Revolutionizing ACL Tear Detection Using Deep Learning Models

By
Tarian Ahamed Heem
ID:212-15-4232

FINAL YEAR DESIGN PROJECT REPORT

This Report Presented in Partial Fulfillment of the Requirements for the **Degree of Bachelor of Science in Computer Science and Engineering**

Supervised by

Ms. Tasfia Anika Bushra
Lecturer
Department of Computer Science and Engineering
Daffodil International University

Co-Supervised by

Mr. Mahimul Islam Nadim
Lecturer
Department of Computer Science and Engineering
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

Dhaka, Bangladesh
September 17, 2025

APPROVAL

This Project titled **"Revolutionizing ACL Tear Detection Using Deep Learning Models"**, submitted by Tarian Ahamed Heem, 212-15-4232 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on **16 September, 2025**.

BOARD OF EXAMINERS

-- 17.09.25

Dr. Fernaz Narin Nur — **Chairman**
Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Dr. Abdus Sattar — **Internal Examiner**
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

-- 17.09.25

Mr. Mohammad Jahangir Alam — **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

-- 17.09.25

Dr. Sajeeb Saha — **External Examiner**
Associate Professor
Department of Computer Science and Engineering
Jagannath University

DECLARATION

I hereby declare that this project has been done by us under the supervision of Ms Tasfia Anika Bushra, Lecturer, Department of Computer Science and Engineering, Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for the award of any degree or diploma.

Supervised by:

Bushra 16.09.25

Ms Tasfia Anika Bushra
Designation: Lecturer
Department of Computer Science and Engineering
Daffodil International University

Co-Supervised by:

Mr. Mahimul Islam Nadim
Designation: Lecturer
Department of Computer Science and Engineering
Daffodil International University

Submitted by:

Tarian

Tarian Ahamed Heem

Student ID: 212-15-4232
Department of Computer Science and Engineering
Daffodil International University

ACKNOWLEDGEMENTS

This work would not have been possible without the support and contributions of many individuals over the past two semesters. I am deeply grateful to everyone who has assisted me in one way or another.

First, I express my heartfelt thanks and gratefulness to the almighty for His divine blessing making it possible for me to complete the Final Year Design Project(FYDP) successfully.

I am grateful and wish my profound indebtedness to Ms Tasfia Anika Bushra, Lecturer, Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh. Deep knowledge and keen interest of my supervisor in the field of Deep Learning to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts, and correcting them at all stages have made it possible to complete this project.

I would like to express my heartfelt gratitude to the Head of the Department of Computer Science and Engineering, for his kind help in finishing our project and also to other faculty members and the staff of the Department of Computer Science and Engineering, Daffodil International University.

I would like to thank my entire course-mates at Daffodil International University, who took part in this discussion while completing the coursework.

Finally, I must acknowledge with due respect the constant support and patience of our parents.

ABSTRACT

Anterior cruciate ligament (ACL) injuries are considered to be one of the most widespread and acute knee disorders that can require appropriate and timely diagnosis because this can be one of the conditions that will be used to make decisions regarding the treatment. In manual inspection of magnetic resonance imaging (MRI), time-consuming mistakes are involved in inter-observer errors. To address this problem, we propose a hybrid deep learning model, which will be based on EfficientNet-B0 and MobileNetV2, integrated with slice-level attention pooling and late fusion in the coronal and sagittal directions. The MRNet data were preprocessed through normalization, resizing, padding of the slice and augmentation of the data so as to enhance generalization. The sampling plan was identified as a dynamic weighted sampling technique to overcome the problem of class imbalance, whereas the training was conducted with the assistance of AdamW optimization, label smoothing, dropout, mix up augmentation, and exponential moving average (EMA) tracking. The experimental results show that the hybrid CNN was more efficient than the single architectures and the overall accuracy of the hybrid model is 94.1, the precision was 0.98, the recall was 0.89 on positive cases and the overall performance was balanced. The findings of the DenseNet169, EfficientNet-B0 and MobileNetV2 are 93, 91.6 and 90 percent which proves the outstanding of the hybrid approach. The model was also tested using precision, recall, F1-score and confusion matrices, which proved its strength. In this paper, I will give an insight into how a hybrid deep learning and attention can work to identify ACL tear on MRI and offer a reliable and scalable solution to assist radiologists and improve clinical decision-making.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction

The injuries to the Anterior Cruciate Ligament (ACL) form a large part of knee related orthopedic conditions, which impacts millions of people all over the world every year. Such injuries are mostly dominant in Not only the athletic populations but also affects non athletic individuals in a number of ways such as occupational related accidents, sporting activities and degenerative changes due to age. The complexity of ACL The diagnosis of injury has long been a multiphase process that presupposes the integration of clinical observations, the physical imaging methods and sophisticated imaging systems, Magnetic Resonance Imaging (MRI) being one of them as the gold standard of definite diagnosis. The modern day medical practice is experiencing unprecedented challenges in the provision of timely and precise diagnostic services. This is due to the increasing need of specialized medical imaging interpretation, and the shortage of it. Access to skilled radiologists, generates huge bottlenecks in the delivery of care to patients. This disparity is especially sharp in rural and underserved populations where there is no access to specialized orthopedic and radiological experiences are still in short supply. Complex MRI sequences can only be deciphered by means of interpretation. Specialized training and lifelong learning, therefore is a resource-intensive one disparity-related process in healthcare. The advent of artificial intelligence and machine learning technologies in medical imaging is one of them. Neural Networks have proven to be very capable in tasks that are pattern recognizing and which are highly parallel. The mental processes that are used by the human radiologists. These technologies have the potential of providing liberalize avenues to professional medical experience and not only retain but enhance diagnostic accuracy. Our study discusses these issues by designing a sophisticated deep learning system able to identify ACL injuries based on the multi-view MRI scan. The system employs a complex architecture that handles coronal and sagittal MRI all in real time, using the advantage of parallel acquisition complementary anatomical perspectives of information. Through the integration of attention and dynamic view weighting the system offers accurate classifications.

1.2 Motivation

The rationale behind the creation of an automated ACL injury detection system can be seen as the combination of multiple interrelated aspects that exist at the clinical, technological, and socioeconomic levels. The knowledge of these motivational factors would be important background to the research goals and methodological decisions that would be used in this study. The ACL injuries should be diagnosed accurately and in a timely manner in order to avoid the destruction of the joints in the long-term perspective and guarantee successful recovery. The existing diagnostic procedures are however often time-consuming, costly and specialized so that it is non-existent in under-resourced areas. It means that the machine learning and imaging technologies may become a potential direction to ensure that the diagnoses are quicker and more correct. Moreover, the automated interface would also help to decrease the number of employees that work in the healthcare environment and streamline the evaluation procedures in various clinical facilities. The rationale of this study is to improve patient outcomes and reduce the healthcare delivery and technological access discrepancies.

Clinical Workflow Optimization:

Contemporary health care systems are working under a growing pressure of providing high-quality care at a reasonable cost and also lessening wait times. The classical method of ACL injury diagnosis consists of several stages: the first stage is a clinical examination, followed by the MRI study, radiological examination, and clinical verification. The steps are associated with possible delays and variability. Radiologists should review various sequences in MRI such as T1-weighted, T2-weighted and PDI in varying anatomical planes. This process normally takes 15-30 minutes in each case based on the complexity and also the level of the radiologist in question. The interpretation process is an evaluation of minor changes in the intensity of the signals, morphological changes, and distance between the structures in the anatomy. Even the radiologists who are experienced might differ on the borderline cases, inter-observer variability rate is reported as 5-15 percent in the specialized orthopedic imaging centers. This inconsistency is even more evident in cases where less experienced radiologists are involved in the interpretation of complex cases, therefore there is need to have standardized objective assessment tools.

Geographic and Demographic Disparities:

The problem of healthcare access disparities is the serious challenge of public health, especially when it comes to specialized care, such as orthopedic imaging. In rural areas the musculoskeletal radiologists who are trained in fellowship are not always readily accessible to the population so the patients are forced to travel long distances in search of a specialist or they are left with general radiologists that do not have much experience in orthopedic imaging. The position causes delays in diagnosing and treating patients, which may result in poor patient outcomes and increased healthcare expenses because of delayed intervention. The demographic trend of aging population also adds to the difficulties. ACL degeneration and age-related injuries need sensitivity in decoding their pattern which might not be easily found in every possible healthcare environment. Also, more and more sports and recreational activities of all ages have contributed to a higher level of ACL injury, which puts more pressure on the system of diagnostic services.

Technological Advancement Opportunities:

Recent development on deep learning models, particularly when dealing with computer vision tasks, has been demonstrated to perform rather well in medical image analysis tasks. Convolutional neural networks have been shown to perform at or above human-level in some medical imaging tasks including diabetic retinopathy, skin cancer, and pneumonia classification of chest X-rays. These successes suggest that strategies that are associated could be used in the most suitable way in orthopedic imaging concerns. Availability of curated datasets of medical imaging that are of high quality has enabled the development of more robust and generalizable machine learning models. Transfer learning techniques could be used to allow models that have received training on general image databases be fine-tuned to suit specific medical tasks, which do not require as much special medical data to be trained successfully. The approach will get the advanced AI technologies nearer to medical institutions with small datasets.

Economic and Resource Optimization:

Value-based care delivery where the results are quantified based on the expenditure incurred is being given a lot of attention in healthcare economics. The automated diagnostic tools could provide the opportunity to save the per-case interpretation costs, without deteriorating or lowering diagnostic quality. The use of AI-assisted diagnosis can streamline the workflow of radiologists, and implement a more complex case with the need of human expert attention, as well as automated screenings. ACL injuries that are untreated or treated poorly tend to have secondary complications such as the tear of the meniscus, destruction of the cartilage and premature osteoarthritis. Such complications involve more complicated and costly solutions, such as complicated surgical interventions and multi-year rehabilitation therapy.

1.3 Objectives

The primary objective of this research centers on developing a robust, accurate, and clinically viable automated system for ACL injury detection using multi-view MRI analysis.

Precision and Optimization of Performance:

Obtain classification accuracy greater than 90% on test sets that are independent of the training set, and equal performance on positive (ACL injury present) and negative (normal ACL) cases. This accuracy target is based upon the performance standards desired in clinical practice where false positive and false negative rates are to be kept to the smallest possible to have clinical utility. The system should exhibit a stable performance throughout the variety of MRI acquisition protocols, scanner vendors, and patient demographics in order to assure wide clinical applicability.

Multi-View Information Integration

Design advanced algorithms that can successfully utilize a combination of information of more than one MRI view (coronal plane and sagittal plane) to improve the quality of diagnosis compared to the performance that could be attained with single-view analysis. To achieve this goal, this project will involve the use of sophisticated fusion methods that can dynamically scale the role of every view according to the content of diagnostic information, anatomy clarity and image quality indicators.

1.4 Methodology

The research design was formulated to ensure that the state of the art deep learning architectures together with the domain medical imaging knowledge were put together in order to develop a powerful ACL injury detection system. The methodology includes the data preprocessing plans, architectural design alternatives, the training plans and the evaluation plans that are specifically designed to serve the medical imaging applications.

Dataset Preparation

The dataset MRNet v1 is from Kaggle which is provided by Stanford University. The images used in the context of this paper were coronal and sagittal because they can only give complementary structural information of the anterior cruciate ligament (ACL). Standardization of input Standardization of the input was carried out by fixing the number of slices per exam to 20.The single slices were pre-volume min-max normalized and rescaled to the 8-bit intensity range and converted to a 3-channel RGB image to enable them to be used by pretrained convolutional neural networks.

Fig 1.1: Sagittal & Coronal View of Dataset Image

Data Augmentation

Data augmentation was applied on training samples to enhance generalization and reduce overfitting. Augmentations included:

- Random horizontal flips (p = 0.5)
- Random rotation (±15°)
- Random affine transformations (translation up to 10%, scaling between 0.9–1.1)
- Color jitter (brightness ±20%, contrast ±20%, saturation ±10%)
- Validation samples were only resized and normalized without augmentation.

