Medical Image Analysis
With Limited Data

Dr. Ata Jodeiri
Medical Bioengineering Department
Tabriz University of Medical Sciences

Roadmap to AI Excellence

- Establish National AI Research Centers partner with global institutions to accelerate knowledge transfer and expertise building.
- Invest in AI Education and Workforce Development
- Develop Strategic AI Policies and Ethical Frameworks including data sharing and intellectual property protection.
- Provide funding and infrastructure for startups in healthcare, agriculture, and smart cities.
- Build a National AI Ecosystem with Global Outreach

AI in Medical Imaging

Benefits

- Improved accuracy and speed of diagnosis
- Reduced healthcare costs
- Enhanced patient outcomes
- Improved efficiency
- Frees Up doctor’s time
- Better utilization of resources
- Improved patient experience

AI in Medical Imaging

Challenges and Limitations

- Data quality and quantity
- Ethics and privacy concerns
- Regulatory and legal issues
- Technical limitations
- Lack of interpretability

AI in Medical Imaging

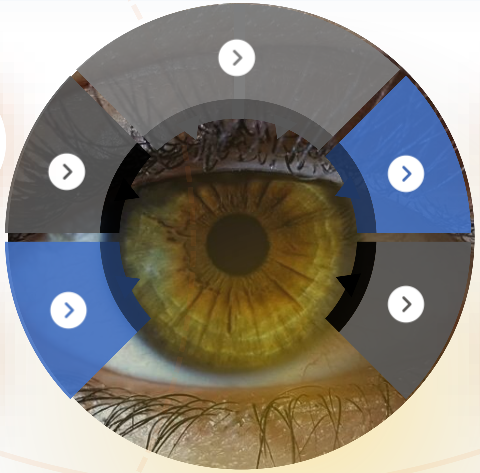
Future

- Integration with electronic health records (EHRs)
- Advancements in image acquisition technology
- Continued development of AI algorithms
- Personalized medicine
- Improved patient outcomes

ADVANCED AI MODELS FOR AMD CLASSIFICATION USING OCT IMAGES

OUTLINE

- Introduction (Age-related Macular degeneration)
- AI role in medical imaging
- AI in age-related Macular degeneration AMD
- Our research methods and results
- Final massage takeaway AI tips



INTRODUCTION

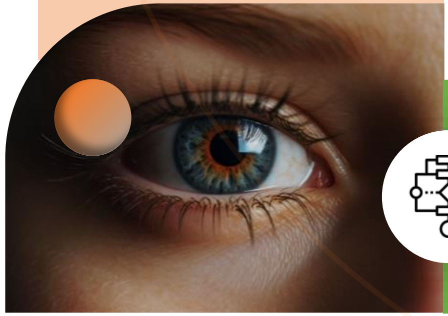
AI AND ML IN MEDICINE:

Many medical applications use artificil intellegence and machine learing to ease solve itriguing problems, reach accurate medical decision or invistaigate certian aspects.
(e.g. Segmentation, disease prediction)

Deep learning's impact on ophthalmology.

An example in research is the use of deep learning in opthamology as it aided in segmentation of images and predicted disease such as age-related macular degeration.

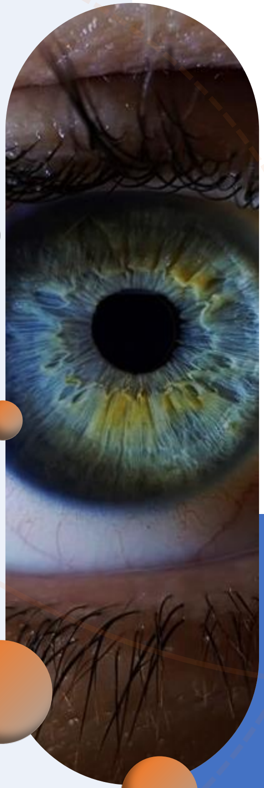
Importance of early diagnosis in AMD.



ML algorithms demonstratied high accuracy rates in identifying drusen and RPE abnormality[1,2]

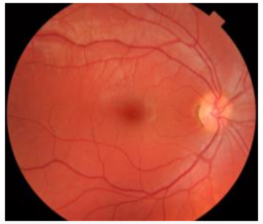
OVERVIEW OF AGE-RELATED MACULAR DEGENERATION (AMD)

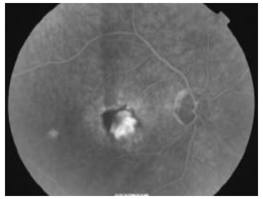
- What is AMD?
- An eye condition that can blur central vision occurs when aging damages the macula (the part of the eye responsible for clear, direct vision.)
- Global impact and prevalence.

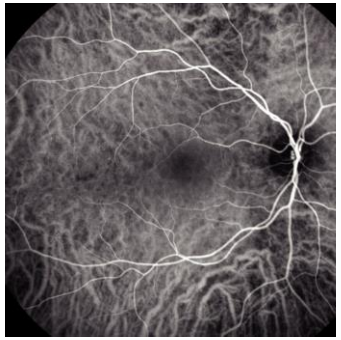


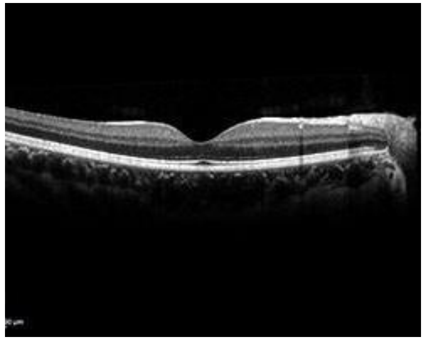
- AMD is one of the principal causes of loss of sight globally, affecting roughly 8.7% of people worldwide (it is affecting 2.5 M in Canada alone in 2023).
- Diagnostic techniques for AMD.
- Fundus photography, fluorescein, indocyanine green angiography (ICG), and optical coherence tomography

OVERVIEW OF AGE-RELATED MACULAR DEGENERATION (AMD)


Fundus photography,


Fluorescein


ICG


OCT

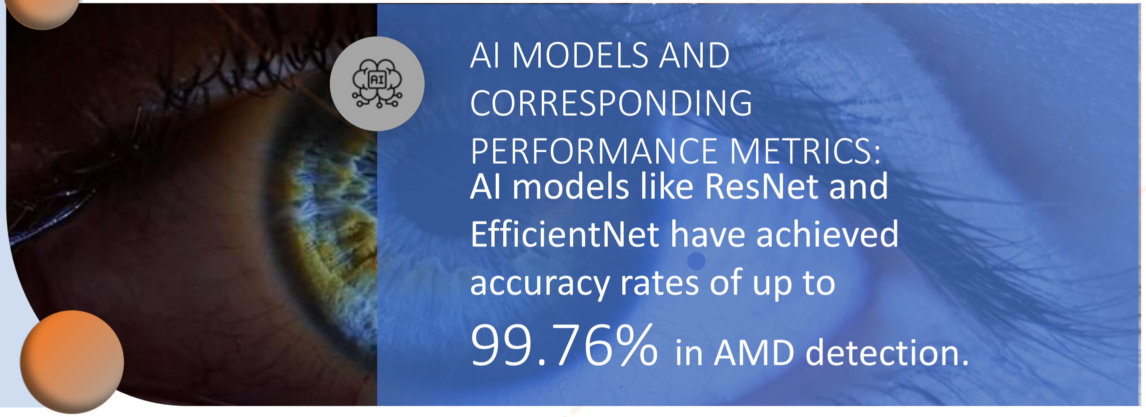
AI IN AMD DIAGNOSIS

AI'S ROLE IN EARLY DETECTION OF AMD.:

deep learning such as convolutional neural networks (CNNs), analyse retinal images and identify key features of AMD, like drusen and Retinal Pigment Epithelium (RPE) abnormalities, with high accuracy.

SUCCESSES AND CHALLENGES IN AI APPLICATIONS FOR AMD:

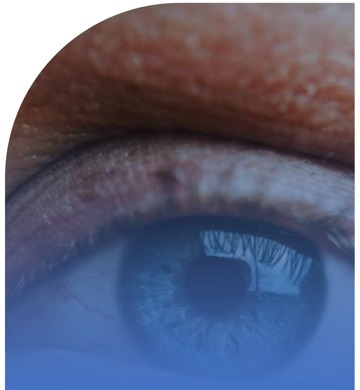
AI (i.e. CNN) has successfully enhanced early AMD detection with high accuracy but faces challenges like occasional misclassification and the need for further refinement and integration into clinical practice.



AI MODELS AND CORRESPONDING PERFORMANCE METRICS:

AI models like ResNet and EfficientNet have achieved accuracy rates of up to 99.76% in AMD detection.

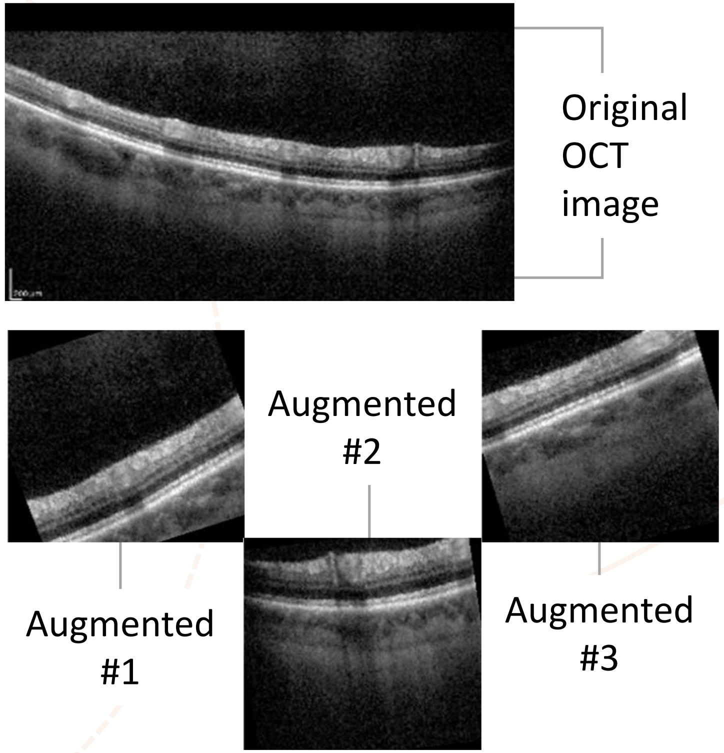
DATASET AND PREPROCESSING



Description of the Noor dataset

Data distribution (Normal, Drusen, choroidal neovascularization CNV).

Data preprocessing steps.



Artistic Break

WATERLILY POND AND JAPANESE FOOTBRIDGE 1899

THE JAPANESE FOOTBRIDGE 1922

DEEP LEARNING MODELS EMPLOYED

OVERVIEW OF THE MODELS USED:

ResNet, EfficientNet, EfficientNet with Attention and ensembled model

ADVANTAGES OF EACH MODEL:

ResNet excels at deep pattern recognition, EfficientNet balances efficiency with accuracy, and EfficientNet with Attention improves focus on critical details for better accuracy.

Ensebled model:

maximizes the strengths of individual models while mitigating their potential weaknesses.

RESULTS – PERFORMANCE METRICS

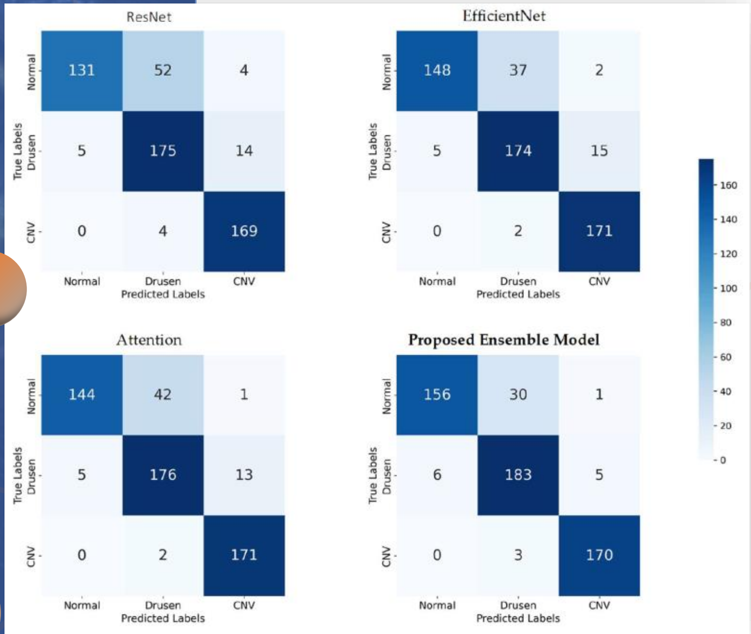
Summary of performance metrics: The proposed ensemble model outperformed other methods, achieving the highest F1 scores of 89% for Normal, 89% for Drusen, and 97% for CNV, with an overall accuracy of 92%.

PROPOSED ENSEMBLE MODEL CONFUSION MATRIX:

Normal:
The ensemble model correctly classified 156 instances as Normal, with 30 misclassified as Drusen and 1 as CNV.

Drusen:
183 instances were correctly classified as Drusen. There were 6 instances misclassified as Normal and 5 as CNV.

CNV:
170 instances of CNV were correctly classified, with 3 misclassified as Drusen.

