Journal of Theoretical and Applied Information Technology
30th April 2025. Vol.103. No.8

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

AN OPTIMIZED DEEP LEARNING MODEL FOR EARLY DETECTION OF RHEUMATOID ARTHRITIS USING KNEE X-RAYS

SALONI FATHIMA[1], G. SHANKAR LINGAM [2]

[1] Research Scholar, Chaitanya Deemed to be University, Department of Computer Science and Engineering, Warangal, India.

[2] Professor, Chaitanya Deemed to be University, Department of Computer Science and Engineering, Warangal, India.

E-mail: [1] salonifathimamd@gmail.com, [2] shankar@chaitanya.edu.in.

ABSTRACT

Rheumatoid arthritis (RA) is a chronic autoimmune condition that requires early diagnosis to prevent irreversible joint damage. While deep learning models have shown promise in automated RA detection, existing approaches suffer from data imbalance, suboptimal feature extraction, and poor generalizability, limiting their clinical applicability. This study proposes an optimized Convolutional Neural Network (CNN) model, incorporating image augmentation and class balancing techniques to improve RA detection accuracy from knee X-ray images. Unlike previous studies, which often overlook the impact of data augmentation on model performance, our work demonstrates its effectiveness in addressing data imbalance and enhancing model robustness. We trained and validated our model using the Kaggle knee X-ray dataset, applying image augmentation to expand training samples and oversampling to balance class distributions. The CNN was optimized through rigorous hyperparameter tuning. Our optimized CNN model achieved a high accuracy of 94%, outperforming baseline deep learning models. Data augmentation and oversampling significantly improved model performance, proving their effectiveness in medical imaging tasks. Our findings establish a novel deep learning framework for RA detection, demonstrating the importance of data augmentation and optimization in improving diagnostic accuracy. This work contributes to the growing field of AI in medical imaging by offering a scalable and interpretable solution for automated RA detection

Keywords: *Rheumatoid Arthritis, knee X-ray images, statistical augmentation, deep learning, CNN.*

1. INTRODUCTION

Rheumatoid Arthritis (RA) is a chronic disease that primarily affects the joints, but can also have widespread effects on other systems of the body. It is characterized by persistent inflammation, which leads to progressive joint destruction, pain, and stiffness, ultimately resulting in significant functional disability and a marked reduction in quality of life. This disease can affect all age peoples, although it is most commonly diagnosed in middle-aged persons, and affected women more than men. The main cause of RA is still not fully identified, but it is believed to involve a combination of genetic problems and environmental changes, such as infections or smoking, which contribute to the abnormal in the immune system.

Early diagnosis and curing of this disease is very important. The disease has many verities of symptoms, like for some patients experiencing mild symptoms while others severe joint damage. The "window of opportunity" concept in RA management suggests that initiating treatment early in the disease can significantly alter its damage to human body, reducing the likelihood of joint damage and improving long-term outcomes. This underscores the importance of accurate and timely diagnosis, particularly during the early stages of the disease when symptoms might be nonspecific.

However, the clinical diagnosis of RA is challenging task. RA's disease, with symptoms that can vary widely among patients, coupled with its similarity to other inflammatory joint diseases, makes early diagnosis difficult. The typical clinical assessment for RA includes evaluating symptoms such as joint pain, stiffness (especially in the

morning), and swelling. Serological tests, including the measurement of rheumatoid factor (RF) and anti-cyclic citrullinated peptide (anti-CCP) antibodies are often used to support the diagnosis. Elevated levels of acute-phase reactants like C-reactive protein (CRP) and erythrocyte sedimentation rate (ESR) can also indicate RA. Additionally, imaging techniques such as X-rays, ultrasound, and magnetic resonance imaging (MRI) are used to assess joint damage and inflammation.

The traditional approaches tried to increase the accuracy of RA diagnosis, but remains challenging, particularly in the early stages. Some symptoms like RF and anti-CCP are not always present in early stages of RA, and while imaging can reveal joint damage, these changes often occur later in the disease process. Moreover, RA symptoms can overlap with those of other conditions, such as osteoarthritis, lupus, and psoriatic arthritis, leading to potential misdiagnosis or delayed diagnosis. These challenges explore the need of more and different diagnostic approaches that can enhance early detection and provide a more accurate assessment of disease activity.

In recent years, the advancements in machine learning (ML) and deep learning (DL) methods have shown improving medical diagnostics. These technologies provide analyzing large volumes of data, which is increasingly available in healthcare through electronic health records (EHRs), medical imaging, and genomic data. ML algorithms can identify patterns and relationships within this data that may not work with traditional statistical methods. For example, in [1,2,3] ML models can analyze a patient's medical history, laboratory results, and imaging studies to predict the likelihood of RA or assess the risk of disease progression. The DL models [4] that uses neural networks with multiple layers, has shown particular promise in image analysis, enabling the automatic detection of abnormalities in radiographs or MRI scans that may indicate early joint damage.

The application of ML and DL in RA diagnostics has its own challenges. One of the primary issues is the need for large, high-quality datasets that include diverse patient [5,6] and comprehensive clinical information. The variability in data quality and the presence of missing data can also affect the performance of these models. Moreover, while ML models can achieve high accuracy in controlled research settings, their integration into routine clinical practice requires careful consideration of issues such as interpretability, reliability, and the potential for bias [7,8,9,10].

Despite these challenges, the merits of ML and DL in RA diagnostics are substantial. These technologies can assist clinicians in making more accurate and timely diagnoses, leading to earlier and more effective treatment interventions. For instance, ML models can help identify patients who are at high risk of rapid disease progression and may benefit from aggressive treatment strategies. Additionally, DL techniques can be used to analyze longitudinal data, tracking changes in disease activity over time and providing insights into the effectiveness of different treatment regimens.