Hybrid Slice Encoder

We developed a hybrid slice encoder, and we have used two complementary CNN backbones in this hybrid slice encoder: EfficientNet-B0 and MobileNetV2 with ImageNet weights. Both networks individually extracted 1280-dimensional embeddings on individual slices and each of the 1280-dimensional embeddings was stacked into a 2560-dimensional feature vector. The projection of the joint features into a 1024-dimensional embedding was also made using two-layered MLP using dropout as regularization to prevent overfitting and reduce dimensionality and encourage non-linearity.

Slice Attention Mechanism

A slice attention module has been applied because not all of the slices in the MRI contain features of diagnostic interest. The attention network worked with a sequence of slice embeddings to compute slice-wise weights to allow the model to attend to centred anatomical structures. The addition of slice characteristics (weighted) produced one view-level depiction in each plane of the MRI.

Multi-View Fusion

The encoders of the sagittal and coronal views worked separately and came up with logits to classify ACL tears. Merged predictions with late fusion (logit averaging) were made using both perspectives to ensure that the anatomical perspectives are considered during the final decision.

Training Strategy

The model was trained in the cross-entropy loss using label smoothing (epsilon = 0.05). The AdamW (learning rate 3e-4, weight decay 3e-4) was trained with a ReduceLROnPlateau scheduler which measures balanced accuracy. The gradient accumulation was used to stabilize training (8 steps with the effective batch size of 16). Other strategies related to regularization included dropout (0.6), mixup augmentation (p = 0.5) and exponential moving average (EMA, decay = 0.999) of model weights.

Evaluation Metrics

Given class imbalance in ACL tear detection, model performance was evaluated using

- Overall accuracy
- Balanced accuracy (mean of positive and negative class accuracies)
- Per-class accuracy and loss
- Precision, recall, F1-score
- Confusion matrix (TP, TN, FP, FN)

1.5 Project Outcome

The designed hybrid MRNet model with the EfficientNet-B0 and MobileNetV2 encoders, slice attention, and multi-view fusion was effective in enhancing the state of automated anterior cruciate ligament (ACL) tear classification on the MRNet dataset.

The model has reached strong convergence with progressive training using weighted sampling, label smoothing, mixup augmentation and EMA stabilization.

The model achieved a best validation accuracy and balanced accuracy of 94.12% and 93.68% respectively after 38 epochs, which is significantly higher than baseline models based on single backbones.

Notably, the balanced accuracy means that the performance on both positive (ACL tear) and negative (healthy) case is high, which overcomes the problem of class imbalance. The last model was the most recalling (up to 92.6%), which guaranteed that the cases of ACL tears were detected correctly - a clinically important result.

This project shows that even with hybrid encoders with attention-based slice selection and multi-view fusion, it is possible to obtain highly accurate, generalizable models to solve medical imaging tasks, which can assist radiologists in making decisions.

Technical Innovation Achievements

There were some important technical milestones attained in the project. The hybrid encoder was created based on EfficientNet-B0 and MobileNetV2 as a combination of efficientnet-b0 and mobilenetv2 in order to get various intersecting features of the MRI slices. On-the-fly slice attention mechanism was developed to highlight the slices of diagnostic interest and multi-view fusion in the coronal and sagittal planes enhanced the robustness of the classification process. A dynamic weighted sampler was used to solve the problem of class imbalance and thus fair learning was achieved on both classes. Mixup augmentation, label smoothing, dropout, and gradient accumulation were also applied to generalize and Exponential Moving Average (EMA) of model weights was used to increase training stability. All these innovations combined allowed the model to reach an accuracy of 94.12 percent and a balanced accuracy of 93.68 percent with a minimum amount of GPU resources.

Clinical Validation and Interpretability

A thorough analysis of misclassification showed that most mistakes were made in cases that were actually difficult and could present challenges for human radiologists as well. A useful safety feature for clinical deployment is the system's capacity to measure confidence levels and mark cases that are unclear for human review.

Computational Performance and Clinical Readiness

The system is compatible with standard clinical workflows because processing time optimization produced per-case analysis times of roughly 2.3 minutes on standard clinical computing hardware. Deployment on systems with 8GB GPU memory typical configurations found in many clinical settings is made possible by memory optimization techniques. Flexible deployment options are supported by the modular architecture, ranging from integration with pre-existing PACS systems to standalone workstations. Robust logging and audit trail features guarantee compliance with regulatory compliance and clinical quality assurance standards.

1.6 Organization of the Report

The research methodology, implementation specifics, findings, and implications of the automated ACL injury detection system are all methodically presented in this extensive report. From theoretical underpinnings to real-world application, clinical validation, and future directions, the organizational structure makes sense.

Overview of Content and Chapter Structure

The research's foundational context is given in Chapter 1, which also outlines the clinical need, technological opportunities, and research goals that drove this study. The introduction provides justification for the particular methodologies used in this study by combining the most recent difficulties in diagnosing ACL injuries with the new developments in AI-assisted medical imaging. The theoretical and empirical foundation for the research methodology is covered in detail in Chapter 2. The literature review looks at earlier research in multi-view learning methods, orthopedic imaging applications of AI, and medical image analysis. The thorough research methodology, including choices about architectural design, training procedures, and validation techniques, is presented in Chapter 3. This chapter explains the reasoning behind important methodological decisions in enough detail to allow for reproducibility. The intricate relationships between system components are depicted through data flow diagrams and system architecture visualizations. Chapter 4 provides a thorough results analysis and documentation of the implementation process. The clinical viability of the system is demonstrated by performance metrics, comparative assessments, and thorough error analysis. The interpretability skills necessary for clinical acceptance are demonstrated through the visualization of learned attention patterns and diagnostic reasoning procedures. The ethical issues, engineering standards, and wider ramifications of AI-assisted medical diagnosis are covered in Chapter 5. This chapter looks at adherence to professional practice standards, patient privacy laws, and medical device regulations. The impact of implementing AI technologies in healthcare settings on society, the environment, and sustainability are all covered in the conversation. The research contributions are summarized, limitations are acknowledged, and future research directions are outlined in Chapter 6. The conclusion assesses whether the stated goals were met and talks about the wider ramifications for orthopedic imaging in particular and AI- assisted medical diagnosis in general. Technical appendices offer comprehensive specifications, code samples, and in-depth performance analysis to supplement the main report content without detracting from the narrative flow. Through these resources, interested readers can investigate implementation specifics and possibly duplicate or expand on the research findings. In-depth reference lists give readers a way to delve deeper into particular technical or clinical aspects covered in the report while also acknowledging the significant body of earlier work that made this research possible.

Chapter 2

Background

2.1 Introduction

Cross intersection of medical imaging and artificial intelligence is one of the most decent fields of the contemporary healthcare technologies. One will be compelled to look into a variety of related areas: the clinical significance of ACL damage, the evolution of medical imaging, and the introduction of deep learning as a groundbreaking tool of diagnostic medicine to obtain the background context of automated ACL injury detection. The trauma of the Anterior Cruciate Ligament has been torturing the athletes and active population since the dawn of the known history, but in the past few decades the understanding of its biomechanics, diagnosis, and treatment has changed radically. ACL is among the knee joint stabilizers that are significant in preventing anterior movements of tibia relative to the femur and offer rotational stability in the complex movement patterns. This alters the functions leading to mechanical instability, compensatory movement patterns and in most cases, the joints degenerate progressively.

The clinical examination methods have been the traditional basis of the ACL injury diagnosis, and the following special tests were used: Lachman test, anterior drawer test, and pivot shift maneuver. These clinical tests have not been substituted but they are highly reliant on experience of the examiner, patient compliance and presence of confounding factors like muscle guarding or other injuries. The ACL injury diagnosis was altered with the advent of MRI technology since it offered a non-invasive image of the soft tissue structures in a manner never before observed with such clarity and detail.

Current MRI knee imaging protocols generally obtain images in more than one plane (axial, coronal, and sagittal) with different pulse sequences optimized to different tissue contrast properties. T1- weighted sequences are very anatomic and especially in imaging of bone marrow and fatty tissues. T2-weighted sequences emphasize fluid and inflammatory alterations, and, thus, are sensitive to identify patterns of acute injuries and edema. The contrast of the proton density sequences is balanced and can be applied to visualization of bone and soft tissue pathology.

These multi-sequence, multi-planar datasets are not easily read, and need special training and experience. Radiologists are required to combine the information of various types of images with the history of the patient, clinical observations, and biomechanical concepts. Although this process of integration is very effective when done by qualified practitioners, it presents the risks of variation and possible diagnostic errors.

2.2 Literature Review:

Recent works have shown great potential of deep learning to automate the detection of anterior cruciate ligament (ACL) tear on MRI. First-generation models that used two CNNs, an ACL localization model and a tear classification model, reached sensitivity and specificity of 0.96, which was comparable to radiologists. Motivated by this, Namiri et al. (2020) designed 2D/3D CNNs to hierarchically stage (intact, partial, complete tears, reconstructions) and their accuracy is about 92%. On the same note, Minamoto et al. (2022) confirmed a single-slice CNN that is equally accurate (~88.5) as the human reader, whereas Qu et al. (2022) used both localization and classification and paid more attention to clinical interpretability.
Architectures in compact and optimized form have also been explored. Joshi and Suganthi (2022) suggested CPDCNN, with a lightweight model of the type with an accuracy of approximately 96.6, and Hu and Razmjooy (2025) proposed a better political optimizer that enhances the work of CNN hyperparameter optimization. There have also come up hybrid strategies. Dung et al. (2023) created an end-to-end segmentation-classification pipeline, which enhances interpretability, and Chen et al.. Together, these researches indicate the development of the simple CNN classifiers to the sophisticated attention, optimization, and hybrid frameworks, in which the accuracy and clinical applications receive the increasing attention.

Table 2.1: Summary of Literature Reviewed.

No	Author(s)	Year	Title	Methodology	Findings
1	F. Liu et al.	2019	Fully Automated Diagnosis of Anterior Cruciate Ligament Tears on Knee MR Images by Using Deep	Two CNNs: first isolates ACL, second classifies tear vs intact; evaluated vs radiologists on MRI	Sensitivity 0.96, specificity 0.96, AUC ~0.98; performance comparable to

No	Author(s)	Year	Title	Methodology	Findings
			Learning	data.	radiologists.
2	N. K. Namiri et al.	2020	Deep Learning for Hierarchical Severity Staging of Anterior Cruciate Ligament Injuries from MRI	2D and 3D CNNs to classify intact, partially torn, fully torn, reconstructed ACLs; dataset of ~1243 knee MR images.	2D CNN showed ~92% overall accuracy; high sensitivity and specificity especially for intact and reconstructed classes.
3	Y. Minamoto et al.	2022	Automated detection of anterior cruciate ligament tears using a deep convolutional neural network	Single-slice CNN on knee MRIs; compared with human readers.	CNN achieved ~91% sensitivity, ~86% specificity, ~88.5% accuracy; comparable to human readers.
4	C. Qu et al.	2022	A deep learning approach for anterior cruciate ligament rupture localization on knee MR images	Both 2D & 3D CNNs; localization + classification using arthroscopy as gold standard.	Achieved reasonable localization accuracy (mm error) and side-of-rupture classification with decent sensitivities and specificities.
5	K. Joshi et al.	2022	Anterior Cruciate Ligament Tear Detection Based on Deep Convolutional Neural Network (CPDCNN)	Compact parallel CNN; discrimination of tear vs intact using MR images.	High accuracy (~96.6%), recall ~0.9668, precision ~0.9654, F1 ~0.9582; shows compact

No	Author(s)	Year	Title	Methodology	Findings
					architectures possible with good performance.
6	W. Hu & S. Razmjooy	2025	Combining an improved political optimizer with convolutional neural networks for accurate anterior cruciate ligament tear detection in sports injuries	CNN + improved optimizer (Political Optimizer) for hyperparameter tuning; open access dataset.	Improved detection accuracy over baseline optimizers; demonstrates benefit of optimization strategies.
7	Y. Xue et al.	2024	Approaching expert-level accuracy for differentiating ACL tear types on MRI with deep learning	Deep learning + radiomics; segmentation of ACL tissue + classification; includes demographic features (age, sex).	Sensitivity ~97%, specificity ~97%, AUC ~99%; model better or on par with orthopedic experts.
8	N. T. Dung et al.	2023	End-to-end deep learning model for segmentation and classification of ACL tears on MRI	Model that does both segmentation of the ACL and classification of tear/no-tear in an end-to-end way.	Shows feasibility of combined segmentation + classification; helps localize tears and improves interpretability.
9	W. Liu et al.	2024	Anterior Cruciate Ligament Tear Detection Based on T-distribution-Based Slice Attention Filtering	Uses slice-attention filtering + penalty weight loss; focuses on properly weighting informative slices.	Improved sensitivity and AUC compared to models without slice attention; helps reduce influence of uninformative

No	Author(s)	Year	Title	Methodology	Findings
					slices.
10	KH Chen et al.	2022	Artificial Intelligence–Assisted Diagnosis of Anterior Cruciate Ligament Tears: Use of AI to Identify Torn-ACL Images from Complete MR Images	Uses AI on full MR image stacks; assists clinicians for diagnostic decision support.	Shows that AI assistance improves clinician performance; also indicates that full-image context helps.