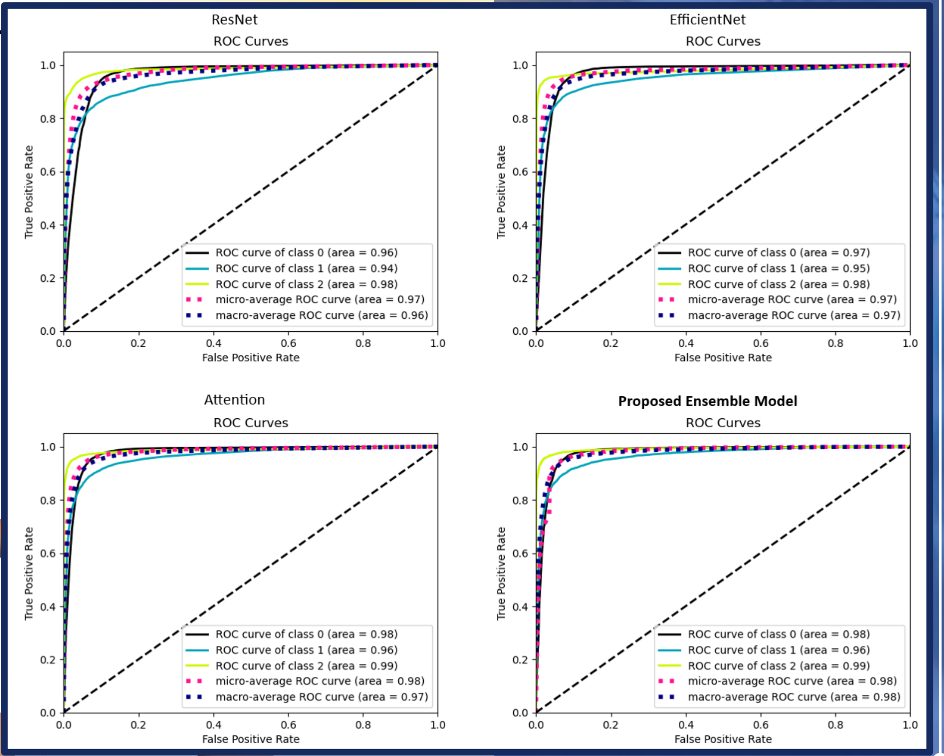


RESULTS - ROC

PROPOSED ENSEMBLE MODEL ROC CURVE:

Class 0 (AUC = 0.98): **Consistently excellent performance.**

Class 1 (AUC = 0.96): **Maintains high accuracy.**



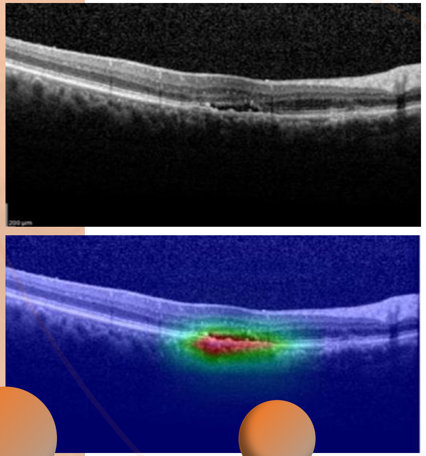
Class 2 (AUC = 0.99): **Outstanding classification accuracy.**

Micro-Average ROC (AUC = 0.98): **Highest overall performance among models.**

Macro-Average ROC (AUC = 0.98): **Most balanced and effective performance across all classes.**

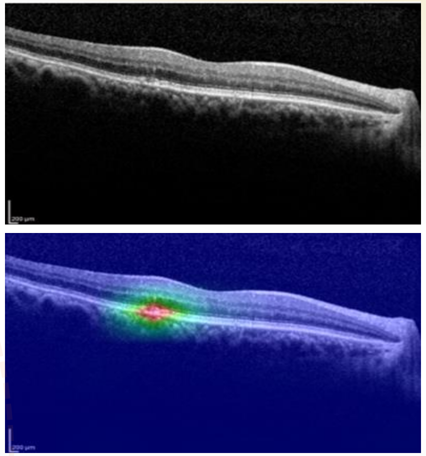
CLASS ACTIVATION MAPS (CAMS)

CNV Detection:
CAMs showed our models pinpointing critical CNV features like neovascular membranes and fluid accumulation in OCT images.



Class:
CNV
Prediction:
CNV

Class Activation Map for CNV



Class:
DRUSEN
Prediction:
DRUSEN

Class Activation Map for DRUSEN

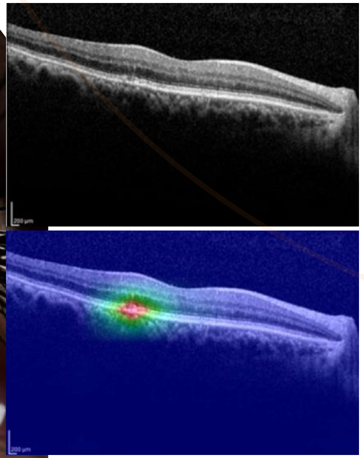
Drusen detection :
CAMs effectively highlighted drusen deposits and retinal changes, aiding precise classification.

Enhanced model interpretability:
CAMs provided insights into model decisions, confirming focus on clinically relevant areas in OCT images.

CLASS ACTIVATION MAPS (CAMS)

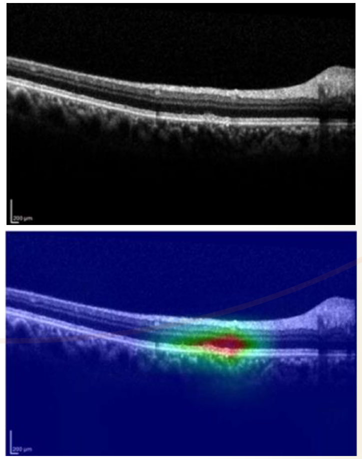
How CAMs provide insights into model decision-making.

Examples of CAMs for CNV and Drusen.



Class: DRUSEN
Prediction: DRUSEN

Class Activation Map for DRUSEN

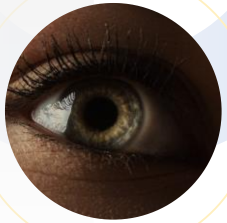


Class: DRUSEN
Prediction: DRUSEN

Class Activation Map for DRUSEN

COLLABORATIVE ERROR ANALYSIS AND DATASET REFINEMENT

Importance of accurate labelling in medical datasets.



Process and outcomes of the collaborative error analysis..

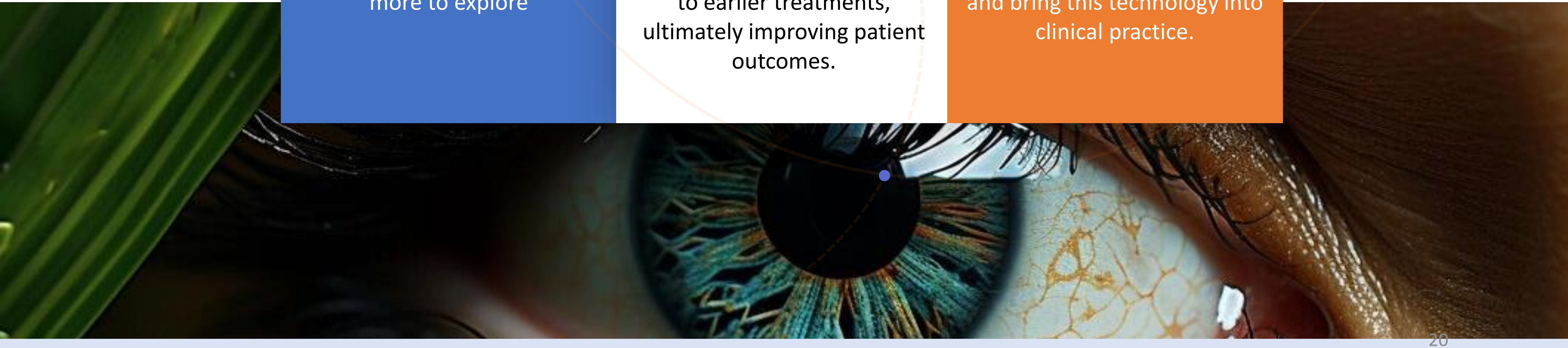
Improvement in model accuracy after refinement.

CONCLUSION AND FUTURE IMPLICATIONS

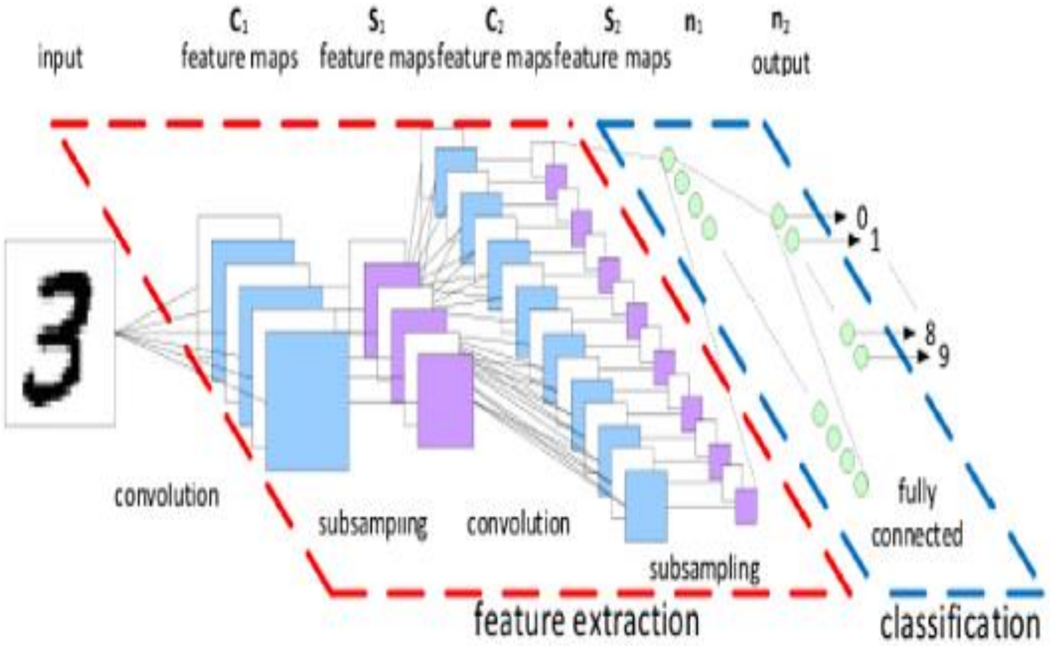
Our ensemble model has set a new benchmark for AMD classification, but there’s still so much more to explore

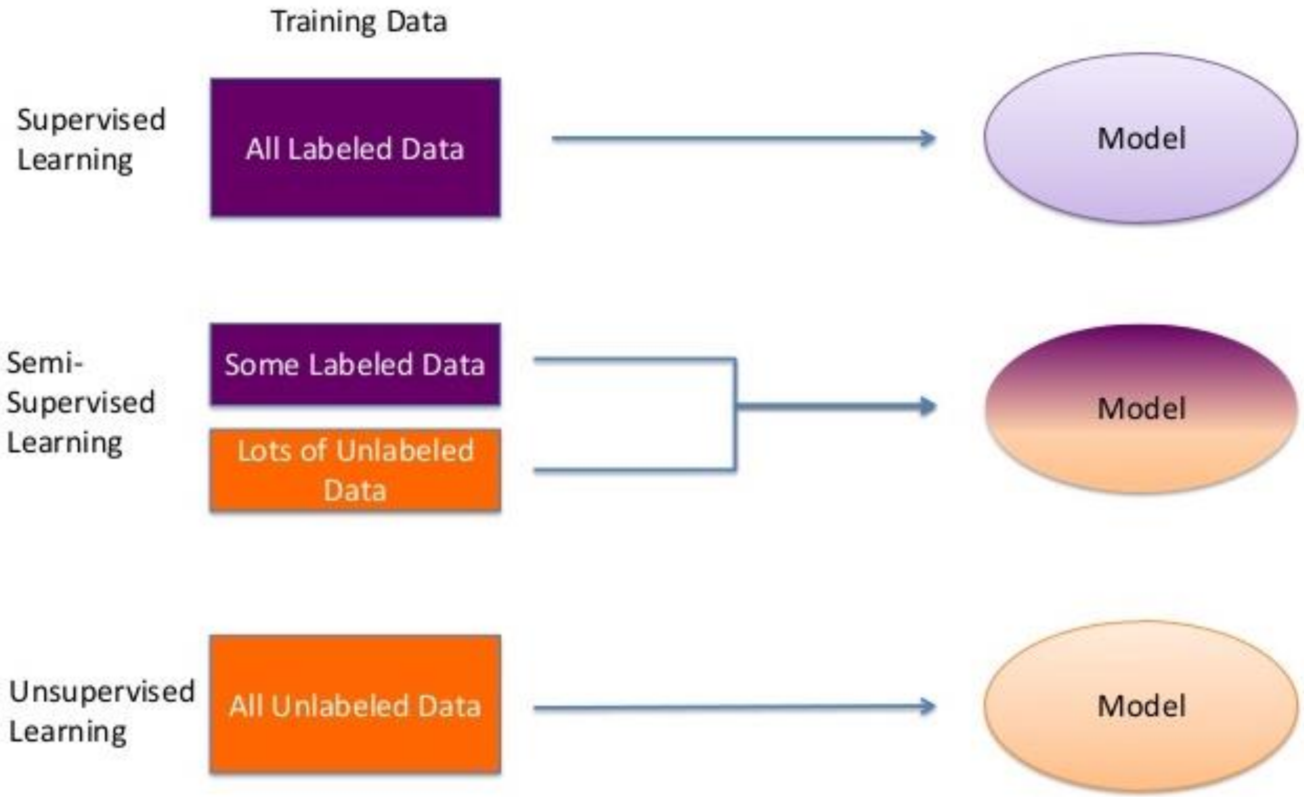
The implications for ophthalmology are profound—better, more accurate diagnoses can lead to earlier treatments, ultimately improving patient outcomes.

Future directions for research and clinical application: we aim to refine these models further, explore additional datasets, and bring this technology into clinical practice.



