{os.path.splitext(filename)[0]}_blurred.png"
    cv2.imwrite(os.path.join(output_folder, blurred_filename), blurred_image)

    # Augmentation 3: Random Rotation
    rotated_image = rotate_image(image)
    rotated_filename = f"{os.path.splitext(filename)[0]}_rotated.png"
    cv2.imwrite(os.path.join(output_folder, rotated_filename), rotated_image)

# Function to adjust brightness and contrast randomly
def adjust_brightness_contrast(image):
```

```
    # Randomly adjust brightness and contrast
    alpha = np.random.uniform(0.8, 1.2)  # Scale factor for contrast (0.8 to 1.2)
    beta = np.random.randint(-20, 20)    # Brightness offset (-20 to 20)
    adjusted_image = cv2.convertScaleAbs(image, alpha=alpha, beta=beta)
    return adjusted_image

# Function to apply Gaussian blur
def apply_gaussian_blur(image):
    # Apply Gaussian blur with random kernel size
    kernel_size = (5, 5)  # Random kernel size
    blurred_image = cv2.GaussianBlur(image, kernel_size, 0)
    return blurred_image

# Function to rotate the image
def rotate_image(image):
    # Randomly rotate the image between -15 and 15 degrees
    angle = np.random.randint(-15, 15)
    height, width = image.shape[:2]
    rotation_matrix = cv2.getRotationMatrix2D((width / 2, height / 2), angle, 1)
    rotated_image = cv2.warpAffine(image, rotation_matrix, (width, height))
    return rotated_image

# Example usage
input_folder = r"C:\Users\Chalie.lijalem\Research\data-after-image-name-updated\RA"
output_folder = r"C:\Users\Chalie.lijalem\Research\data-after-data-augmenation_final\RA"

augment_data(input_folder, output_folder)
```

Appendix F: Sample code for Proposed Model

```
# Define the SEModule class
```

```
class SEModule(nn.Module):
    def __init__(self, channels, reduction=16):
        super(SEModule, self).__init__()
        self.avg_pool = nn.AdaptiveAvgPool2d(1)
        self.fc1 = nn.Conv2d(channels, channels // reduction, kernel_size=1, padding=0)
        self.relu = nn.ReLU(inplace=True)
        self.fc2 = nn.Conv2d(channels // reduction, channels, kernel_size=1, padding=0)
        self.sigmoid = nn.Sigmoid()

    def forward(self, input):
        x = self.avg_pool(input)
        x = self.fc1(x)
        x = self.relu(x)
        x = self.fc2(x)
        x = self.sigmoid(x)
        return input * x

# Define the Bottle2neck class
class Bottle2neck(nn.Module):
    expansion = 4

    def __init__(self, inplanes, planes, stride=1, downsample=None, baseWidth=26, scale=4,
stype='normal', se=True):
        super(Bottle2neck, self).__init__()

        width = int(math.floor(planes * (baseWidth/64.0)))
        self.conv1 = nn.Conv2d(inplanes, width*scale, kernel_size=1, bias=False)
        self.bn1 = nn.BatchNorm2d(width*scale)

        if scale == 1:
          self.nums = 1
```

```
        else:
          self.nums = scale - 1
        if stype == 'stage':
            self.pool = nn.AvgPool2d(kernel_size=3, stride=stride, padding=1)
        convs = []
        bns = []
        for i in range(self.nums):
          convs.append(nn.Conv2d(width, width, kernel_size=3, stride=stride, padding=1, bias=False))
          bns.append(nn.BatchNorm2d(width))
        self.convs = nn.ModuleList(convs)
        self.bns = nn.ModuleList(bns)

        self.conv3 = nn.Conv2d(width*scale, planes * self.expansion, kernel_size=1, bias=False)
        self.bn3 = nn.BatchNorm2d(planes * self.expansion)

        self.relu = nn.ReLU(inplace=True)
        self.se = SEModule(planes * self.expansion) if se else None
        self.downsample = downsample
        self.stype = stype
        self.scale = scale
        self.width  = width