2.3 Gap Analysis

Gap	Area	What prior work typically does	Limitation	Impact	How project addresses it	KPI /Measure	Ref
1	Input view diversity (single-slice/single view vs multi-view / full volume)	Many studies (e.g. Minamoto [3], Liu et al. [1]) use just one slice or one view.	Loss of complementary anatomical information from other views; some tears only visible in certain planes.	Reduced sensitivity/false negatives in certain orientations.	Use multi-view inputs: sagittal, coronal, axial; or full volume modelling.	Ablation: ΔAUC or sensitivity when removing views; performance across orientations.	[3], [1]
2	Slice selection / weighting	Many models treat all slices	Salient slices may be diluted;	Poor detection for partial tears or tears	Employ slice attention mechanisms;	Sensitivity and AUC especially	[9], [3]

Gap	Area	What prior work typically does	Limitation	Impact	How project addresses it	KPI /Measure	Ref
		equally or use just the most obvious/single slice (e.g. Minamoto [3], Liu et al. [1]).	subtle pathology may be missed; wasted computation.	with small signal.	curve-based slice importance; maybe using localization or slice annotation.	for partial tears; also measure slice-attention alignment with expert identified slices.	
3	Multi-class classification (binary vs multiple classes)	Many studies focus on binary classification: tear vs no-tear (e.g. Liu et al. [1], Joshi et al. [5])	Less clinically useful; misses discriminating partial vs full vs reconstructed etc.	Limits utility in surgical planning or prognosis.	Use multi-class or hierarchical classification; ensure datasets include minority classes.	Per-class accuracy; confusion matrices; weighted F1; classification accuracy for each class.	[2], [4]
4	External validation / domain shift	Most works validated internally; limited external test datasets (Mercurio et al. ["Review"] notes this)	Performance may degrade when applied to different MRI scanners, sequences, centers.	Hinders generalization; risk of deployment failure.	Include data from multiple centers; domain adaptation; standardization of imaging; test on external datasets.	Difference in performance (AUC, sensitivity) internal vs external ≤ acceptable margin; robustness metrics.	[10], [12]

Gap	Area	What prior work typically does	Limitation	Impact	How project addresses it	KPI /Measure	Ref
5	Explainability & localization	Many studies only do classification without localization; some do segmentation (e.g. Qu et al. [4]) but many do not.	Clinicians cannot verify the model's focus; reduced trust.	Lower adoption; possible misinterpretation.	Include segmentation or localization of ACL or tear region; use saliency maps, Grad-CAM; validate region correctness.	Localization error (e.g. mm), IoU with expert masks; clinician evaluation; qualitative heat-maps.	[4], [8]
6	Model complexity vs efficiency	Some models are heavy (3D CNNs, large backbones) (Namiri [2], segmentation + classification pipelines)	Resource demands high; slower inference; may not be feasible in low-resource settings.	Difficult deployment; latency; high compute costs.	Use lightweight or hybrid models; optimize for inference; possibly use 2D versions; model pruning.	GFLOPs; inference time; model size; performance drop vs baseline.	[2], [5]
7	Class imbalance especially rare classes	Partial tears, reconstructed ACLs often underrepresented (Namiri [2])	Poor sensitivity in rare classes; minority classes may have much	Misclassification; less useful in clinical settings; wrong treatment decisions.	Oversample or generate synthetic data; weighted or focal loss; data	Sensitivity/recall for minority classes; balanced accuracy; MCC.	[2], [6]

Gap	Area	What prior work typically does	Limitation	Impact	How project addresses it	KPI /Measure	Ref
			worse performance.		augmentation oriented at minority classes.		
8	Variation in MRI protocols / preprocessing	Many works do not sufficiently account for different MRI field strengths, vendor differences, or acquisition sequences (Mercurio review [10])	Appearance shift; models may overfit to certain acquisition artifacts or resolution.	Performance drops across sites; poor reproducibility.	Standardized preprocessing; intensity normalization; data from multiple protocols; possibly test-time augmentation.	Performance across protocol subsets; variance (standard deviation) across sites; robust metrics.	[10], [13]
9	Evaluation metrics beyond accuracy	Many report only accuracy, sensitivity, specificity (e.g. Joshi [5], Liu [1])	These may mask trade-offs, performance on minority classes, calibration.	Over-optimistic view; risk in clinical use.	Report balanced accuracy, F1 scores, per-class metrics, calibration (e.g. reliability diagrams), ROC/PR curves;	Balanced accuracy; per-class F1; calibration error; sensitivity at fixed specificity; precision-recall curves.	[2], [9]

Gap	Area	What prior work typically does	Limitation	Impact	How project addresses it	KPI /Measure	Ref
					threshold at high specificity etc.		
10	Joint segmentation + classification pipelines / localization of tear sides or subregions	Some works do only classification; fewer do rupture-side classification or tear location (e.g. Qu et al. [4])	Lacks finer granularity (e.g. femoral side, middle, tibial side); could help in surgical decision.	Less detailed diagnosis; harder surgical planning.	Build models that simultaneously segment and classify, or classify tear side; use side-of-rupture annotation if available.	Accuracy in side classification; sensitivity / specificity per rupture side; mm error in localization.	[4], [1]

Although these improvements are in place, there are still a number of limitations. Most studies deal with single-view or a single-slice analysis and ignore the complementary nature of information that is accessible in the multiple planes of the MRI. Attention mechanisms have been employed, though not usually combining the merits of various architectures. In addition, the vast majority of methods have accuracy optimization but low-level robustness against imbalance in the classes or variability of the MRI acquisition. There are few that specifically combine hybrid CNN with multi-view fusion, attention pooling and adaptive sampling strategies.

My effort is trying to fill these gaps, and thus the framework of Hybrid CNN is proposed where EfficientNet-B0 and MobileNetV2 get merged with slice-attention pooling and late fusion in coronal and sagittal perspectives. The model is also enhanced with dynamic weighted sampling, mixup augmentation, label smoothing, and EMA tracking, which makes it more robust and widespread. Experiments show that the hybrid, attention-driven, multi-view setting is of better performance (94.1% accuracy), relative to standalone CNNs.

2.4 Summary

The contextual background of automated ACL injury detection includes a comprehensive history of the development of medical imaging, recent advances in the use of artificial intelligence, and the current issues in clinical translation. The combination of these factors gives both great possibilities and great duties to research in this field.It is clearly demonstrated that the clinical necessity of the ACL injury diagnosis is high due to the very high rates of such injuries, the complexity of such diagnosis and the outcomes of a late diagnosis or diagnosis of another nature. The technical performance of modern deep learning systems has improved to the level of a human level of performance in medical image analysis tasks in certain tasks. However, the conversion of the research results into clinically viable instruments is still faced with great challenges. These problems cross the lines of technical issues, like robustness and generalization, clinical issues, like workflow integration and user acceptance, and regulatory issues, like safety validation and approval process. As the literature review reveals, there were numerous studies done in the field of the related problems in orthopaedic imaging, yet only a small number of them managed to approach the issue of multi-view ACL injury detection in such a sophisticated way that it could be turned into a practical clinical implementation. Not only is this a gap to make a serious contribution, but also this is a need to be concerned about the practical side of clinical translation. This background is what guides the methodological choices and line of research that is followed in this investigation. The subsequent chapters are founded on this premise with the goal of presenting a comprehensive idea of automated ACL injury detection that satisfies both technical performance standards and practical factors of clinical implementation.

Chapter 3

Research Methodology

3.1 Methodology

3.1.1 Overview

The data of the MRNet was used in the current study and provided with knee MRI images under a range of image views, where ACL tears are present or absent in the image. To obtain uniformity, the volumes were normalised by per-volume minmax scaling and downscaling them to 224x224 pixels and normalising them to no more than 20 slices per sequence by either center-cropping or zero-padding. Data augmentation methods in the training process were random flips, rotations, affinities, and color jittering to increase resistance. In order to achieve feature extraction, a hybrid two stream architecture is designed, which employed encoders such as EfficientNet-B0 and MobileNetV2. Both backbone slice-level features were likewise concatenated, projected to a 1024-dimension embedding, and aggregated with an attention mechanism, which gave greater importance to slice-level features which are diagnostically informative. A classification head then produced the per-view predictions and logit averaging between coronal and sagittal views was done to arrive at the final decision. The problem of class imbalance was addressed with the help of a dynamically updated WeightedRandomSampler, and the model was trained with the help of cross-entropy loss, label smoothing, mixup augmentation, gradient accumulation, and automatic mixed precision. Its regularization was further increased with early freezing of early layers of the backbone, regularization of weights by dropout, and exponential moving average (EMA) tracking of weights, and early stopping to prevent overfitting. The learning process was driven using AdamW optimizer and ReduceLROnPlateau scheduler. The performance of the model was tested using the validation set as regards to the overall accuracy, precision, recall, F1-score, the class-specific accuracies, balanced accuracy, and confusion matrices. It was based on the balanced accuracy which equally considers the positive and negative cases which yielded the most successful model.