AI Applications in Diagnostic:

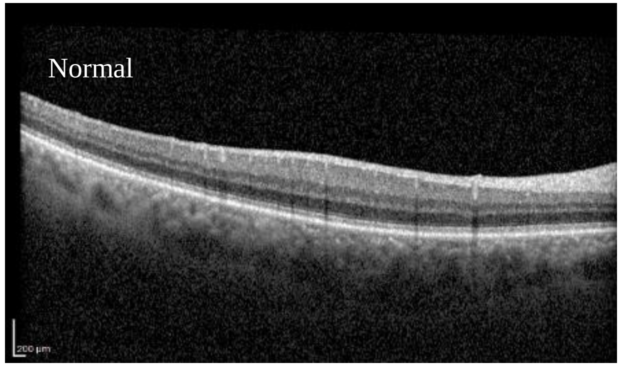


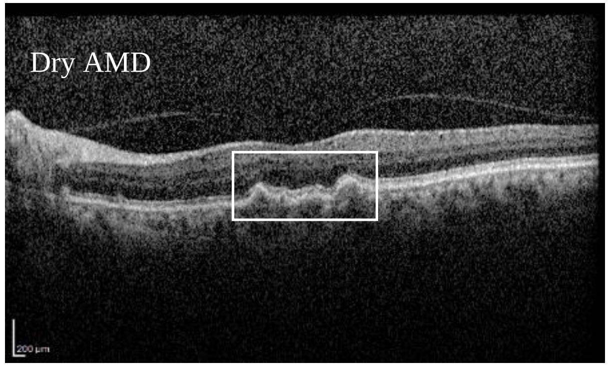


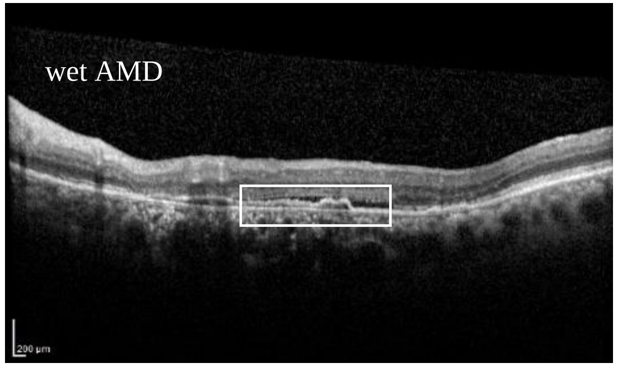
Objectives:

1. Supervised model design & development for AMD detection
2. Choosing supervised optimized parameters
3. Semi-supervised model design & development to obtain SL results
4. Results evaluation: accuracy, precision,…

Dataset:



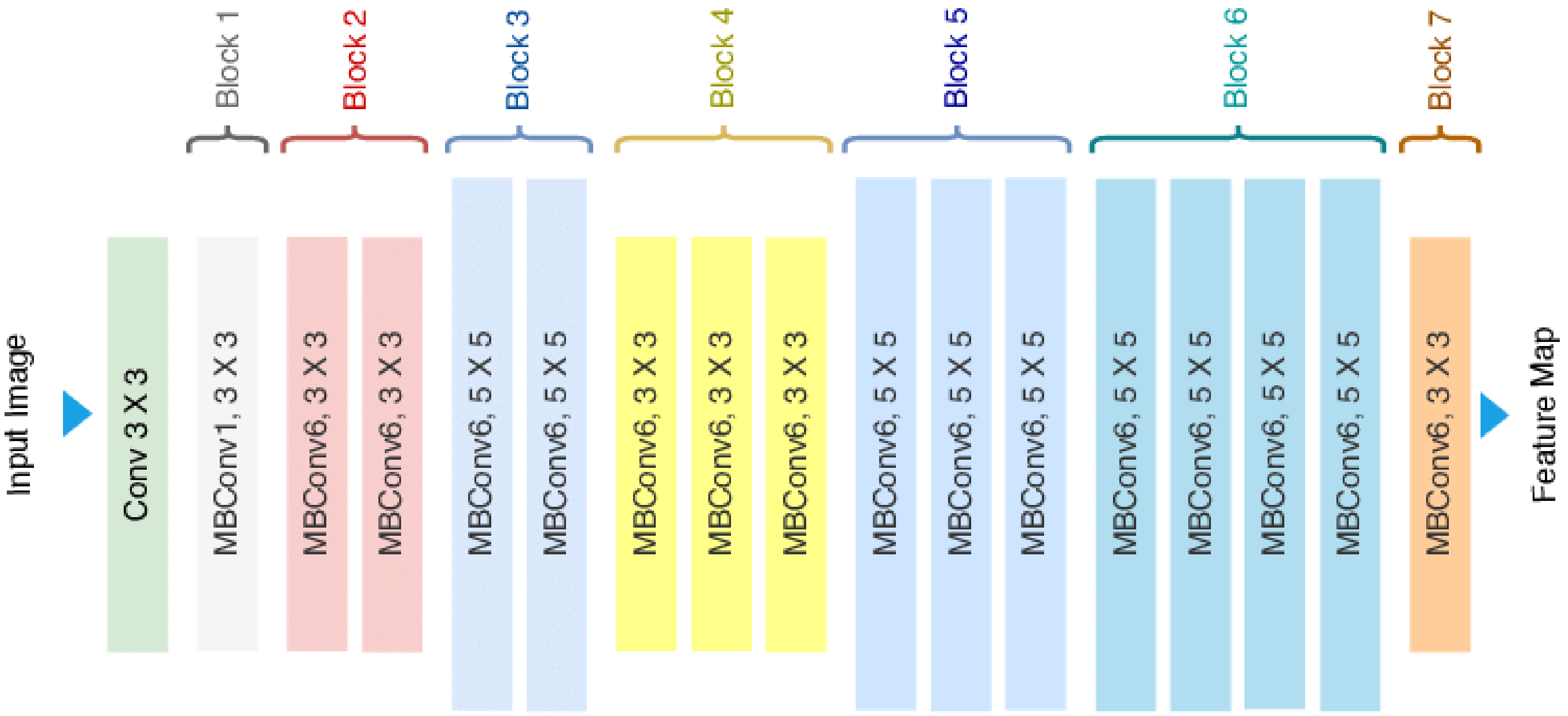




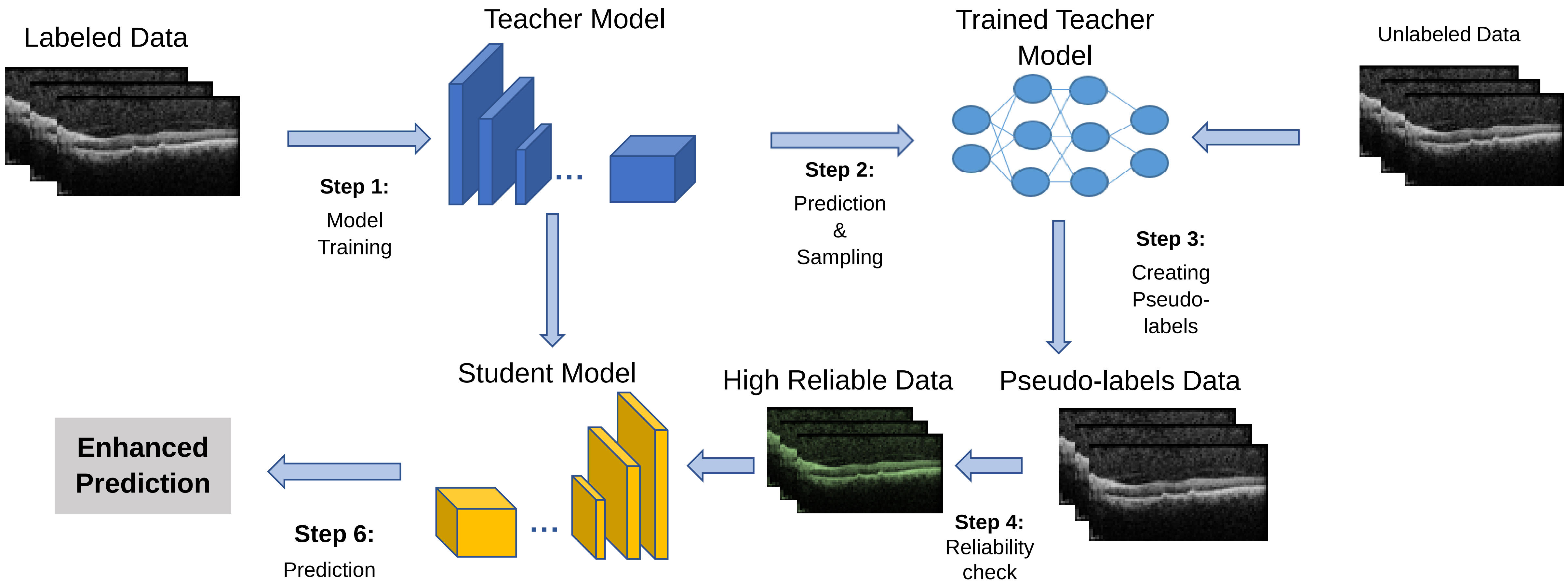
16822 OCT images of AMD cases

Proposed model: Optimized EfficienNet

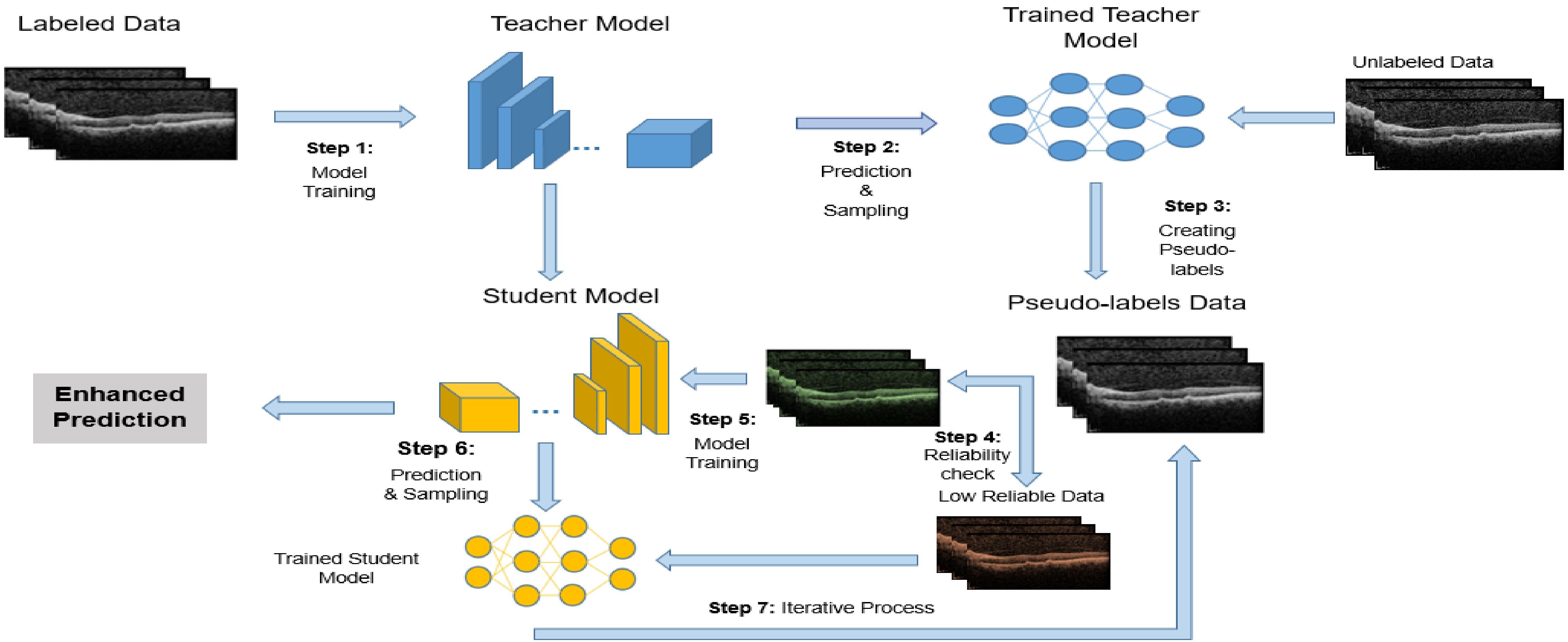
- Transfer Learning
- Data Augmentation
- Pre-trainable Layers



Proposed model: Teacher/Student



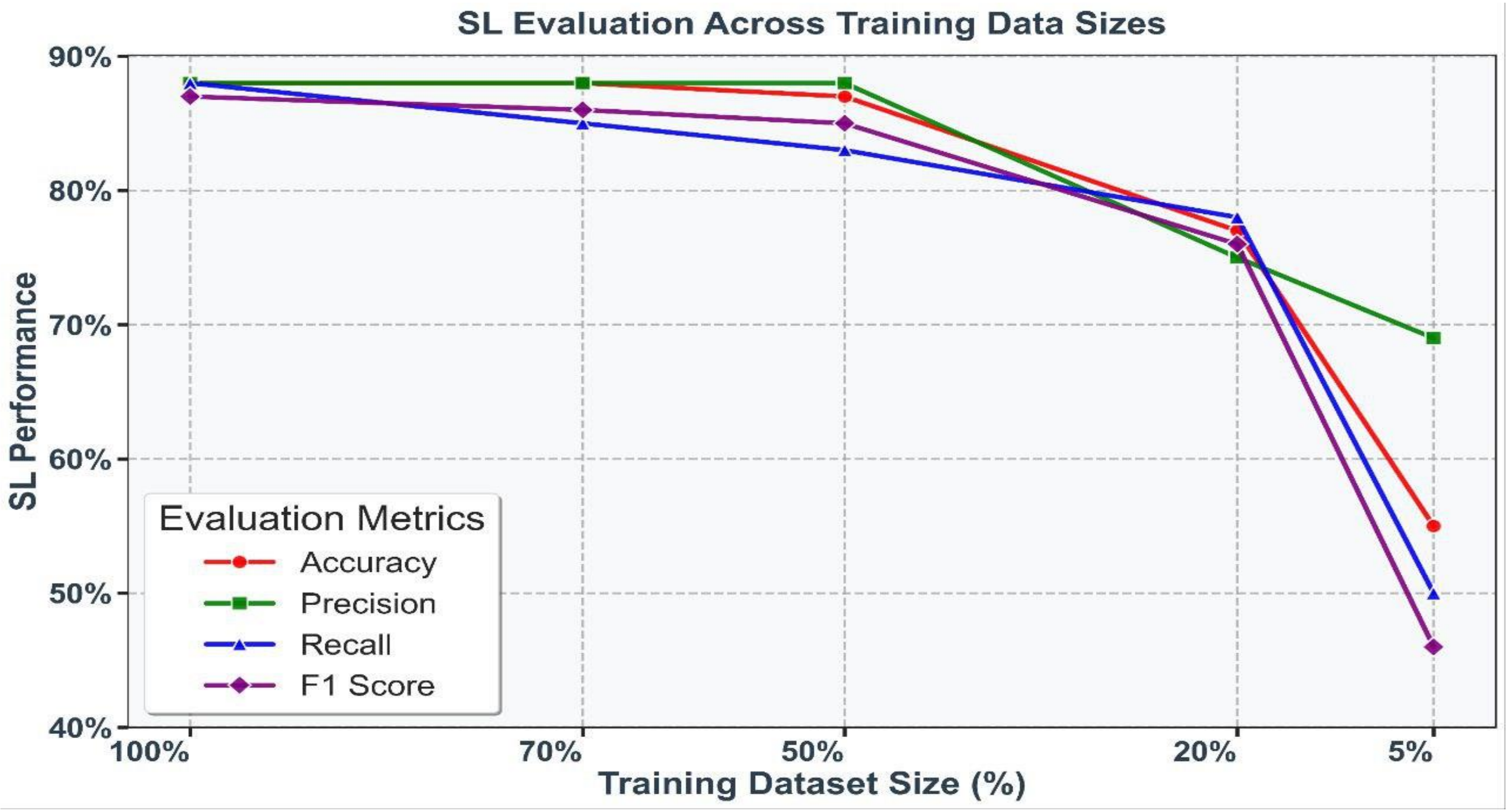
Proposed model: Iterative Teacher/Student