    def forward(self, x):
        residual = x

        out = self.conv1(x)
        out = self.bn1(out)
        out = self.relu(out)

        spx = torch.split(out, self.width, 1)
        for i in range(self.nums):
```

```
        if i == 0 or self.stype == 'stage':
          sp = spx[i]
        else:
          sp = sp + spx[i]
        sp = self.convs[i](sp)
        sp = self.relu(self.bns[i](sp))
        if i == 0:
          out = sp
        else:
          out = torch.cat((out, sp), 1)
      if self.scale != 1 and self.stype == 'normal':
        out = torch.cat((out, spx[self.nums]), 1)
      elif self.scale != 1 and self.stype == 'stage':
        out = torch.cat((out, self.pool(spx[self.nums])), 1)

      out = self.conv3(out)
      out = self.bn3(out)

      if self.se is not None:
          out = self.se(out)

      if self.downsample is not None:
          residual = self.downsample(x)

      out += residual
      out = self.relu(out)

      return out

class Res2Net(nn.Module):
```

```
def __init__(self, block, layers, baseWidth = 26, scale = 4, num_classes=1000,se=False):
    self.inplanes = 64
    super(Res2Net, self).__init__()
    self.baseWidth = baseWidth
    self.scale = scale
    self.conv1 = nn.Conv2d(3, 64, kernel_size=7, stride=2, padding=3,
                           bias=False)
    self.bn1 = nn.BatchNorm2d(64)
    self.relu = nn.ReLU(inplace=True)
    self.maxpool = nn.MaxPool2d(kernel_size=3, stride=2, padding=1)
    self.layer1 = self._make_layer(block, 64, layers[0])
    self.layer2 = self._make_layer(block, 128, layers[1], stride=2)
    self.layer3 = self._make_layer(block, 256, layers[2], stride=2)
    self.layer4 = self._make_layer(block, 512, layers[3], stride=2)
    self.avgpool = nn.AdaptiveAvgPool2d(1)
    self.fc = nn.Linear(512 * block.expansion, num_classes)

    for m in self.modules():
        if isinstance(m, nn.Conv2d):
            nn.init.kaiming_normal_(m.weight, mode='fan_out', nonlinearity='relu')
        elif isinstance(m, nn.BatchNorm2d):
            nn.init.constant_(m.weight, 1)
            nn.init.constant_(m.bias, 0)

def _make_layer(self, block, planes, blocks, stride=1):
    downsample = None
    if stride != 1 or self.inplanes != planes * block.expansion:
        downsample = nn.Sequential(
            nn.Conv2d(self.inplanes, planes * block.expansion,
                      kernel_size=1, stride=stride, bias=False),
            nn.BatchNorm2d(planes * block.expansion),
```

```
        )

        layers = []
        layers.append(block(self.inplanes, planes, stride, downsample=downsample,
                        stype='stage', baseWidth = self.baseWidth, scale=self.scale))
        self.inplanes = planes * block.expansion
        for i in range(1, blocks):
            layers.append(block(self.inplanes, planes, baseWidth = self.baseWidth, scale=self.scale))

        return nn.Sequential(*layers)

    def forward(self, x):
        x = self.conv1(x)
        x = self.bn1(x)
        x = self.relu(x)
        x = self.maxpool(x)

        x = self.layer1(x)
        x = self.layer2(x)
        x = self.layer3(x)
        x = self.layer4(x)

        x = self.avgpool(x)
        x = x.view(x.size(0), -1)
        x = self.fc(x)

        return x
def res2net50(pretrained=False, model_path='res2net50_26w_4s.pth', **kwargs):
    model = Res2Net(Bottle2neck, [3, 4, 6, 3], baseWidth=26, scale=4, **kwargs)
    if pretrained:
        # Load pre-trained weights
```

```
    pretrained_dict = torch.load(model_path, map_location=torch.device('cpu'))
    # Filter out keys related to SEModule layers
    pretrained_dict = {k: v for k, v in pretrained_dict.items() if 'se' not in k}
    # Load filtered weights
    model.load_state_dict(pretrained_dict, strict=False)
return model
```