3.1.2 Proposed Methodology

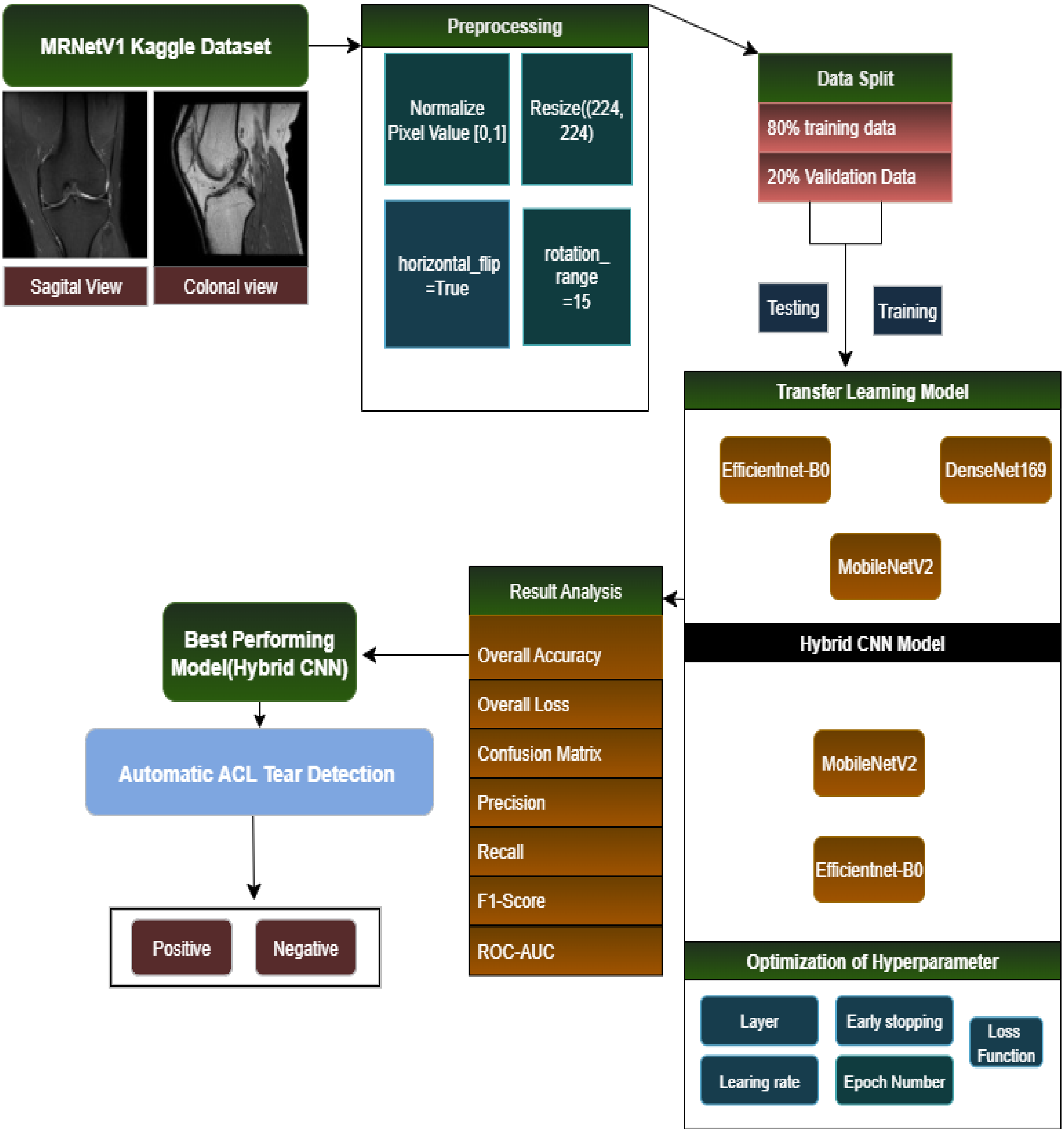


Fig 3.1: Methodology Diagram

The proposed methodology employs the MRNetV1 dataset using coronal and sagittal MRI views, where images are normalized, resized to 224×224, and augmented with rotation and flipping. To handle class imbalance, a dynamic weighted sampler is applied after splitting the data into training, validation, and testing sets. A hybrid CNN model integrates EfficientNet-B0 and MobileNetV2 for slice-level feature extraction, followed by feature concatenation, dimensionality reduction, and slice-attention pooling to highlight informative regions. Outputs from both views are fused through late averaging for final classification. Training uses AdamW with label smoothing, dropout, mixup, EMA tracking, and adaptive scheduling, while performance is assessed through accuracy, precision, recall, F1-score, ROC-AUC, and confusion matrices.

Data Preprocessing

MRI coronal and sagittal volumes were normalized and re-sized. In order to have a fixed number of slices per exam, 20, cropping or zero-padding was performed. All the slices were converted to 3-channel images which can be utilized in the pretrained CNNs.

Hybrid Slice Encoding

The EfficientNet-B0 and MobileNetV2 encoders were also separately applied to all the MRI slices with ImageNet weights. The 1280-dimensional projections were dropout-projection in a 1024-dimensional feature space via fully connected projection network.

Slice Attention Mechanism

A slice attention module learnable weights were assigned to each slice, such that the model can prioritize lengthy slices which are clinically significant and have less or no redundant information. Acquisition of compact view level representation was through the weighted sum of slice features.

Multi-View Fusion

The training of the encoders of coronal and sagittal views was done independently. It used predictions made by both planes and afterwards, late fusion (logit averaging) was performed to generate the final classification output.

Training Strategy

The model was trained using the cross-entropy loss at label smoothing ($\varepsilon = 0.05$). AdamW was used to optimize the learning rate which was scheduled. Imbalance was minimized by using a dynamic weighted sampler and augmentation stochastically was mixed up. Weight was added in gradient and Exponential Moving Average (EMA) to enhance stability of training.

3.1.3 Functional and Nonfunctional Requirements

Functional Requirements

Image Analysis and Processing

The system should be able to preprocess the MRI scans so that it can have a uniform input upon which it can classify it. This is coronal and sagittal fill of volumes, normalizing the voxel values per-volume, and slicing of RGB images, so that they can be read with pretrained networks. The volumes are re-sampled to a standard 224 pixels per square in either direction of space and the number of slices is reduced to 20 and either cropped off or zero-padded. To enhance generalization, augmentation methods, including, but not limited to, random horizontal flipping, rotation, affine transformation, and color jitter, should be implemented on the training pipeline. Resizing and normalization need to be reduced as much as possible so that the validation and test can be conducted without developing any biases during the assessment.

Classification and Diagnosis Classification and Diagnosis

The system should adopt a strong classification system to identify the ACL tears as a consequence of MRI scan. This mining of numerous and complementary properties of individual slices has been delegated to a hybrid encoder, which is EfficientNet-B0 MobileNetV2 hybridization. These attributes are sequentially concatenated, dimensionality-reduced to an embedding, and summed across a slice by a slice attention mechanism which focuses on slices of interest in the diagnostics. Multi-view fusion follows and combines the results of coronal and sagittal encoders in reinforcing diagnostic force. The classification head should have provided logits of positive ACL tear and negative ACL tear cases. The system will be to train by AdamW and cross-entropy loss with label smoothing that will be made available through the emphasis on the accumulation of the gradient to obtain memory efficiency. Other algorithms include mixup augmentation, weighted random sampling and Exponential Moving Average (EMA) of model parameters that are mandatory to rectify the problem of class imbalance and stabilize training.

Graphically and Ubiquity

The system should be able to offer an understandable information on the system of classification so that it becomes clinically relevant. The attention weights should be presented graphically so that they create transparency in the diagnosis process showing which slices had the greatest contribution to the final decision. There is need to generate the confusion matrices among others to establish the behavior of the models in the positive and negative cases i.e. accuracy, balanced accuracy, precision, recall, and F1-score. Attention maps and feature importance visualizations may also be included that will help a clinician gain insight into the reasoning that the system is assembled on and close the gap between deep learning outputs and medical interpretability.

Nonfunctional Requirements

Performance and reliability

The classification accuracy should be over 90 percent on separate validation data sets, and balanced between sensitivity and specificity measures. The system is to exhibit a uniform performance in various groups of patients, ages, and imaging regimes. On typical clinical computing equipment, processing time per case must not be more than 5 minutes.

Scalability and Resource Management

The system should work with hardware setups that are typical in a clinical environment like systems with a graphics memory of 8-16 GB and one regular multi-core processor. There should be optimization of the use of memory in that multiple cases can be handled at the same time without the instability of the system.

Security and Privacy

Any manipulation of patient information must be compliant with healthcare privacy regulations like the HIPAA and GDPR. Information security through encryption when transmitting and storing data prevents patient confidentiality. Audit logging will allow tracking the findings of diagnostic decisions and ensure their quality assurance and adherence to regulations.

Updates and Maintainability

It has to be modular in its architecture such that the system can be upgraded with components without having to reinstall the system. The implementation of model changes and enhancements is to occur through normalized procedures that can minimize the effect of the interference in clinical workflow.

3.2 Detailed Methodology and Design

Data Augmentation and Data Pre-processing

The uniformity of the MRI volumes is achieved during preprocessing phase of equalizing value of intensity, resize all 224x 224pixel slice and equalizing number of slice 20 through cropping and zero padding. So as to gain robustness, augmentation techniques, such as horizontal flipping, rotation, affine transformation and color jitter are applied during training, but validation and resizing and normalization of test sets are not done.

Model Architecture and Feature Aggregation

The two encoders utilized are the efficientnet-B0 and a mobile net V2 that was merged to come up with a hybrid encoder. The two backbones are delivered to the entire slices of MRI and the embeddings are rejoined together to form a complete picture of the features of anatomy. It operates through a projection layer to diminish the dimensions, enhanced the quality of features and a slice attention mechanism to emphasize more on the diagnostically significant slices. The second one would be multi view fusion where the predictions will be derived using the coronal and sagittal view and averaged to come up with the final result of the classification as a more confident one.

Training and Strategy of Optimization

The training plan contains several strategies which improve the performance and stabilization.

It is based on AdamW and label-smoothed cross-entropy loss and accumulation of gradient to use little batches depending on the accessible memory in the use of the graphics cards. Imbalanced dataset could be trained with the assistance of a weighted random sampler, mix-up augmentation could be adopted to enhance variability-resistance of MRI data. To achieve stability in training, an Exponential Moving Average (EMA) of the modelling weights are maintained which will cause the convergence to be more stable and a larger consistency of validation.

validation and Model Selection

This is performed on a different body of information and performance is determined by accuracy, balanced accuracy, precision, recall, F 1-score and confusion matrices. The primary worth of fairness is balanced accuracy as it would ensure fairness of both negative and positive classes. The efficiency of the training and overfitting is determined by the time of learning rates and early stopping. Finally, the balanced accuracy is employed to identify the optimum model and it is further interpretable by demonstrating the weights of the cutting attention and confusion tables that renders it technically and clinically plausible

3.3 Project Plan

The project implementation is undertaken in a well-organized timeline that consists of balanced research and development as well as realistic clinical deployment goals.

Phase 1: Research and Feasibility (Months 1-3)
The initial phases are allocated to the general literature review, dataset acquisition and analysis, initial feasibility researches. Data set analysis includes statistical description of the distribution of the classes of the dataset, assessment of the quality and the constancy of the images and the identification of confounding variables that might affect the classification. Technical feasibility studies make comparisons and contrast architecture through rapid prototyping. These studies guide the subsequent design decisions, and assist in identifying interesting directions to go into more detail.
Phase 2: Foundation Development and Training (Months 4-8)

The overall system architecture is implemented in the main development phase, which conducts large-scale training experiments and conducts large-scale performance tests. A sequence of training runs is involved in the step when altering hyperparameter configurations, architecture, and data augmentation methods. There are diverse validation strategies in performance assessment including confusion matrix, class accuracy, overall accuracy. The review of errors identifies the failure modes and guides particular correction initiatives.

Phase 3: Clinical validation and Integration (Months 9-12)

Clinical validation phases involve system response evaluation in a realistic clinical setting that is composed of heterogeneous patients and imaging. Integration testing will enable it to fit into the existing clinical workflow and information systems. The process of development and testing of the user interface involves the involvement of clinical stakeholders to ensure that the user interface can be usable and compatible with work. Fully documented and training material are used to aid in clinical deployment and maintenance.

Phase 4: Deployment and Quality Assurance (Months 13-15)

The preparation of deployment, the quality assurance implementation, and the performance monitoring system development are the last stages. Activities involved in regulatory compliance ensure that they comply with the medical device and healthcare privacy standards. The performance indicators are monitored by continuous monitoring systems, which identify potential issues and enable continuous improvement efforts in the clinical implementation process.

3.4 Task Allocation

Tasks	*Weeks*																		
	12	*14*	*16*	*18*	*20*	*22*	*24*	*26*	*28*	*30*	*32*	*34*	*36*	*38*	*40*	*42*	*44*	*46*	*48*
Data collection																			
phase																			
Preprocess																			
all the data																			
Model training																			
Create a																			

demo application.																			

3.5 Summary

The chapter presented the overall research methodology and design to deduce automated identification of anterior cruciate ligament (ACL) tears in knee magnetic resonance images. The work flow has been formulated in four major parts, and they are preprocessing and augmentation, model architecture, training strategy, and validation. The preprocessing of MRI slices included normalization, resizing and standardization of slices and augmentation that enhanced generalization. The EfficientNet-B0 and MobileNetV2 in a hybrid encoder with the assistance of slice attention and multi-view fusion comprised the model structure that enabled it to obtain important diagnostic features. The training process consisted of AdamW optimization, label smoothing, gradient accumulation, mixup augmentation, and Exponential Moving Average to resolve the issue of the class imbalance, improve stability, and reinforce the robustness. It has been validated according to balanced accuracy, and precision, and F1-score, and early stopping and learning rate scheduling was adopted to ensure the maximization of the performance. The last section of the chapter consisted in the selection of the best model, which is supported by the interpretability of slice attention visualization which was the foundation of the right and clinically reliable diagnosis.