SL results:

F1 score (%)	Recall (%)	Precision (%)	Accuracy (%)	Augmentation	Pre-trainable layers	Weights	No.
87/33	87/66	87/33	87/14	✓	✓	ImageNet	1
86	84/66	88/33	86/27	×	✓	ImageNet	2
76/66	75	76/66	77/31	✓	×	ImageNet	3
78/33	77	80/33	79/19	×	×	ImageNet	4
85	85/33	86/66	86/27	✓	✓	×	5
69	66/33	73	69/39	×	✓	×	6

SL Limited Data Results:



TS Results: 70% of Train Data

F1 score (%)	Recall (%)	Precision (%)	Accuracy (%)	Total images	Pseudo labels	Pre-trainable layers	SL model
86	85/33	87/66	88/44	7535	×	✓	Teacher
87	85/66	88/33	88/71	9094	1559	✓	Student
71	69	76/23	75/70	7535	×	×	Teacher
72	69/66	76/66	76/55	7776	241	×	Student

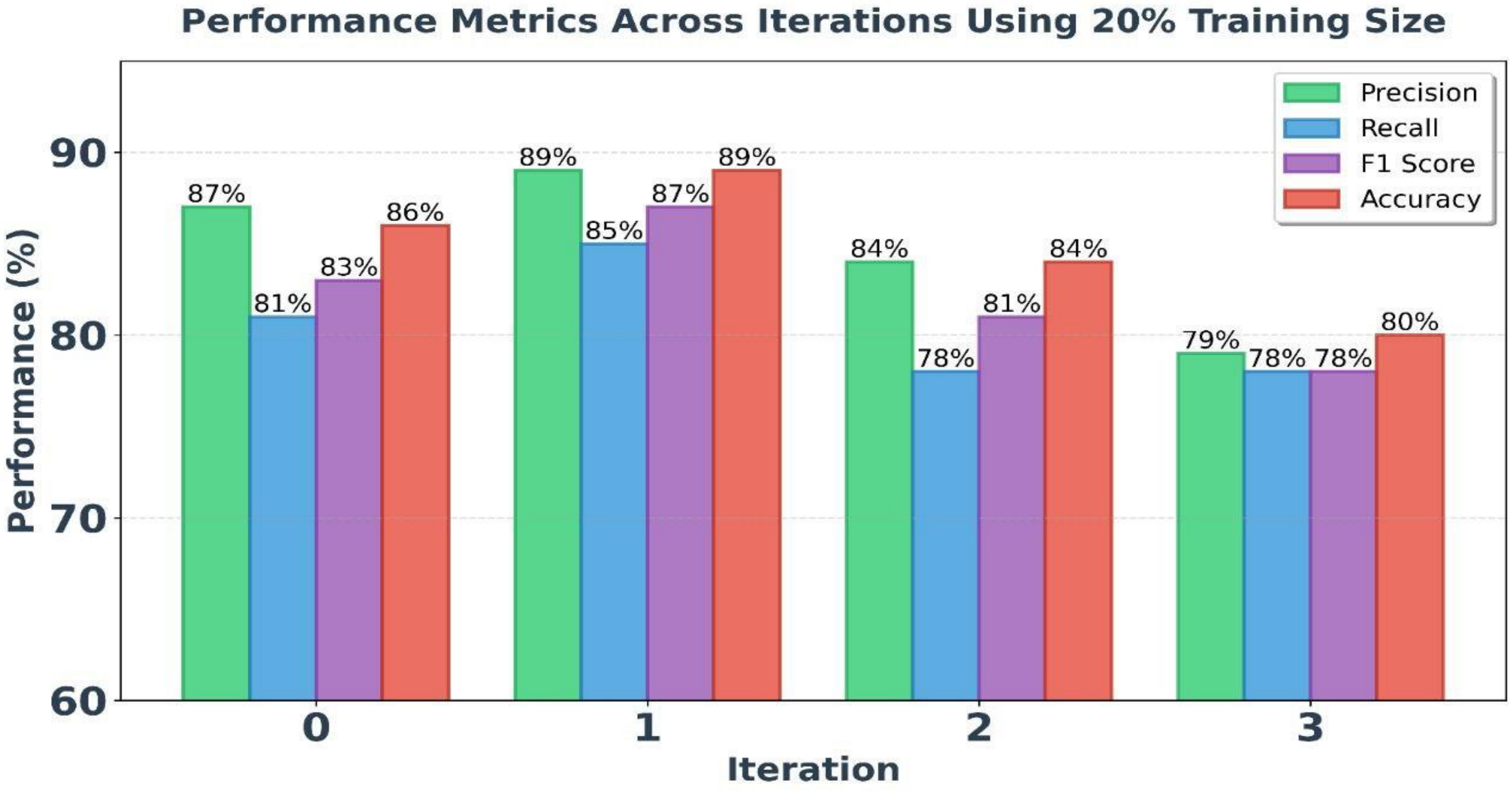
TS Results: 50% of Train Data

F1 score (%)	Recall (%)	Precision (%)	Accuracy (%)	Total images	Pseudo labels	Pre-trainable layers	SL model
85/33	83/33	88	87/10	5255	×	✓	Teacher
85	83	89	88/32	7160	1905	✓	Student
71	68/66	76	75/28	5255	×	×	Teacher
70/33	68	75/33	75/03	5666	411	×	Student

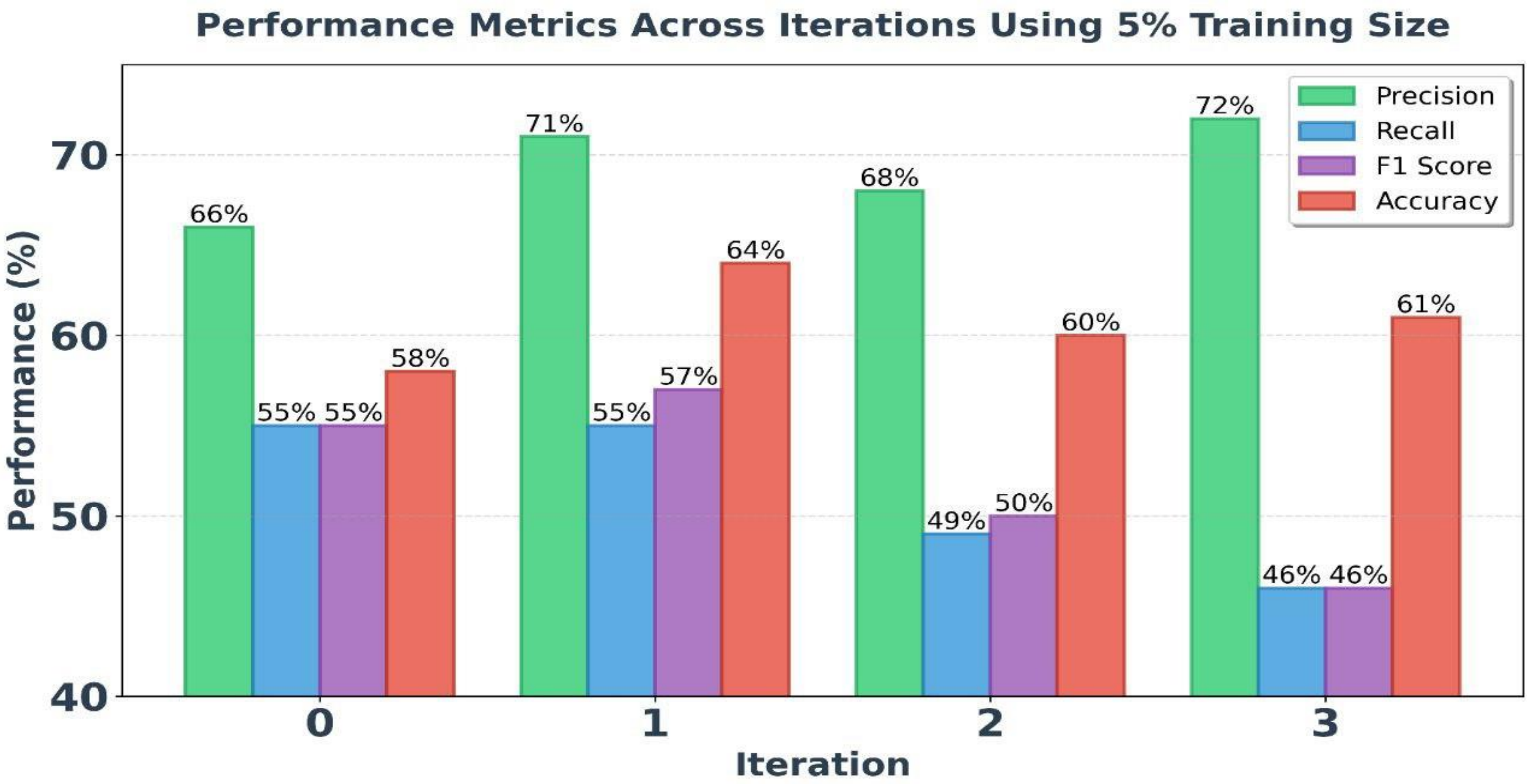
TS Results: 20% of Train Data

F1 Score (%)	Recall (%)	Precision (%)	Accuracy (%)	Pseudo labels	Confidence level
79	77	84	82/15	8648	0%
82	81	85	84/25	5440	90%
83	81/66	87/33	86/01	4529	95%
82	80	89	85/31	2816	99%

ITS Results: 20% of Train Data



ITS Results: 5% of Train Data



Discussion:

Discussion

Best parameters , Transfer Learning , Data Augmentation

Critical role of Pre-Trainable Layers in increasing the SL accuracy by 7% to 10%

selected parameter settings stability even in accuracy reduction in limited datasets

Discussion:

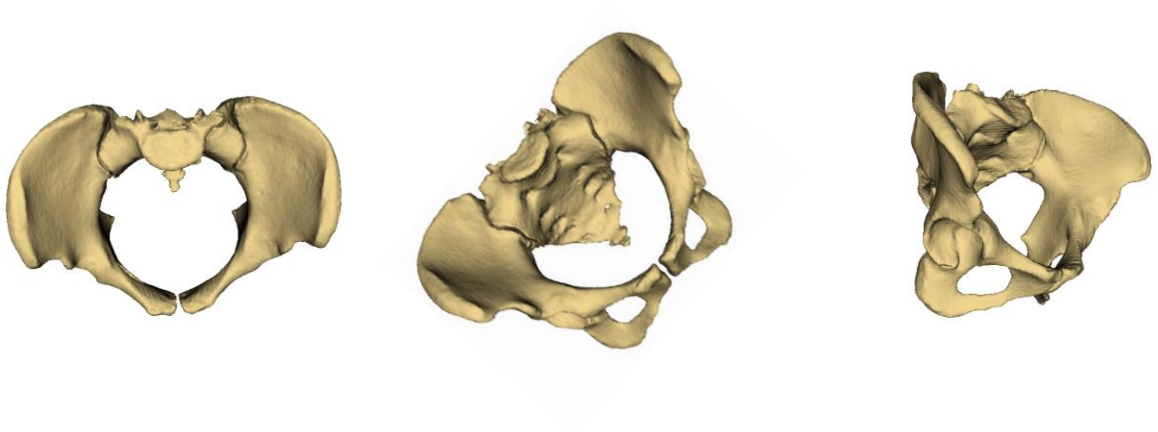
Discussion

- 14% Increasing of the SSL model accuracy with Pre-Trainable Layers
- effective pseudo-labeling strategy in SSL models
- strict confidence level in generating pseudo-labels
- necessity of balanced approach in adding high pseudo-labels to new train dataset

AI in Medical Imaging

Applications | Pelvic Tilt Estimation

- Total Hip Arthroplasty (THA) is one of the most prevalent orthopedic operations that is mostly used for the treatment of osteonecrosis, osteoarthritis, and developmental dysplasia of the hip
- In THA, pelvic tilt in standing position is an important factor in cup alignment planning.

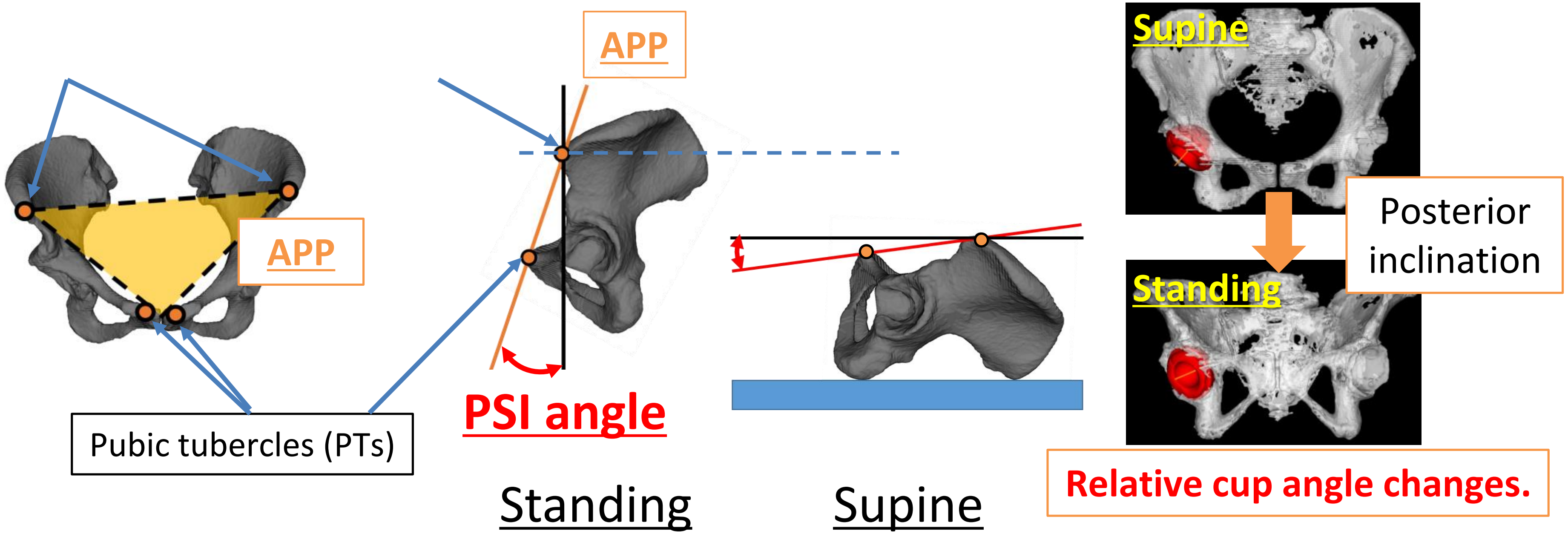


Ata Jodeiri, Yoshito Otake, et. al. - EPiC Series in Health Sciences (2018)
Estimation of Pelvic Sagittal Inclination From Anteroposterior Radiograph Using Convolutional Neural Networks

AI in Medical Imaging

Applications | Pelvic Tilt Estimation

- PSI angle is defined by angle between anterior pelvic plane (APP) and vertical direction.

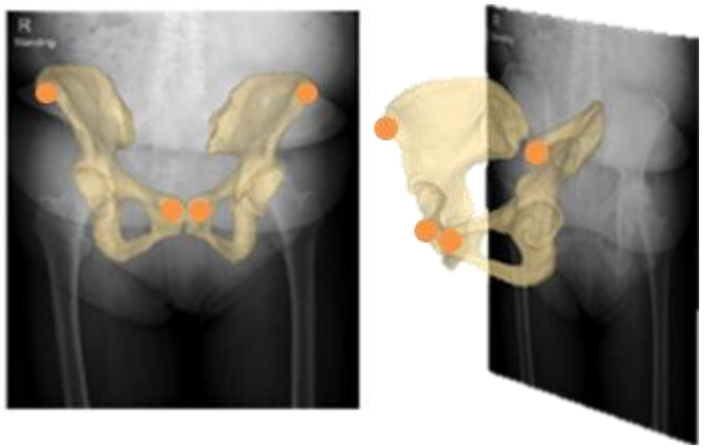


AI in Medical Imaging

Applications | Pelvic Tilt Estimation

Automated 2D-3D registration of radiograph (standing) and CT image (supine position).

Uemura, Keisuke et al., “Change in Pelvic Sagittal Inclination From Supine to Standing Position Before Hip Arthroplasty,” *J. Arthroplasty*, 32,8 (2017): 2568–2573.

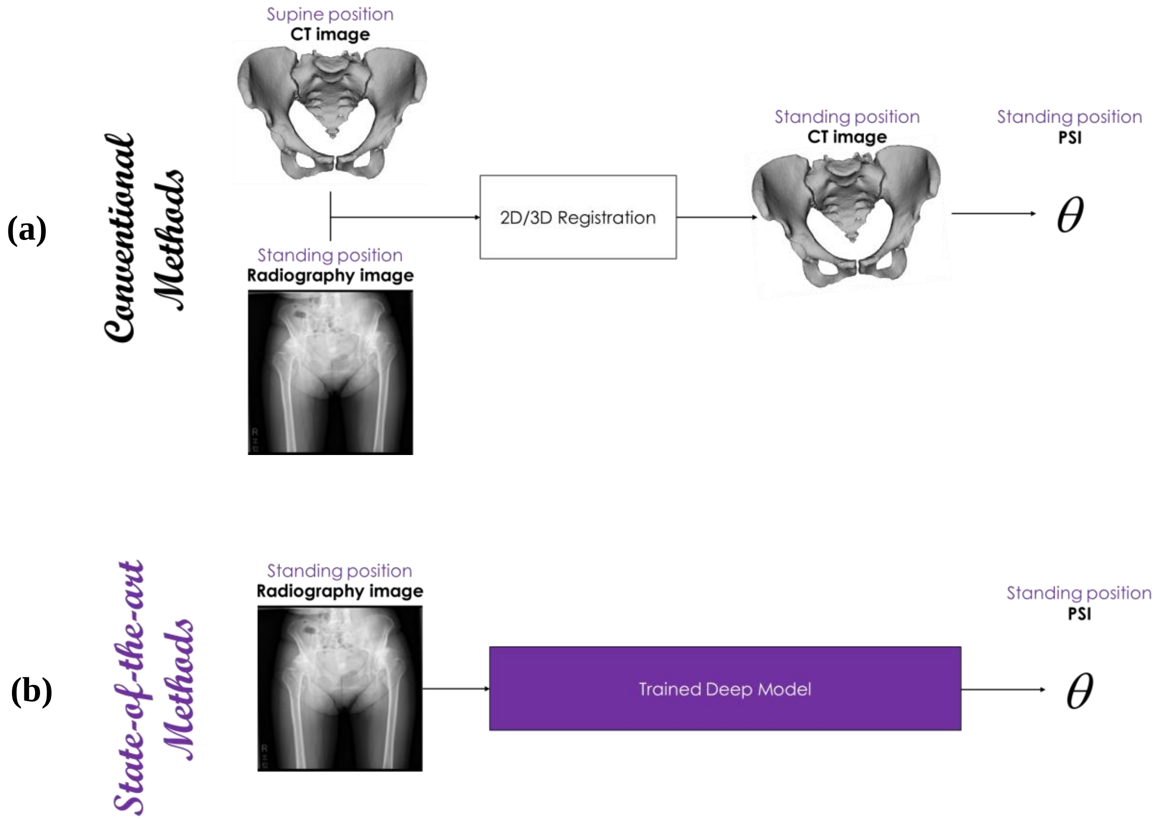


AI in Medical Imaging
Applications | Pelvic Tilt Estimation

Summary of CT dataset	
Hospital	Osaka Univ. Hospital, Dept. of Orthopaedic Surgery
Study population	Patients who are subjected to THA surgery
Number of cases	475 cases（Male:69, Female:406, Average age：59 y.o.）
Patient position	Supine
Matrix size and voxel dimension	Approx. 512 x 512 x 550 [voxels] (0.7 x 0.7 x 1 [mm^3])