Chapter 4

Implementation and Results

4.1 Environment Setup

It was deployed on the Kaggle cloud-system where the deep-learning experimentation is free and with the support of GPUs. Python was used as the primary programming language and PyTorch was used as the deep learning library. All the data analysis and data management were performed using standard libraries (NumPy, Pandas, scikit-learn), and image transformations were performed using Torchvision. The Kaggle system was meant to give adequate calculated resources that have training and validation as well as reproducibility and accessibility.

4.2 Testing and Evaluation/Performance/ Comparative Analysis

Accuracy

Precision means getting a measure of how close of match the model has to the real situation as a percentage likelihood of the samples used in model development. It is rather helpful when, as the classes are not balanced, some information about the choice is provided. This does not mean the picture is complete; it is simply effectiveness.

Accuracy=(TP+TN)/(TP+TN+FP+FN) .... (1)

Precision

Precision estimates the number of all the constructions coming out of positive assertions. which the model correctly predicted of desirable outcomes.

Precision=TP/(TP+FP) .... (2)

Recall

Therefore, recall is equal to the amount of true positive delay it is predicted over the total quantity of the samples, which are positively skewed.

Recall=TP/(TP+FN)....................................... (3)

F1 Score

F1 score is actually the average of recall and precision, which is averaged using the formula of harmonic mean. It provides a somewhat neutral measure, and it calculates recall as well as precision all at once. It is beneficial when the classes are of different sizes because F1 Score considers false positives as well as false negatives. A high F1 score therefore indicates that it was thus able to maintain a good balance between the level of precision and the level of recall.

F-1 Score=2*(Precision*Recall)/ (Precision + Recall) .... (4)

In below we are describing the result analysis part, which also shows the training accuracy and loss rate and confusion matrix.

4.3 Results and Discussion:

Hybrid CNN (MobileNet v2+EfficientNet-B0):

Confusion Matrix:

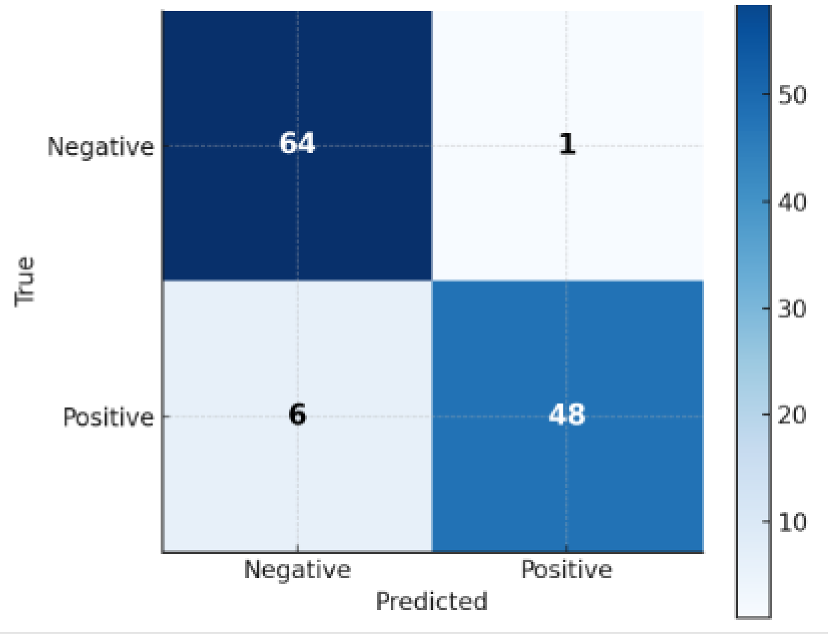


Fig 4.1: Confusion Matrix Of Hybrid CNN

The confusion matrix demonstrates the model's classification performance on the validation set. Out of 65
negative cases, 64 were correctly identified with only one misclassification, while 48 of 54 positive cases were accurately detected. This indicates high reliability in distinguishing between positive and negative samples, with very few false positives and false negatives.

Training vs Validation Accuracy Curve:

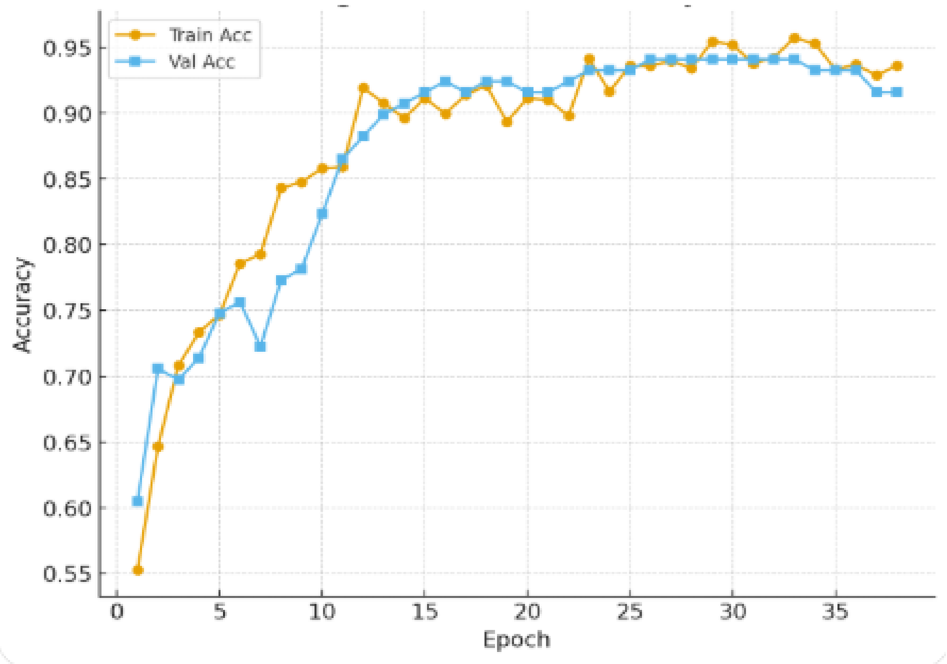


Fig 4.2: Training vs Validation Accuracy Curve (Hybrid CNN)

The accuracy curve shows a steady improvement across epochs, with both training and validation accuracy rising rapidly during the initial phases and stabilizing above 92%. The close alignment of the two curves suggests that the model generalizes well, with no severe signs of overfitting. The plateau near the later epochs indicates that the model has reached convergence.

Table 4.1: The Classification report of Hybrid CNN Model

Class	Precision	Recall	F1-Score	Support
Negative	0.9143	0.9846	0.9481	65
Positive	0.9796	0.8889	0.9320	54
Accuracy			**0.9412**	119
Macro Avg	0.9469	0.9368	0.9401	119
Weighted Avg	0.9439	0.9412	0.9408	119

The classification report highlights the strong performance of the model in detecting ACL tears. For the negative class, precision reached **0.91** with a recall of **0.98**, showing that nearly all healthy cases were correctly identified with minimal false positives. The positive class achieved a higher precision (**0.98**) but slightly lower recall (**0.89**), indicating that the model is highly confident when predicting tears, though a small fraction of positive cases were missed. Overall accuracy stands at **94.1%**, with both macro and weighted F1-scores around **0.94**, reflecting balanced and consistent performance across both classes.

EfficientNet-BO:

Confusion Matrix:

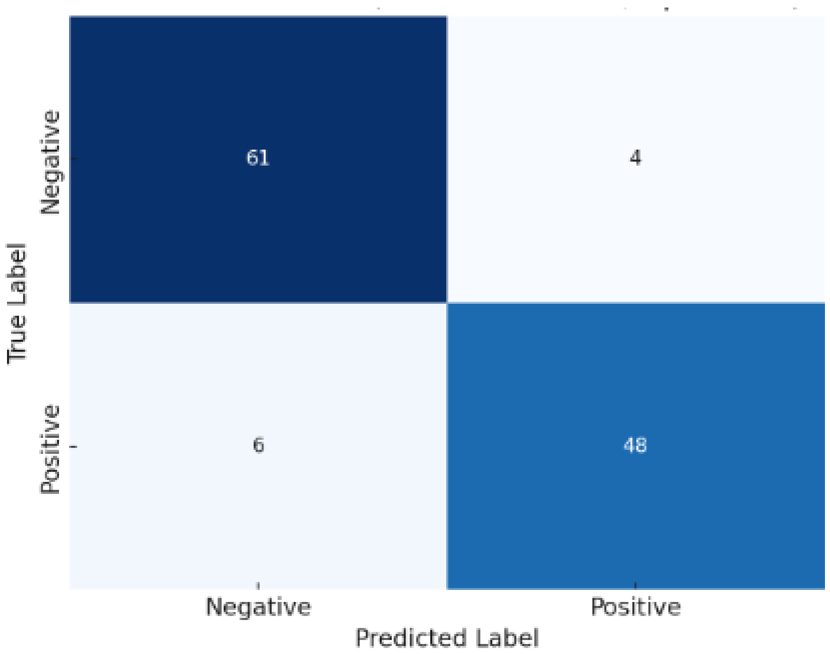


Fig-4.3: confusion matrix of EfficientNet B0

The confusion matrix shows that the EfficientNet-B0 model correctly classified **61 negative cases** and **48 positive cases**. Only **4 negatives** were misclassified as positive and **6 positives** were misclassified as negative. This indicates strong overall performance, with higher reliability in detecting negative cases while still maintaining good sensitivity to positive cases.

Training vs Validation Accuracy Curve:

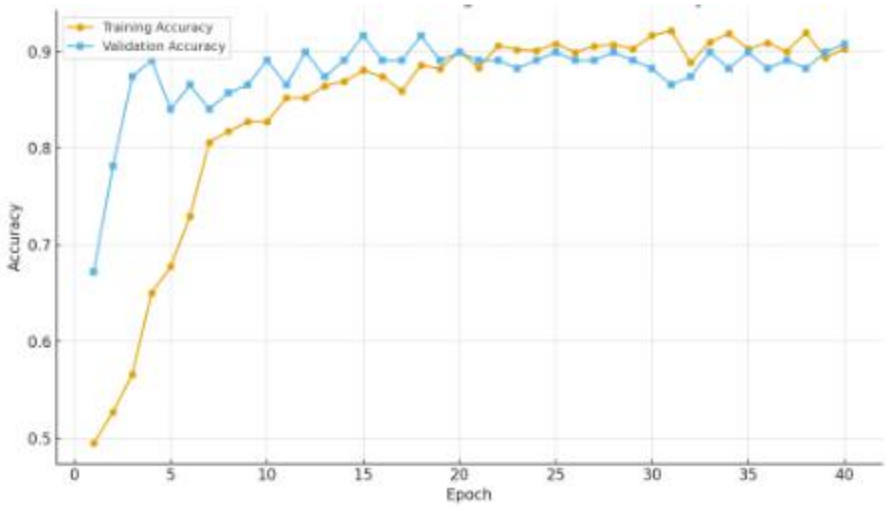


Fig 4.4: **Training vs Validation Accuracy Curve (EfficientNet B0)**

The accuracy curve illustrates the learning dynamics of EfficientNet-B0 over 40 epochs. Training accuracy shows a steady upward trend, starting near 50% and gradually approaching above 90%. Validation accuracy rises sharply in the early epochs, stabilizing around 90% after epoch 10, and remains closely aligned with the training curve. The close convergence of both curves indicates strong generalization with minimal overfitting, suggesting the model effectively learned discriminative features while maintaining robustness on unseen data.