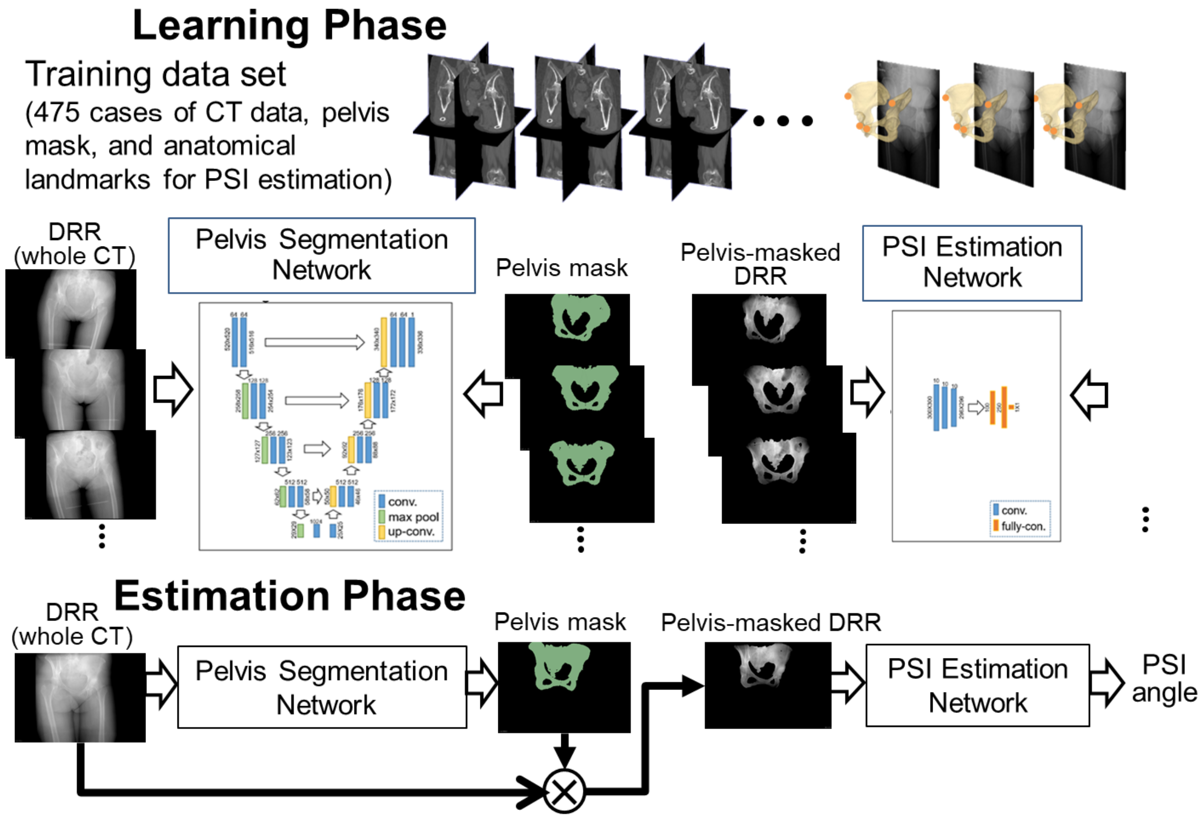
AI in Medical Imaging

Applications | Pelvic Tilt Estimation

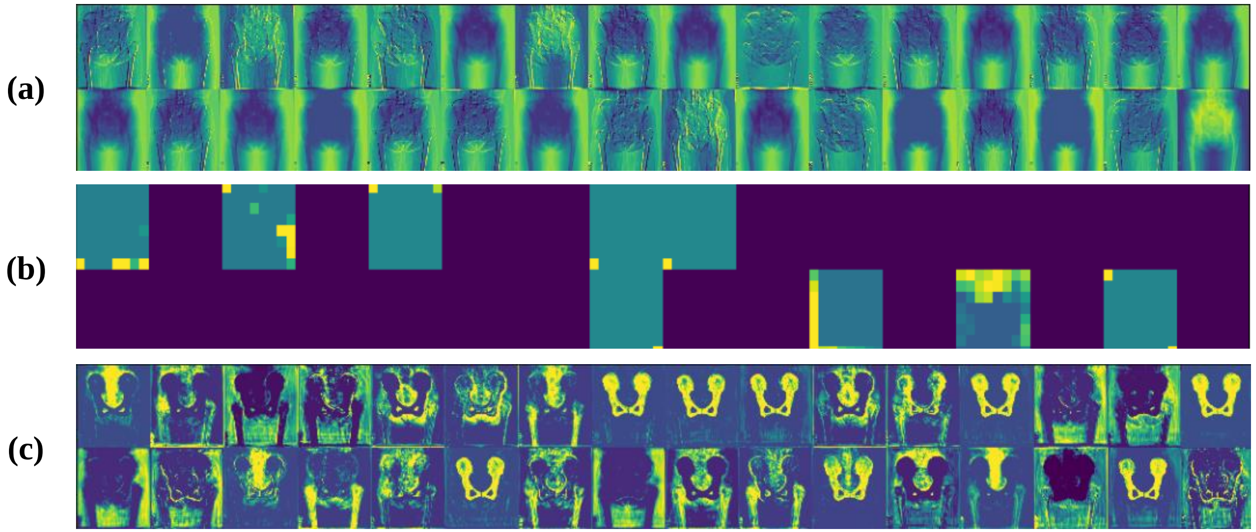


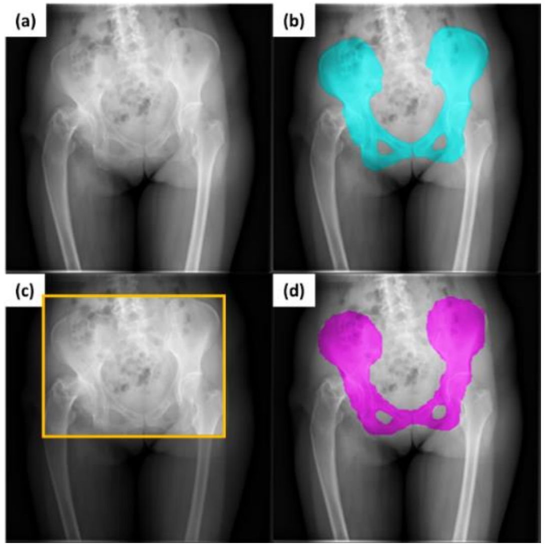
AI in Medical Imaging

Applications | Pelvic Tilt Estimation



AI in Medical Imaging
Applications | Pelvic Tilt Estimation



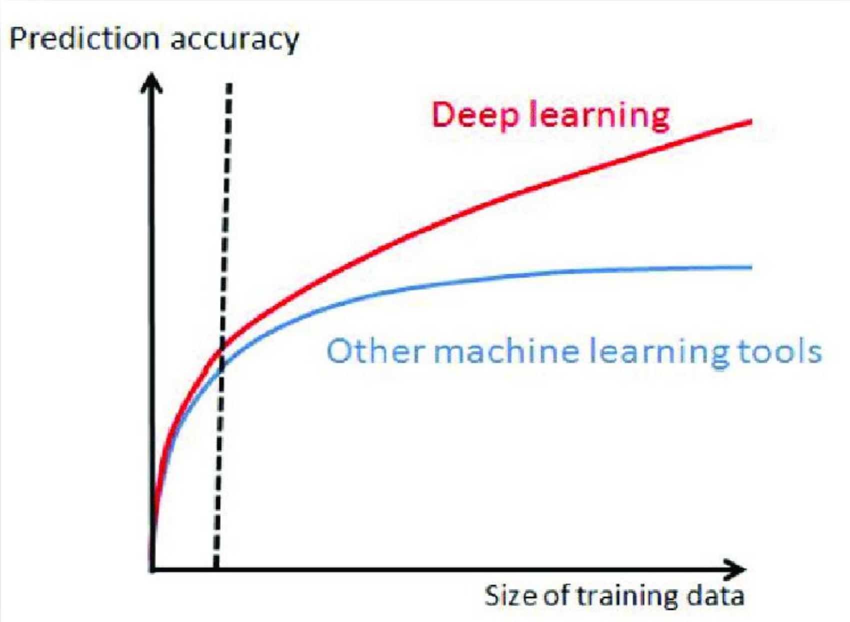


AI in Medical Imaging

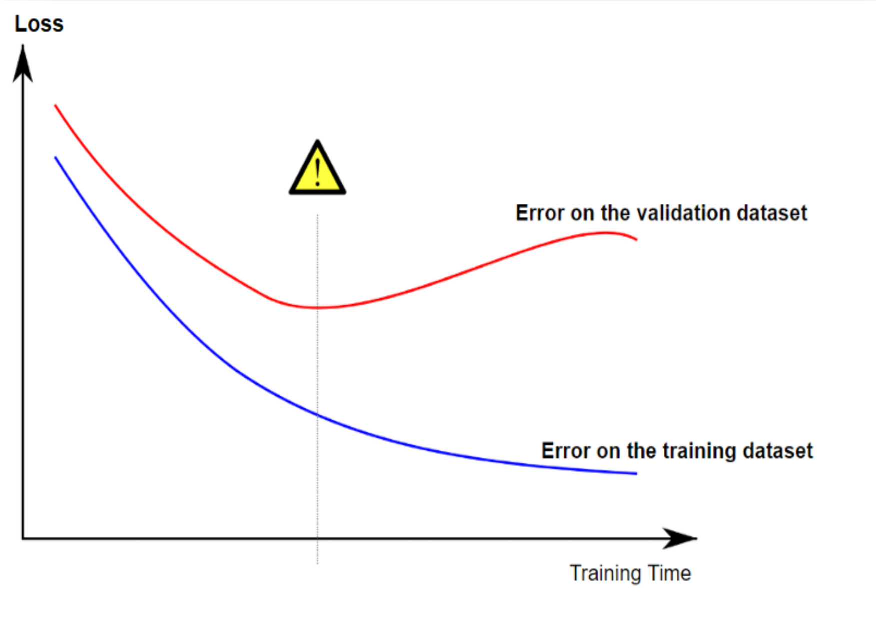
Applications | Pelvic Tilt Estimation

Image	Ground-truth	Network Prediction	Six Surgeon Prediction					
			#1	#2	#3	#4	#5	#6
1	**P** (-13.57)	**P** (-13.30)	**P**	**P**	**P**	**P**	**U**	**P**
2	**A** (+6.02)	**A** (+0.90)	**A**	**A**	**P**	**A**	**A**	**A**
3	**A** (+7.82)	**A** (+8.87)	**A**	**A**	**A**	**P**	**A**	**P**
4	**A** (+10.54)	**A** (+9.79)	**A**	**P**	**A**	**A**	**A**	**A**
5	**P** (-10.34)	**P** (-5.94)	**A**	**P**	**P**	**P**	**P**	**P**
6	**P** (-5.12)	**P** (-6.74)	**A**	**P**	**P**	**A**	**A**	**A**
7	**P** (-25.38)	**P** (-17.99)	**P**	**P**	**A**	**P**	**P**	**P**
8	**A** (+5.73)	**A** (+1.19)	**P**	**A**	**A**	**A**	**A**	**A**
9	**A** (+8.71)	**A** (7.80)	**A**	**A**	**A**	**A**	**P**	**A**
10	**P** (-6.69)	**P** (-8.41)	**P**	**A**	**P**	**P**	**P**	**P**
11	**A** (+5.62)	**A** (+4.09)	**A**	**A**	**P**	**A**	**P**	**A**
12	**P** (-22.46)	**P** (-18.06)	**P**	**P**	**P**	**P**	**P**	**P**
13	**P** (-5.83)	**P** (-2.09)	**U**	**P**	**A**	**A**	**U**	**A**
14	**P** (-8.44)	**P** (-3.57)	**P**	**P**	**A**	**P**	**P**	**P**
15	**A** (+6.74)	**A** (+7.40)	**A**	**A**	**P**	**U**	**A**	**A**
16	**A** (+6.74)	**A** (+7.40)	**U**	**P**	**P**	**P**	**A**	**P**
17	**P** (-2.17)	**P** (-1.53)	**A**	**A**	**P**	**A**	**U**	**A**
18	**A** (+3.33)	**A** (+3.23)	**A**	**A**	**P**	**A**	**P**	**P**
19	**P** (-17.39)	**P** (-18.21)	**P**	**P**	**P**	**P**	**P**	**P**
20	**P** (-0.82)	**P** (-1.91)	**U**	**P**	**A**	**P**	**P**	**P**

Data Scarcity

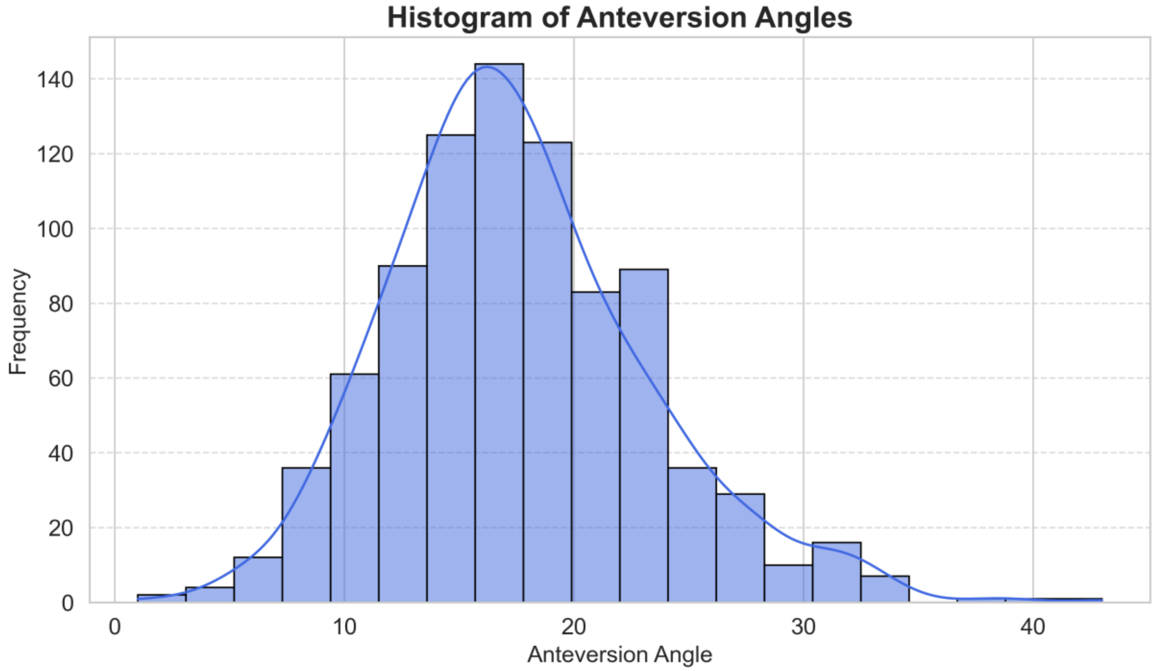


Noisy Labels

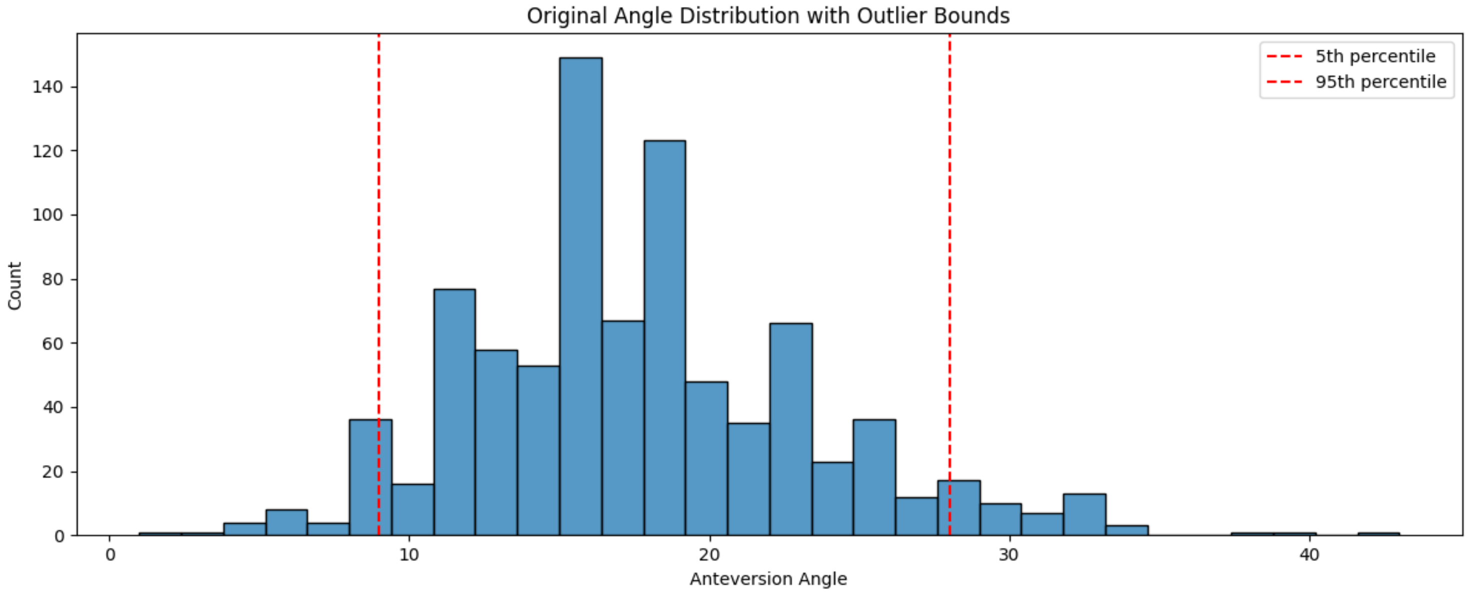


Data

Age Group	N	Age	Anteversion Angle	Side Distribution
0-24	118	19.84 ± 3.42	15.42 ± 4.80	R: 66.95%, L: 33.05%
25-49	291	38.51 ± 7.19	16.87 ± 5.02	R: 60.48%, L: 39.52%
50-99	461	72.15 ± 11.75	18.52 ± 6.11	R: 53.8%, L: 46.2%
All	870	53.80 ± 22.48	17.55 ± 5.70	R: 57.82%, L: 42.18%



Outlier Detection (percentile-based approach)



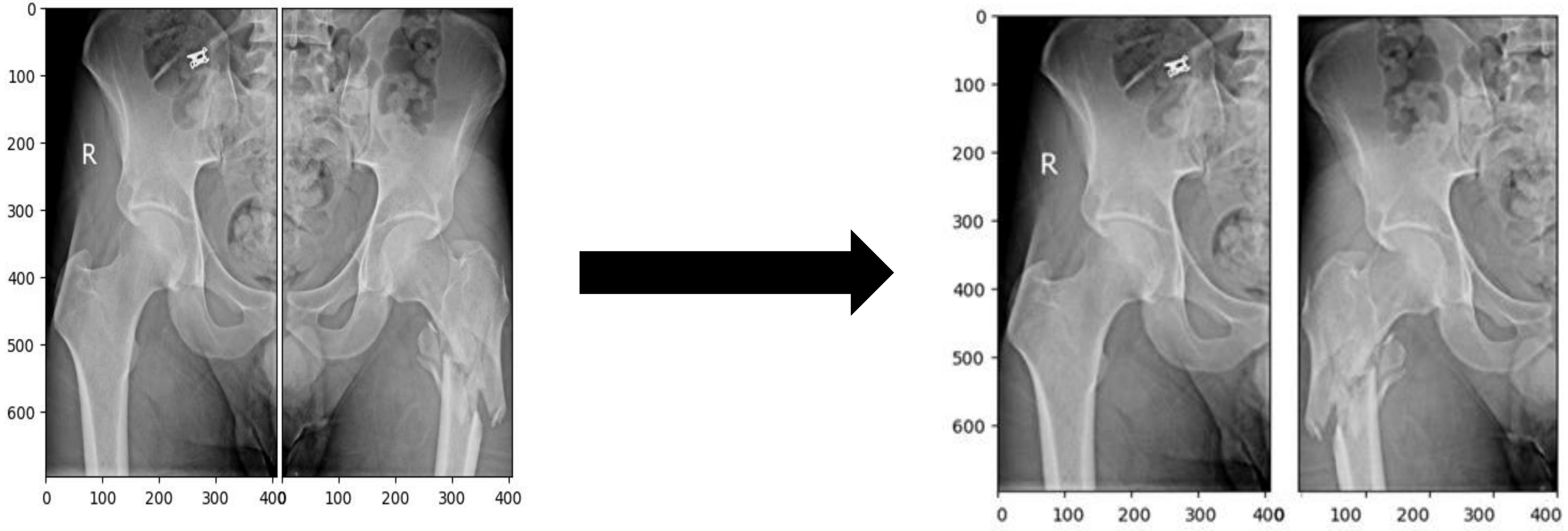
Summary Statistics	
Original samples	870
After outlier removal	806

Gaussian Mixture Model (GMM) Clustering

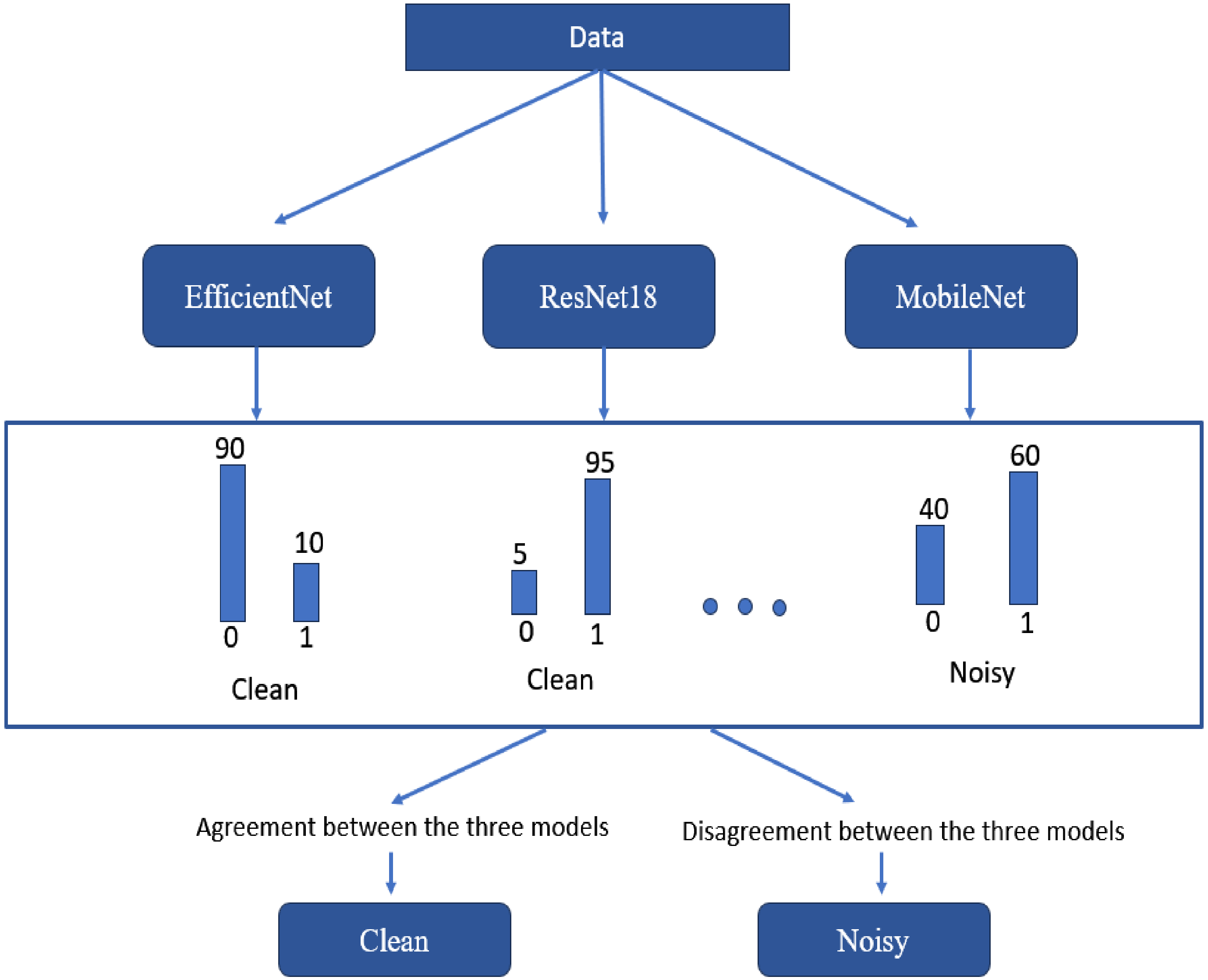
Angle Category	Distribution	Angle Range
High	237	14.0° - 19.0°
middle	392	9.0° - 13.0°
low	177	20.0° - 28.0°

414

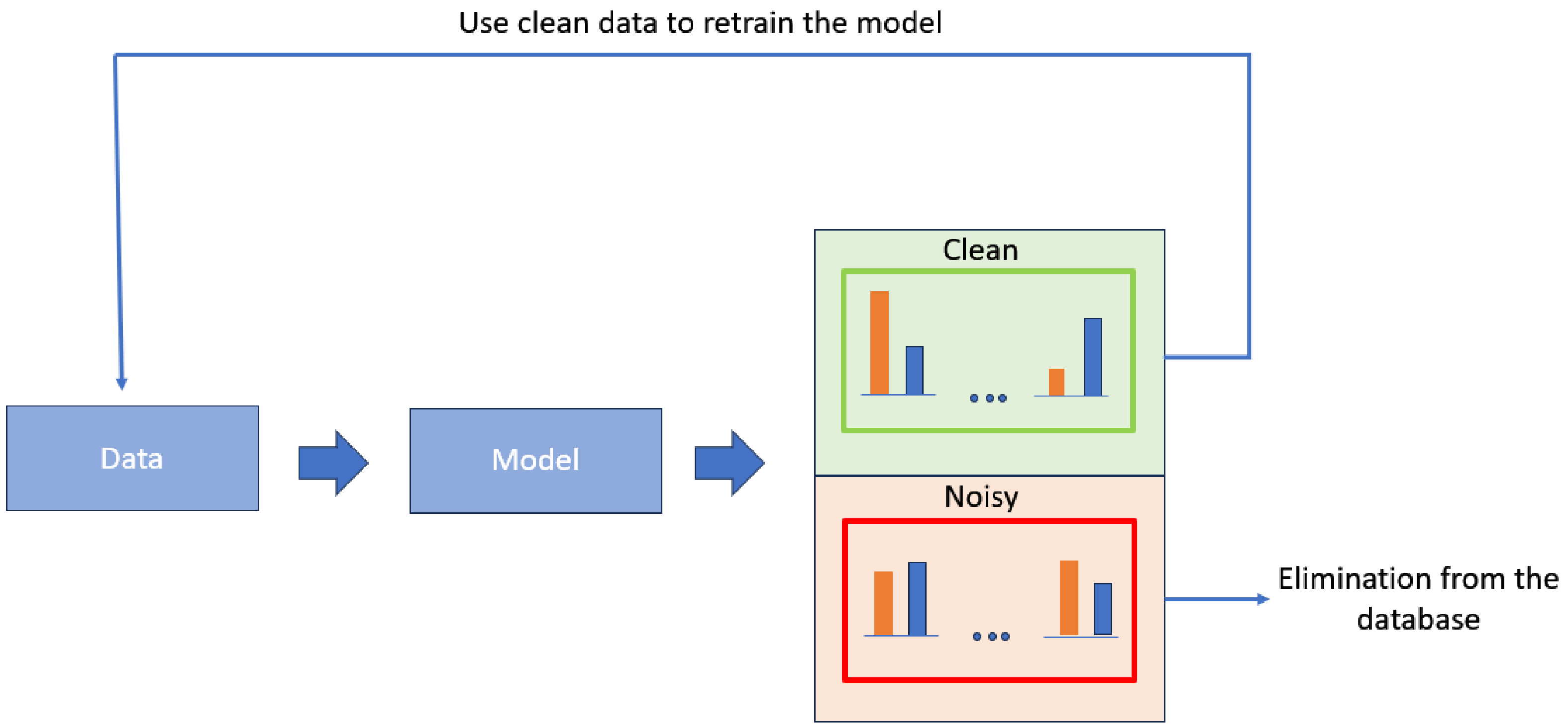
Data Preprocessing



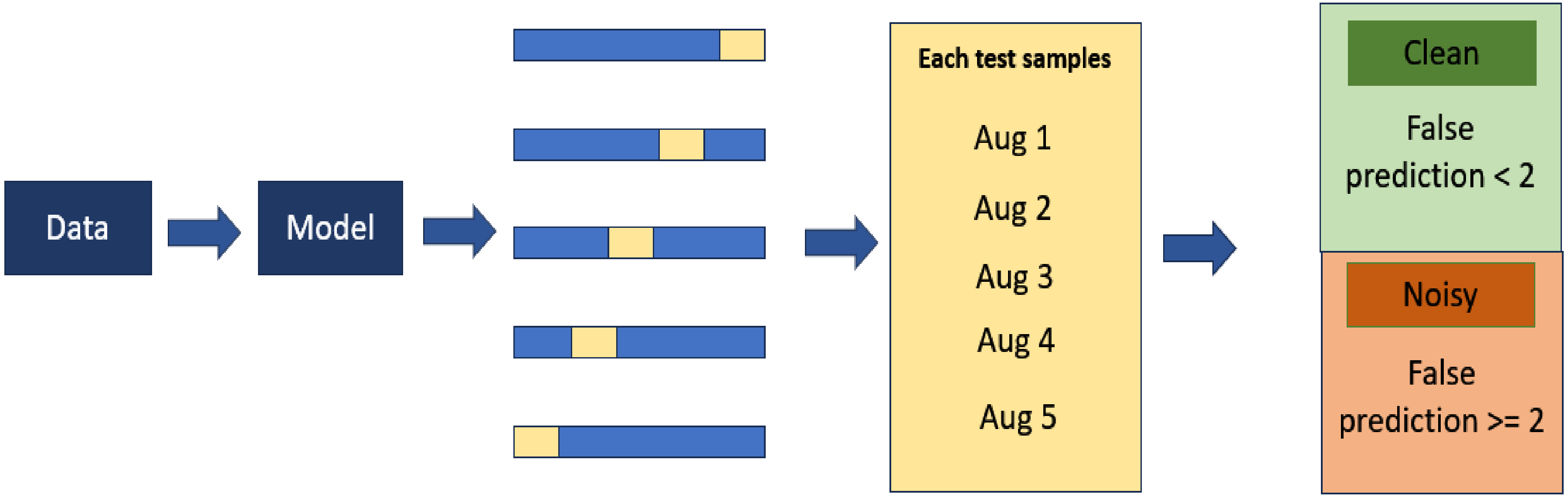
1. Cross-Model Comparison for Noise Detection)



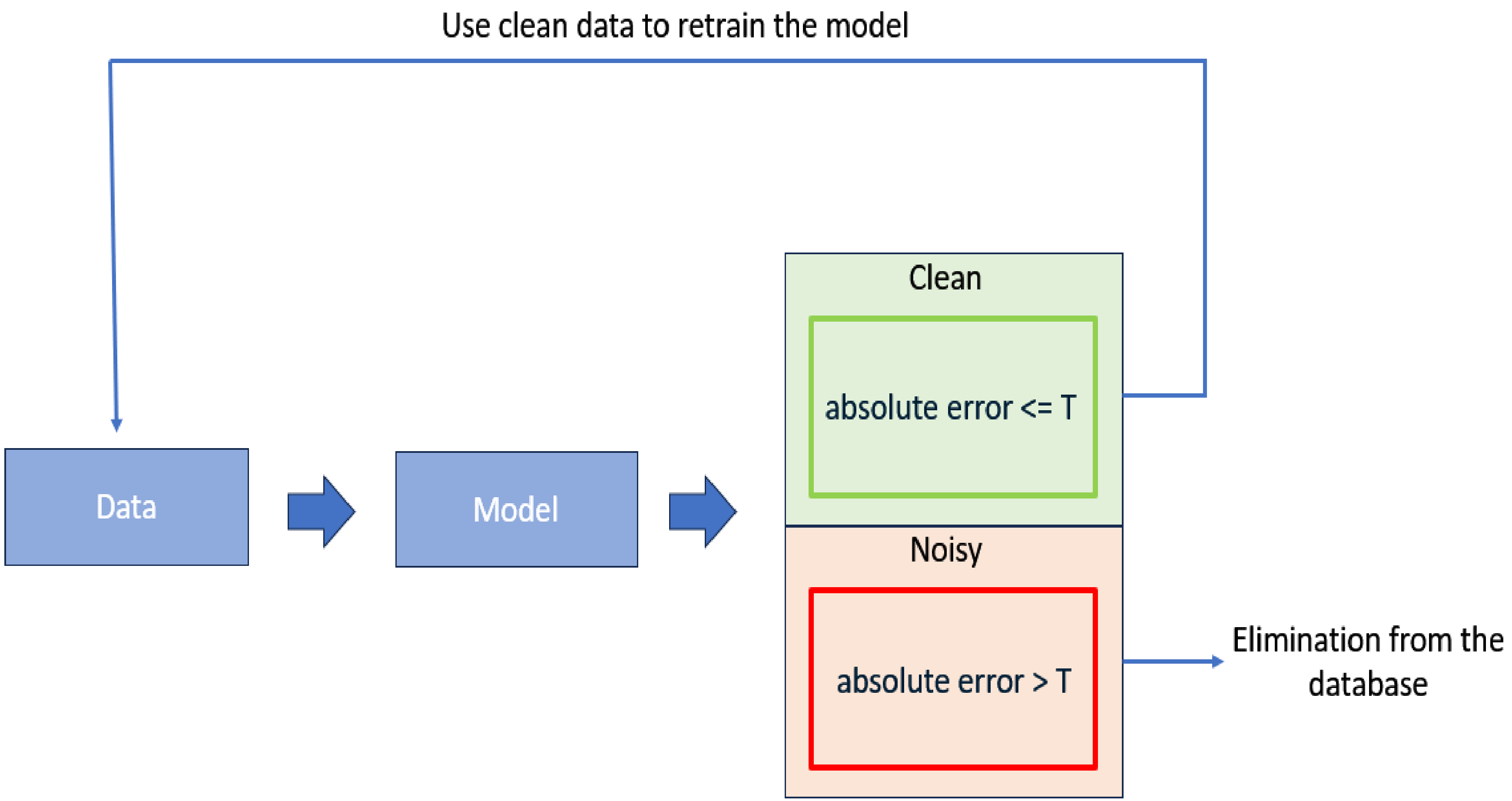
2. Iterative Noise Detection and Refinement