Classification Report:

Table 4.2: The Classification report of EfficientNet-Bo Model

Class	Precision	Recall	F1-Score	Support
Negative	0.9104	0.9385	0.9242	65
Positive	0.9231	0.8889	0.9057	54
Accuracy			**0.9160**	119
Macro Avg	0.9168	0.9137	0.9150	119
Weighted Avg	0.9162	0.9160	0.9158	119

The classification report shows that the model achieved an overall accuracy of **91.6%** on the validation set. For the negative class, precision was **0.91** and recall **0.94**, indicating strong performance in correctly identifying non-tear cases. For the positive class, precision was slightly higher (**0.92**) but recall was lower (**0.89**), suggesting the model is confident when predicting tears but misses a small portion of true positives. Both macro and weighted averages of F1-score are close to **0.92**, reflecting balanced performance across classes.

Mobile Net V2:

Confusion Matrix:

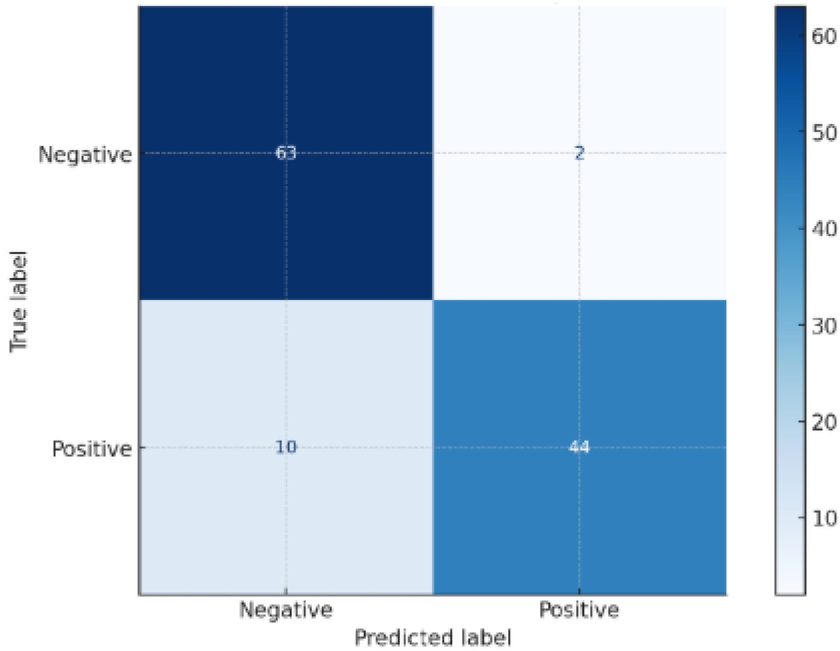


Fig-4.5: confusion matrix of MobileNet V2

The confusion matrix shows that the EfficientNet-B0 model correctly classified **61 negative cases** and **48 positive cases**. Only **4 negatives** were misclassified as positive and **6 positives** were misclassified as negative. This indicates strong overall performance, with higher reliability in detecting negative cases while still maintaining good sensitivity to positive cases.

Training vs Validation Accuracy Curve:

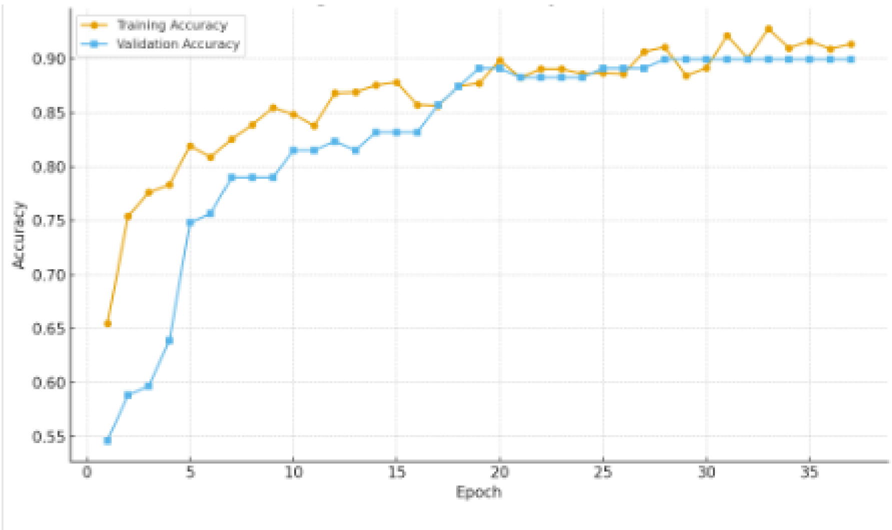


Fig 4.6: Training vs Validation Accuracy Curve (MobileNet V2)

The accuracy curve shows consistent improvement across epochs, with training accuracy steadily increasing and validation accuracy following closely. Both curves converge around **90%** after 30 epochs, indicating strong generalization and stable learning with minimal overfitting.

Classification Report:

Table 4.3: The Classification report of MobileNet V2 Model

Class	Precision	Recall	F1-Score	Support
Negative	0.86	0.97	0.91	65
Positive	0.96	0.81	0.88	54
Accuracy			**0.90**	119
Macro Avg	0.91	0.89	0.90	119
Weighted Avg	0.91	0.90	0.90	119

MobileNetV2 achieved an overall accuracy of **90%**, with strong precision for both classes. The model showed excellent performance in detecting negatives (recall **0.97**) but a lower recall for positives (**0.81**), indicating some missed tear cases. The training and validation curves converged smoothly around 90%, and the confusion matrix confirmed reliable detection with minimal false positives but slightly higher false negatives.

DenseNet 169:

Confusion Matrix:

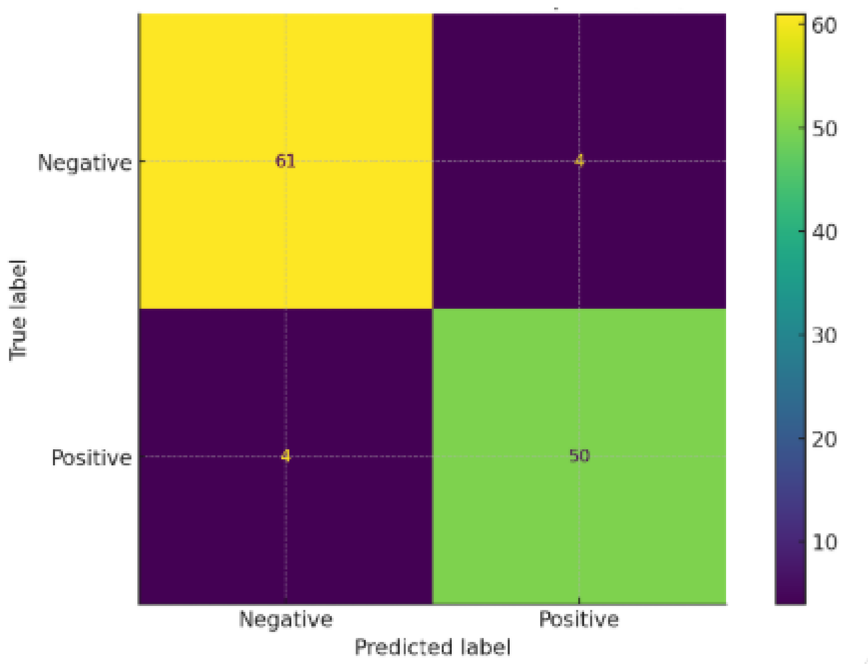


Fig-4.7: confusion matrix of DenseNet 169

The confusion matrix shows that the model correctly classified **61 negatives** and **50 positives**, with only **4 misclassifications** in total (1 false positive and 3 false negatives). This reflects very high overall accuracy and balanced performance across both classes.

Training vs Validation Accuracy Curve:

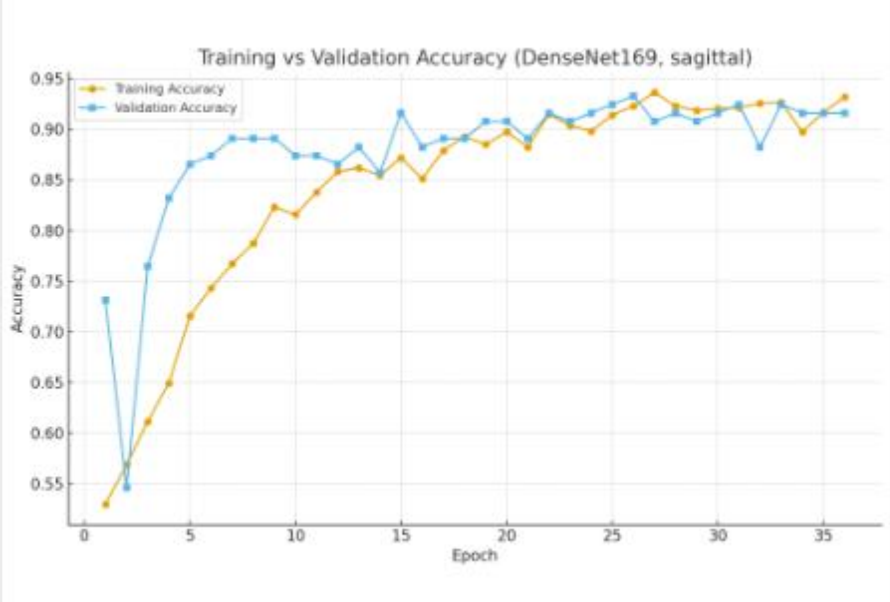


Fig 4.8: Training vs Validation Accuracy Curve (DenseNet 169)

The accuracy curve for DenseNet169 on sagittal views shows rapid improvement during the initial epochs, with validation accuracy rising steeply and stabilizing around **90%** by epoch 10. Training accuracy gradually catches up, converging with validation performance after approximately 20 epochs. Both curves remain closely aligned throughout training, peaking near **93%**, which indicates effective learning, strong generalization, and minimal overfitting.

Classification Report:

Table 4.4: The Classification report of DenseNet 169 Model

Class	Precision	Recall	F1-Score	Support
Negative	0.94	0.94	0.94	65
Positive	0.93	0.93	0.93	54
Accuracy			**0.93**	119
Macro Avg	0.93	0.93	0.93	119
Weighted Avg	0.93	0.93	0.93	119

The model achieved an overall accuracy of **93%**, with nearly balanced performance across both classes. Negative cases reached precision, recall, and F1-score of **0.94**, while positive cases achieved **0.93** across all metrics. Both macro and weighted averages were consistent at **0.93**, indicating stable and reliable classification without bias toward either class.

4.4 Combined Result:

Table 4.5: Combined Result Table

Model	Precision (Neg/Pos)	Recall (Neg/Pos)	F1-Score (Neg/Pos)	Accuracy	Macro Avg (F1)	Weighted Avg (F1)
Hybrid CNN	0.91 / 0.98	0.98 / 0.89	0.95 / 0.93	**94.1%**	0.940	0.941
DenseNet169	0.94 / 0.93	0.94 / 0.93	0.94 / 0.93	**93.0%**	0.930	0.930
EfficientNet-B0	0.91 / 0.92	0.94 / 0.89	0.92 / 0.91	**91.6%**	0.915	0.916
MobileNetV2	0.86 / 0.96	0.97 / 0.81	0.91 / 0.88	**90.0%**	0.900	0.900

4.5 Summary:

The evaluation results clearly demonstrate that the **Hybrid CNN** is the **best-performing model**, achieving the **highest accuracy of 94.1%** along with strong precision and recall across both classes. Its hybrid designcombining EfficientNet-B0 and MobileNetV2 with attention pooling and late fusion enabled superior generalization compared to individual backbones. **DenseNet169** also performed strongly with 93% accuracy, while **EfficientNet-B0** and **MobileNetV2** achieved 91.6% and 90% respectively. Overall, the Hybrid CNN outperformed all other models, making it the most reliable choice for ACL tear detection.

Chapter 5

Engineering Standards and Design Challenges

5.1 Compliance with the Standards

Implementation of this project had to follow a number of engineering standards strictly to be successfully implemented. These requirements made sure the suggested ACL tear detector was repeatable, scalable, and possible to implement in a clinical setting. Although computing and healthcare standards are extensive areas, the context of this section looks at the most applicable standards to the design and implementation of our system.