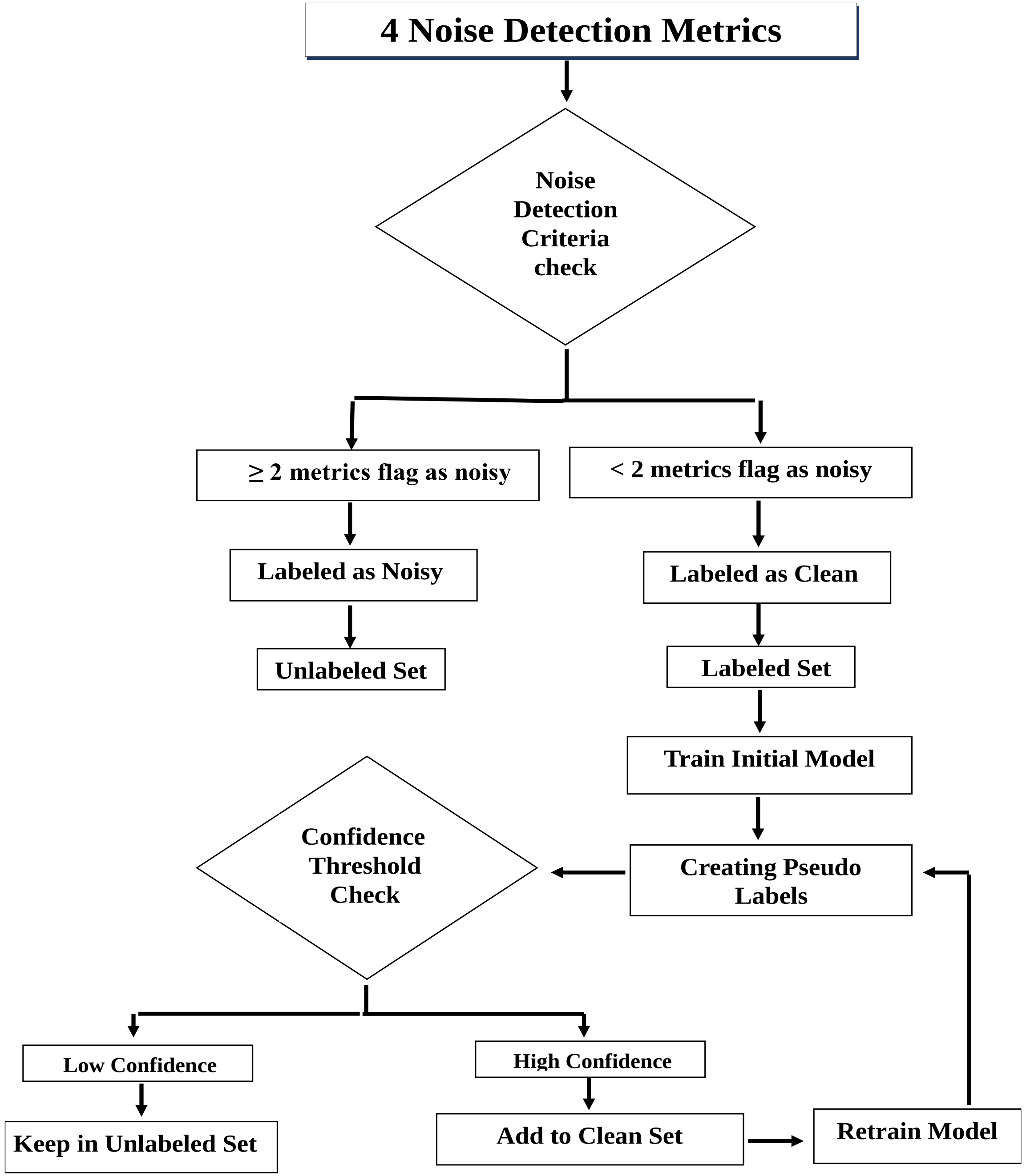


3. Augmentation Consistency Metric

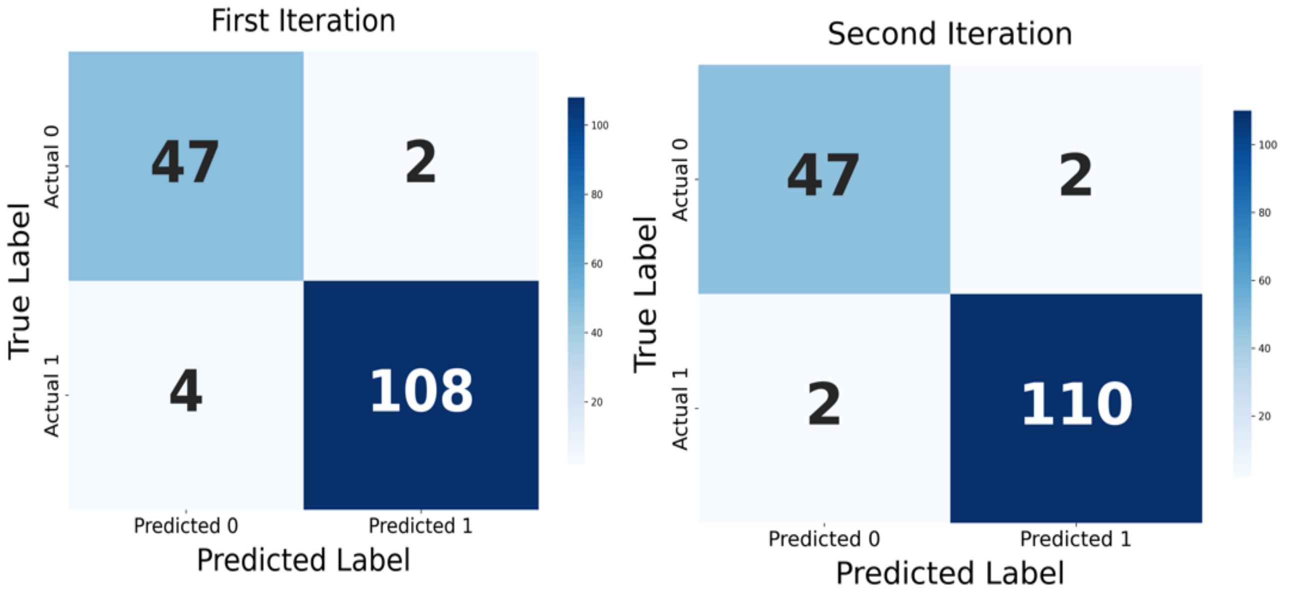


4. Using Prediction Errors to Identify Noise



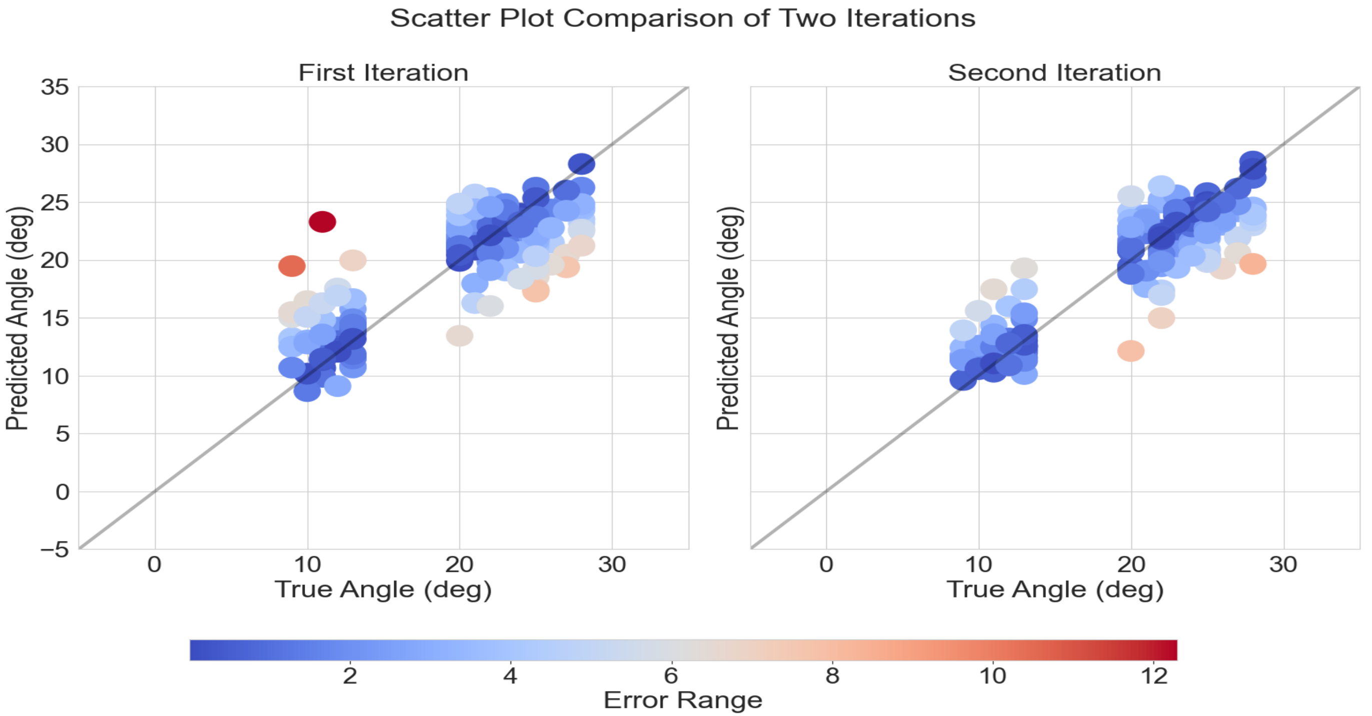

4 Noise Detection Metrics
Noise Detection Criteria check
≥ 2 metrics flag as noisy
< 2 metrics flag as noisy
Labeled as Noisy
Labeled as Clean
Unlabeled Set
Labeled Set
Train Initial Model
Confidence Threshold Check
Creating Pseudo Labels
Low Confidence
High Confidence
Keep in Unlabeled Set
Add to Clean Set
Retrain Model

Classification Results



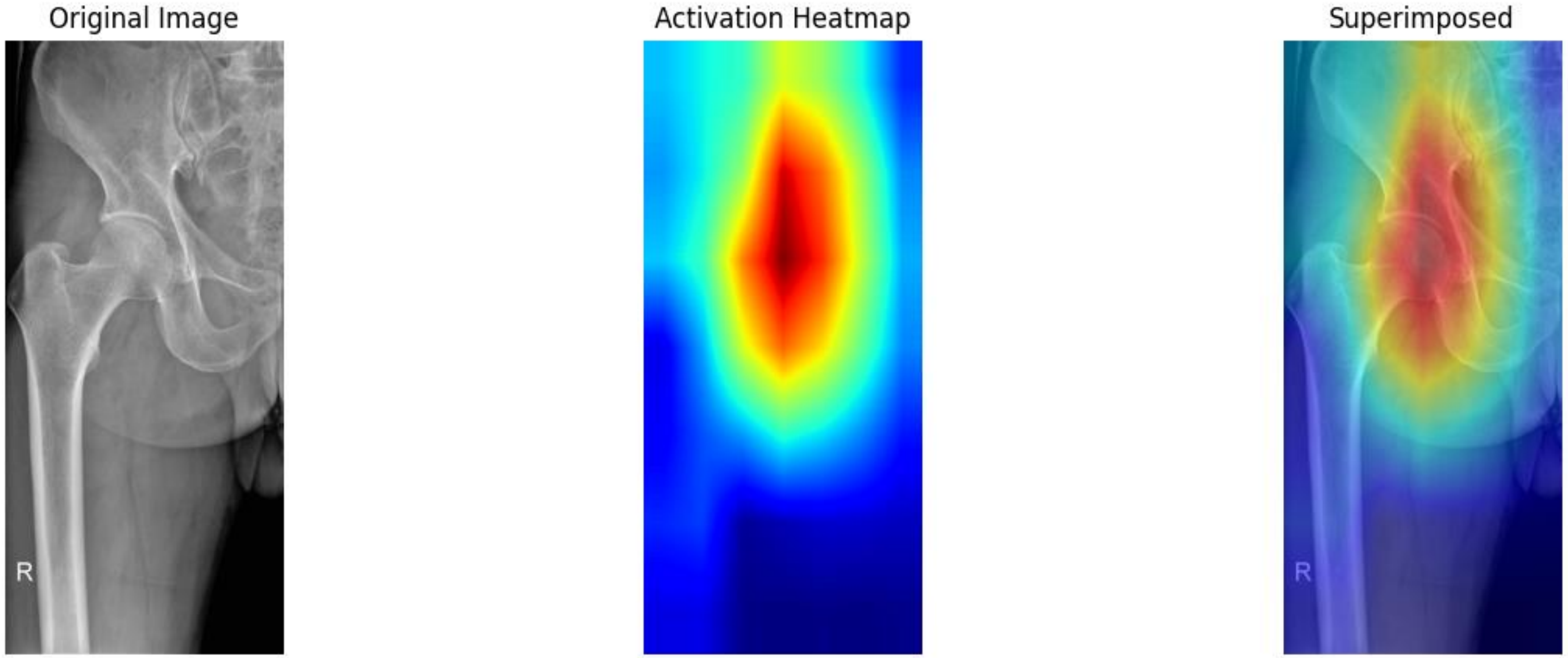
Iteration	Precision	Recall	F1 Score	Accuracy	support
First Iteration	0.95	0.96	0.96	0.96	161
Second Iteration	**0.97**	**0.97**	**0.97**	**0.98**	161

Regression Results



Iteration	MAE	MSE	RMSE	MAPE	R^2
First Iteration	2.64	11.73	3.42	15.69	0.66
Second	**2.13**	**7.62**	**2.76**	**11.74**	**0.77**

Model Activation and Localization Results



Method	Classification					Regression			
	Precision	Recall	F1_score	Accurac y	support	MAE	MSE	MAPE	R^2
Cross-Model Agreement	**0.95**	**0.91**	**0.93**	**0.95**	120	2.61	12	15	0.63
Confidence-Based Filtering	0.72	0.71	0.71	0.72	294	4.12	27.10	27.36	0.27
Augmentation Consistency	0.87	0.84	0.86	0.87	249	3.34	18.17	21.98	0.52
Regression Error-Based Filtering	0.88	0.89	0.88	0.90	183	**2.4**	**9.94**	**14.9**	**0.65**

Thank you
for your attention

A.Jodeiri@tbzmed.ac.ir
Ata.Jodeiri@gmail.com