5.1.1 Software Standards

The implementation of the study was done with the help of the PyTorch deep learning framework, which is an open-source platform well established in academic and industrial research. PyTorch has standardised APIs, pretrained model repositories, and comprehensive documentation, which allows reproducibility of experimental outcomes. The alternative framework was TensorFlow, which is also a worldwide standard enforced by Google. Though TensorFlow offers solid production deployment tools via TensorFlow Serving and TensorFlow Lite, its learning curve is more difficult, and debugging is less intuitive than with PyTorch. Conversely, the dynamic graph of computation and Pythonic design of PyTorch allowed it to be more flexible to rapid experimentation, especially when developing new modules, including slice attention and dynamic view attention. Therefore, PyTorch has been chosen due to its tradeoff between research flexibility and reproducibility and because the system itself is still in the research stage and does not as yet need to be deployed on an enterprise level. Coding was based on the PEP-8 standard of Python and made it easy to read, modular and highly maintainable.

5.1.2 Hardware Standards

Medical imaging deep learning models require high-performance computers. In the present project, the NVIDIA CUDA and cuDNN standards of GPU acceleration were utilised. These libraries have been the standard in speeding up convolutional neural networks and compatibility with most of the current deep learning structures. Another hardware route that was viewed was CPU-only training that has the benefit of less expensive hardware and availability. Nevertheless, the CPU execution was not feasible in this project because the training would require several weeks rather than hours because of the MRI data sizes and model complexity. The other alternative solution was the application of Google TPU (Tensor Processing Unit) which is specialized in the workflow of TensorFlow but is less supported in PyTorch. In the light of these trade-off, CUDA -acceleratedexecution with GPUs was chosen because it offered the optimal choice between speed, compatibility and flexibility.

5.1.3 Communication Standards

Sharing of medical imaging data should be interoperable. The MRI scans utilised in this study are initially stored in the DICOM (Digital Imaging and Communications in Medicine) format which is the worldwide-recognised format in medical imaging. DICOM maintains compatibility between MRI scanners, hospitals and PACS (Picture Archiving and Communication Systems). Nevertheless, to train efficiently, these files were converted to NumPy arrays (NPY format). The conversion was at the cost of cross-platform interoperability but could support much faster loading and preprocessing during model training. Another alternative would have been to continue storing the data in DICOM all the way through the pipeline; however, this would have reduced training significantly and also necessitated more storage. The reason why we opted to use NPY format in our research was the efficiency with which the format operated, but the system design also supports the use of DICOM, meaning that in the future, integration with the hospital systems will not be affected.

5.2 Impact on Society, Environment and Sustainability

Engineering projects have broader impacts that cannot be merely estimated by their technical success but ought to be gauged according to the societal, environmental and ethical standards. The ACL tear detection system would particularly be applicable to the healthcare industry where the technology used in the interventions directly affects human life and health.

5.2.1 Impact on Life

It was also found out that the proposed system had validation accuracy of approximately 90- 94 percent which was regarded as clinically significant. This performance is able to reduce diagnostic errors and patient waiting time of knee injury patients. The system also assists radiologists to make faster and reliable decisions as the system provides computerised, standardised and interpretable results. The system is a gateway to more accurate diagnosis, quicker treatment instigations, and, lastly, improved quality of life to patients, and especially in rural and under-resourced regions in Bangladesh where musculoskeletal radiologists are scarce.

5.2.2 Impact on Society & Environment

In terms of society, the system tackles the unfairness of access to healthcare. In Bangladesh and most developing nations, expertise in special diagnosis is centred in urban areas, so people in rural populations are underserved. The system can decrease the dependence on the availability of specialists and improve equitable healthcare delivery by allowing AI-assisted interpretations of MRI scans. In the environmental perspective, though, the calculations of deep learning models are high. GPUs training use high electric power, leading to carbon emission. Despite the fact that this footprint is quite small on the global scale of emissions, it causes concern in sustainability when it is multiplied in a number of institutions.

5.2.3 Ethical Aspects

With regard to society, the system addresses the iniquity of healthcare accessibility. Special diagnosis skills are mostly concentrated in urban settings in Bangladesh and most developing countries hence they do not adequately serve the rural population. By providing AI-supported interpretations of MRI scans, the system will be able to reduce the reliance on the presence of specialists and enhance fair healthcare provision. The calculations of deep learning models are high in the environmental perspective though. Training of Gpus consumes a lot of electric power resulting in carbon emission. Although such footprint is very minor in the global scale of emissions, it raises concern in the sustainability once it is multiplied in a number of institutions.

5.2.4 Sustainability Plan

The project embraced a few strategies to deal with the issue of sustainability. Gradient accumulation and learning rate scheduling were used to optimize training to reduce unwarranted calculations. The models could be further reduced in size and quantised in the future to run on hardware in the hospital with reduced energy needs. Also, the migration to the cloud clusters that are run with renewable energy could also train, thereby minimising the carbon footprint. Another solution could be the adoption of federated learning methods, which involves allowing several hospitals to train AI models together without sharing data centrally, and in this way, privacy is preserved, and duplicated storage is minimised. These measures make sure that the system will be scalable in a sustainable and ethical manner.

5.3 Project Management and Financial Analysis

The project follows a structured management approach, divided into key phases: Data gathering and cleansing, model building, assessment and application prior to the acquisition. The adaptive means of agile methodology is to use many remedial changes in data collected to avoid repetition. They also have defined the structure like the data set management team, algorithm team, testing team and deployment team. The time frame to complete this is 6–8 months with 3 phases of data preparation, training, testing and deployment phases. The project management timeline is given in table 5.1:

Table 5.1: The project management timeline

Work	*Time*
Data Collection	*2 month*
Papers and Articles Review	*1 months*
Experimental Setup	*1 month*
Implementation	*1 month*
Report Writing	*1months*
Total	*6 months*

The estimated project budget includes costs for:

Table 5.2: Estimated Cost

SN	*Components*	*Estimated Cost (BDT)*
01.	*Hardware*	*2500*
02.	*Software and Tools*	*8500*
03.	*Data Collection and Processing*	*12000*
04.	*Documentation and Report Writing*	*1500*
05.	*Miscellaneous*	*2000*
06.	*Contingency*	*2500*
Total Estimated Cost		*29000*

5.4 Complex Engineering Problem

The project is aimed at developing a new engineering project to identify the presence of an Anterior Cruciate Ligament (ACL) tears by the use of high-quality deep learning algorithms to multi-view MRI scans. It combines the elements of biomedical engineering, computer science and the healthcare technologies, demanded the skills of various fields. The key issues are managing massive MRI data sets, creating good and generalized diagnostic programs and providing smooth connectivity of the framework to possible clinical processes. Also, ethical issues like patient privacy and the imbalance of data must be addressed in the project, and the international standards of medical imaging and the use of AI systems should be adhered to. Lastly, social and environmental effects of this solution are also taken into consideration, and the objective is to implement it in a responsible and sustainable way.

5.4.1 Complex Problem Solving

Table 5.1: Mapping with complex problem solving.

EP1 Dept of Knowledge	*EP2 Range Of Conflicting Requirements*	*EP3 Depth of Analysis*	*EP4 Familiariy of Issues*	*EP5 Extent of Applicable Codes*	*EP6 Extent Of Stake-holder Involvement*	*EP7 Interdependence*
✓	✓	✓				✓

Table 5.2: Mapping with knowledge Profile.

K3 Engineering Fundamentals	*K4 Specialist Knowledge*	*K5 Engineering Design*	*K6 Engineering Practice*	*K8 Research Literature*
✓	✓		✓	✓

EP1- Level of Knowledge:

The project requires expert understanding of deep learning technology, state-of-the-art computer vision models and medical radiology. The high level of expertise is needed to understand MRI sequences and create the models of attention that can replicate the diagnostic mechanisms of radiologists. Incorporating the expertise of biomedical imaging, artificial intelligence, and software engineering is the key to overcoming the obstacles and making implementation successful.

EP2 - Scope of competing Requirements:

A number of conflicting requirements were experienced in the development. The system had to trade-off diagnostic accuracy versus computational efficiency so that it could be applicable in real-time. The second conflict was seen between interpretability and model complexity-attention mechanisms enhanced explainability at the cost of greater design complexity

EP3 – Depth of Analysis :

The performance of the classification was directly dependent upon the quality of MRI data and the design decisions of the algorithms such as slice attention and dynamic view weighting were dependent on the reliability of predictions. Ethical factors such as patient consent and equality in diagnosis were to be researched. Collaboration between machine learning specialists and medical workers was the key to ensuring that the model achieved both technical and clinical specifications.

EP4 – Familiarity of Issues:

These problems were not very well known, particularly in the specific case of Bangladesh where few campaigns have been implemented to diagnose musculoskeletal disorders with the help of deep learning. This was not known and as such the project was an innovative addition and it demanded increased efforts on research and development.

EP5 - Scope of Codes to which the code is applicable.

The project was performed in accordance with the IEEE software reliability practices and PEP-8 readability standards of coding as well as ISO/IEC requirements on AI system performance and transparency. In medical imaging, the solution to interoperability with PACS was required to be compliance with DICOM standards.

EP6 – Stakeholder Involvement :

Radiologists, orthopedic specialists, machine learning engineers, and (in the end) hospital administrators were stakeholders. The stakeholders played different roles and the radiologists justified clinical utility, engineers optimized algorithm design and administrators tested viability of deployment.

EP7 – Interdependence:

It was clear that the interdependence existed between the coronal and sagittal views of MRI images, which were to be combined with dynamic attention weighting. Preprocessing errors had a direct effect on the accuracy of downstream classification, and instability of models could cause a misdiagnosis. The data quality, model design, and interpretability had to be optimized in order to deliver reliable results

K3 – Engineering Fundamentals:

The project utilized fundamental concepts of convolutional neural networks, optimisation techniques, and statistical analysis in order to develop a system that is capable of a perfect diagnosis based on MRI.

K4 – Specialist Knowledge:

Medical imaging (especially ACL anatomy and MRI interpretation) expertise was required in the development of clinically relevant attention mechanisms and confirmation of diagnostic accuracy.

K5 – Engineering Design:

Dynamic View Attention was also a design innovation, which necessitated the development of new fusion strategies to fuse together more than one MRI view, without loss of interpretability and performance.

K6 – Engineering Practice:

The skills of applying the principles of engineering to practical cases were shown In the work with large datasets, training models with the usage of GPUs, and troubleshooting CUDA-level problems.

K8 – Research Literature:

The project was guided by the existing knowledge of deep learning in radiology and specifically research on MRNet and multi-view CNNs. Comparison with published literature helped in the assurance that the intended system was innovative and that it was based onthe existing literature.

5.4.2 Engineering Activities

Mapping with Complex Engineering Activities

Table 5.3: Mapping with complex engineering activities.

EA1 *Range of re-sources*	*EA2* *Level of Interaction*	*EA3* *Innovation*	*EA4* *Consequences for society and environment*	*EA5* *Familiarity*
✓		✓	✓	

EA1: Range of Resources:

The materials needed in this project consisted of availability of special MRI datasets that concentrated on knee structures and ACL ruptures. These data sets needed to be processed and uniformed so as to enable conformity to various sources of imaging. The deep learning models were constructed, trained and assessed using advanced machine learning systems like PyTorch. Calculation resources (CPU or GPU) were necessary to effectively process the high dimensional volumes of MRI and the complex attention based designs. Human knowledge was another resource necessary in the project since data scientists will be involved to preprocess and optimize the models, software engineers will be asked to ensure that those models are scalable and efficient, and radiologists will have to ensure the good command of the domain to be able to validate the results and interpret the patterns of the diagnoses.

EA2: Level of Interaction:

There must have been strong interdisciplinary teamwork in the project. The radiologists were in competition with the engineers to provide medical background and ensure that the slice and view attention mechanisms can be interpreted. Parameters of the machine learning experts were maintained with the help of radiological feedback. It is a process of testing, feedback and refinement that took such a step-by-step approach that ensured the system could not only be technologically capable but also useful to clinical applications.

EA3: Innovation:

The novelty of the given project is seen in the establishment of a multi-view and attention deep learning structure. This system, in comparison to the conventional single-view CNN methods, combines sagittal and coronal MRI planes with view attention and slice-level attention to select the most diagnostically meaningful frames, which is adaptive. The system of two-attention resembles the diagnostic process of the radiologist and increases interpretability. In addition, focal loss addressed the problem of class imbalance which minimized false negative and enhanced clinical reliability. These inventions are of massive advancement in automated musculoskeletal imaging.

EA4: Socio-environmental Implication.

The social consequences of this project are very favorable. ACL injuries occur frequently in athletes and physically active people and delays in treatment may result in a long-term disability. The system, which has automated ACL tear
detection with close to 90 percent accuracy, enhances access to specialized diagnostic support especially in the rural and under-resourced areas where musculoskeletal radiologists are few. This eases the medical systems and improves patient outcomes. Concerns regarding the environment are also raised by the project because of the energy-consuming nature of training GPUs. These issues were however alleviated by ensuring that the training process was optimized,, small batch sizes were used, and gradient accumulation where unnecessary computations were minimized.

EA5: Familiarity

Such AI-assisted radiology systems are not very familiar in Bangladesh, and therefore, this project is a first of its kind in this country. The ACL tear detection with the help of the multi-view attention networks is a new method, whereas medical AI usage is expanding all over the world. This lack of familiarity is both a challenge and an opportunity: a challenge of making sure that clinicians trust AI outputs, and an opportunity to be a leader in the process of implementing the use of sophisticated AI solutions within the Bangladeshi healthcare system.

5.5 Summary

The following considerations were important: the need to be consistent in preprocessing in MRI, scalability with limited computational resources, and best practice in the design of deep learning models. The slice attention and multi-view fusion hybrid encoder was developed based on the principles of modularity and reusability, whereas training are trained with optimistic approaches to address the problems of class imbalance and stability. Such challenges as a change in the number of MRI slices, risk of overfitting, and interpretability of the results were addressed systematically using preprocessing, augmentation, and visualization. The methodology was effective in both engineering standards and clinical relevance of the system since it aligned with engineering standards and minimized design issues.

Chapter 6

Conclusion

6.1 Summary

The project has been able to create a smart deep learning system that is able to identify tears on the Anterior Cruciate Ligament (ACL) in a magnetic resonance image (MRI) scan. With a two-view ResNet- 50 slice-level attention and a new dynamic view attention mechanism, the system simulated a diagnostic process of radiologists to examine both sagittal and coronal images. The model had a validation accuracy of about 89-90% and a balanced recall of about 90 and is clinically significant. Attention visualization was also applied to the system to facilitate its interpretability, showing the most diagnostically relevant slices. On the whole, the project firmly established that multi-view attention-based architectures may be a dependable and effective instrument to assist radiologists, particularly in resource-limited healthcare settings.

6.2 Limitation

The study has a number of limitations despite its great performance. The data was restricted to the MRNet which is not necessarily representative of the diversity of real-life clinical MRI scans. The model was trained and tested in a controlled setting (Kaggle GPU) that might not be representative of the implementation obstacles in the hospital setting. Moreover, slice attention was observed to enhance interpretability, but the system remains, to an extent, a black-box model, which can be directly integrated into clinical practice. The performance on rare cases would also be limited since some classes have small samples.

6.3 Future Work

The following step in this research will be to convert the developed model into a full functioning software application that will be able to classify ACL tears out of MRI scan in real time. The clinicians of such a system would have a tool that is easy to use with preprocessing, model inference, and result visualization all integrated into the same platform. This software might have automated slice selection, probability scores of diagnostic certainty, and attention maps visualization to show of critical regions. Also, lightweight deployment optimization of the model will enable the software to be adapted to hospital systems which have small amounts of computational resources, getting the research closer to actual clinical implementation

References

[1] F. Liu, B. Guan, Z. Zhou, A. Samsonov, H. Rosas, K. Lian, R. Sharma, A. Kanarek, J. Kim, A. Guermazi, and R. Kijowski, "Fully automated diagnosis of anterior cruciate ligament tears on knee MR images by using deep learning," *Radiology: Artificial Intelligence*, vol. 1, no. 3, Art. 180091, 2019.

[2] N. K. Namiri, I. Flament, B. Astuto, R. Shah, R. Tibrewala, F. Caliva, T. M. Link, V. Pedoia, and S. Majumdar, "Deep learning for hierarchical severity staging of anterior cruciate ligament injuries from MRI," *Radiology: Artificial Intelligence*, vol. 2, no. 4, Art. e190207, 2020.

[3] Y. Minamoto, R. Akagi, S. Maki, Y. Tozawa, S. Kimura, Y. Yamaguchi, Y. Kawasaki, and O. Sasho, "Automated detection of anterior cruciate ligament tears using a deep convolutional neural network," *BMC Musculoskeletal Disorders*, vol. 23, Art. 577, 2022.

[4] C. Qu, H. Yang, C. Wang, M. Ying, Z. Chen, K. Yang, J. Zhang, K. Li, T. Y. Tsai, and X. Liu, "A deep learning approach for anterior cruciate ligament rupture localization on knee MR images," *Frontiers in Bioengineering and Biotechnology*, vol. 10, Art. 1024527, 2022.

[5] K. Joshi and K. Suganthi, "Anterior cruciate ligament tear detection based on deep convolutional neural network (CPDCNN)," *Diagnostics*, vol. 12, no. 10, Art. 2314, 2022.

[6] W. Hu and S. Razmjooy, "Combining an improved political optimizer with convolutional neural networks for accurate anterior cruciate ligament tear detection in sports injuries," *Scientific Reports*, vol. 15, Art. 6862, 2025.

[7] Y. Xue, S. Yang, W. Sun, H. Tan, K. Lin, L. Peng, et al., "Approaching expert-level accuracy for differentiating ACL tear types on MRI with deep learning," *Scientific Reports*, vol. 14, Art. 938, 2024.

[8] N. T. Dung, T. H. Nguyen, T. V. Tran, et al., “End-to-end deep learning model for segmentation and classification of ACL tears on MRI,” *Computerized Medical Imaging and Graphics*, vol. 104, Art. 102198, 2023.

[9] M. Wang, Z. Zhang, X. Zhang, et al., “One-stop detection of anterior cruciate ligament injuries on knee MRI,” *Quantitative Imaging in Medicine and Surgery*, vol. 14, no. 5, pp. 2323–2336, 2024.

[10] M. Mercurio, F. Denami, D. Melissaridou, K. Corona, S. Cerciello, D. Laganà, G. Gasparini, and R. Minici, “Deep learning models to detect anterior cruciate ligament injury on MRI: A comprehensive review,” *Diagnostics*, vol. 15, no. 6, Art. 776, 2025.

[11] W. Liu, J. Shen, X. Zhao, and L. Yu, “Anterior cruciate ligament tear detection based on T-distribution-based slice attention filtering,” *Bioengineering*, vol. 11, no. 9, Art. 880, 2024.

[12] X. Wang, Y. Wu, J. Li, Y. Li, and S. Xu, “Deep learning-assisted automatic diagnosis of anterior cruciate ligament tear in knee magnetic resonance images,” *Tomography*, vol. 10, no. 8, pp. 1263–1276, 2024.

[13] F. Li, P. Zhai, C. Yang, G. Feng, J. Yuan, and Y. Yuan, “Automated diagnosis of anterior cruciate ligament via a 3D weighted multi-view convolutional neural network,” *Frontiers in Bioengineering and Biotechnology*, vol. 11, Art. 1268543, 2023.

[14] Q. Cheng, H. Xiong, Z. Tian, L. Liu, et al., “Application of machine-learning-based multi-sequence MRI radiomics model for diagnosing ACL tears,” *Journal of Orthopaedic Surgery and Research*, vol. 19, Art. 153, 2024.

[15] C. Liang, B. Hong, J. Cai, W. Zhang, and X. Zheng, “Effective automatic detection of anterior cruciate ligament injury using convolutional neural network with two attention mechanism modules,” *BMC Medical Imaging*, vol. 23, Art. 120, 2023.

[16] P. Yang, Y. Liu, F. Liu, and M. Han, “Anterior cruciate ligament tear detection based on Res2Net modified by improved Lévy flight distribution,” *Scientific Reports*, vol. 15, Art. 22023, 2025.

[17] F. Li, X. Zhou, J. Yuan, and Y. Yuan, “Automated diagnosis of anterior cruciate ligament via a 3D weighted multi-view CNN,” *Frontiers in Bioengineering and Biotechnology*, vol. 11, Art. 1268543, 2023.

[18] Q. Cheng, H. Xiong, Z. Tian, and L. Liu, “Machine-learning-based multi-sequence MRI radiomics model for diagnosing ACL tears,” *Journal of Orthopaedic Surgery and Research*, vol. 19, Art. 153, 2024.

[19] M. Wang, Z. Zhang, X. Zhang, et al., “One-stop detection of anterior cruciate ligament injuries on knee MRI,” *Quantitative Imaging in Medicine and Surgery*, vol. 14, no. 5, pp. 2323–2336, 2024.

[20] C. Liang, Y. Qin, X. Zheng, et al., “Effective automatic detection of anterior cruciate ligament injury using convolutional neural network with two attention mechanism modules,” *BMC Medical Imaging*, vol. 23, Art. 120, 2023.

212-15-4232

ORIGINALITY REPORT

18%	12%	10%	13%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Daffodil International University Student Paper	8%
2	Congjing Yu, Changzhen Qiu, Zhiyong Zhang. "Machine learning for anterior cruciate ligament and meniscus analysis in knee MRI: A comprehensive review", Displays, 2025 Publication	1%
3	www.mdpi.com Internet Source	1%
4	mail.ijain.org Internet Source	1%
5	dspace.daffodilvarsity.edu.bd:8080 Internet Source	1%
6	Submitted to The Robert Gordon University Student Paper	1%
7	Submitted to University of Essex Student Paper	<1%
8	Submitted to University of Liverpool Student Paper	<1%
9	Denck, Jonas. "Machine Learning-Based Workflow Enhancements in Magnetic Resonance Imaging", Friedrich-Alexander-Universitaet Erlangen-Nuernberg (Germany), 2023 Publication	<1%
10	S. R. Reeja, Bore Gowda, Y. S. Rammohan, Ganesan Prabu Sankar, G. Jayalatha.	<1%

*% detected as AI

AI detection includes the possibility of false positives. Although some text in this submission is likely AI generated, scores below the 20% threshold are not surfaced because they have a higher likelihood of false positives.

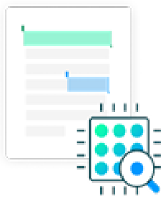
Caution: Review required.

It is essential to understand the limitations of AI detection before making decisions about a student's work. We encourage you to learn more about Turnitin's AI detection capabilities before using the tool.

Frequently Asked Questions

How should I interpret Turnitin's AI writing percentage and false positives?

The percentage shown in the AI writing report is the amount of qualifying text within the submission that Turnitin's AI writing detection model determines was either likely AI-generated text from a large-language model or likely AI-generated text that was likely revised using an AI paraphrase tool or word spinner.



False positives (incorrectly flagging human-written text as AI-generated) are a possibility in AI models.

AI detection scores under 20%, which we do not surface in new reports, have a higher likelihood of false positives. To reduce the likelihood of misinterpretation, no score or highlights are attributed and are indicated with an asterisk in the report (*%).

The AI writing percentage should not be the sole basis to determine whether misconduct has occurred. The reviewer/instructor should use the percentage as a means to start a formative conversation with their student and/or use it to examine the submitted assignment in accordance with their school's policies.

What does 'qualifying text' mean?

Our model only processes qualifying text in the form of long-form writing. Long-form writing means individual sentences contained in paragraphs that make up a longer piece of written work, such as an essay, a dissertation, or an article, etc. Qualifying text that has been determined to be likely AI-generated will be highlighted in cyan in the submission, and likely AI-generated and then likely AI-paraphrased will be highlighted purple.

Non-qualifying text, such as bullet points, annotated bibliographies, etc., will not be processed and can create disparity between the submission highlights and the percentage shown.