Many of the researchers have integrated ML and DL methods [11,12,13] with other emerging technologies, such as wearable devices and remote monitoring tools, offers exciting possibilities for personalized medicine in RA. Patients with RA could benefit from continuous monitoring of their disease activity, with real-time data being analyzed by ML models to provide personalized treatment recommendations. This approach has the potential to revolutionize RA management, moving from a reactive to a proactive model of care, where treatment is tailored to the individual patient's needs and adjusted dynamically based on real-time data.

Despite advancements in deep learning, existing AI-based RA detection models face key limitations:

1. Data Imbalance: Many RA datasets suffer from class imbalances, leading to biased model predictions.
2. Limited Generalizability: Most deep learning models lack robustness when tested on diverse datasets.

To address these issues, this study proposes a customized Convolutional Neural Network (CNN) model optimized with data augmentation and class balancing techniques. The model is trained on knee X-ray images from the Kaggle dataset, ensuring better generalizability and improved accuracy. Our approach enhances RA detection performance by reducing bias, increasing robustness, and improving classification reliability.

Contribution:

- Implemented image augmentation for Knee sample expansion. And to balance all classes with over sampling method.
- Trained a well-optimized CNN with augmented data that increase performance up to 0.94.

- Proposed model consistently outperformed other models, showcasing the augmentation strategy's effectiveness.

2. RELATED WORK

In recent years, the use of ML and DL approaches for detecting and predicting RA has shown significant progress. Üreten Ket al. [14] implemented a customized Convolutional Neural Network (CNN) model trained on 135 samples, for automated RA detection using X-ray data and got an accuracy of 73.3%. Similarly, Li, Y. [15] implemented CNNs model on radiographs, first they did image pre-processing to enhance accuracy, so the accuracy increased to 67.80%.

Moreover, the integration of EH into automated model has proven better performance. Like Scientific, L. L. [16] implemented a DL like RCNN model to predict RA using features extracted from EHR data, and got accuracy of 92.81%. This approach highlights the potential of utilizing large-scale clinical data for disease prediction. Same way Peng, Y. et al. [17] and Wu, Q., and Dai, J. [18] implemented ML and advanced deep learning models like GoogLeNet and VGG16 and trained clinical data, achieving accuracies of 96.15%, 97.12 and 91.20%, respectively.

Rajesh, G., et al. [19] and Gandrup, J., et al [20] implemented an automated evaluation and early detection of RA using radiographs, achieving accuracies of 80.56% and AUC of 0.82 respectively. In addition to CNNs and DL models Chen, Y. J et al. [21] implemented ML algorithms to predict RA disease activity using clinical data and biomarkers, and got an optimal accuracy of 86.5%. Moradmand, H., and Ren, L[22] and Venäläinen, M. Set al [23] implemented transformers model with hybrid model to predict RA activity and biologic-free remission, achieving accuracies of 0.99% with hybrid model. Furthermore, some the researchers have worked on MRI samples images for RA classification and disease prediction with RA2-DREAM algorithm and got RMSE as 0.35. Likhith, R., et al. [24] studied various model on RA and deep learning models for early detection. And Alam, A., et al. [25] implemented DL models like CNN to classify RA from knee and hand MRI images, achieved an optimal accuracy.

Tu, J. B., et al [26] proposed ML models for RA diagnosis using clinical and genetic data, achieving an AUC of 0.773. In this work they integrated genetic information with clinical data to enhance diagnostic accuracy. Similarly, Mao, Y., et al. [27] Implemented custom models combined with genomics data for predicting RA outcomes, achieving a high AUC ranged from 0.941%.

In terms of imaging, Stoel, B. C., et al. [28] studied an image processing-based DL model for RA detection using X-ray images, and worked on comparative study on all types of samples. Ma, Y., et al [29] also focused on early RA classification using multi-feature fusion from hand X-ray images like RA or not , achieving an accuracy of 0. 872%.

Ho, C. S., et al [30] extracted different features for RA identification in hand radiographs, achieving an accuracy of 80%. In this they explained critical role of feature engineering in improving model performance with DeepDXA a CNN-based model. Hassanzadeh, T., et al.[31] applied fully automated deep learning methods for predicting RA disease activity using hand MRI images, achieving true positive and false negative rate. Some works have also explored the prediction of treatment outcomes in RA. Duquesne, J., et al [32] applied ML models to predict treatment response, using clinical and treatment data, with an AUROC of 0.72. Li Het al.[33] and Okino T et al. [34] implemented ML methods to predict biologic-free remission and remission in RA patients treated with methotrexate, achieving Pearson's correlation 0.711 and GSS progression ($p = 0.004$) respectively. The use of radiographic images has also been increased by researcher like Kato K et al. [35] proposed DL model for RA detection using ultrasound images, with a notable results like R square as 0.986.

Despite significant advancements in AI-based RA detection, existing studies have notable limitations. Many CNN-based models, such as those by Üreten K et al. [14] and Li, Y. [15], suffer from low accuracy (73.3% and 67.8%, respectively) due to small, imbalanced datasets and lack of augmentation. While hybrid approaches integrating EHR and clinical data, like those by Scientific, L. L. [16] and Peng, Y. et al. [17], achieve high accuracies (92.81%–97.12%), they primarily rely on non-image data, limiting their applicability for radiographic RA diagnosis. Advanced models using transformers and hybrid techniques, such as those by Moradmand, H. and Ren, L. [22], report near-perfect accuracy (99%), but their reliance on complex architectures raises concerns about real-world deployment and generalization. Similarly, MRI- and ultrasound-based methods (e.g., Alam, A. et al. [25] and Kato K et al. [35]) offer high precision but are costly and less accessible than X-ray-based diagnosis.

Additionally, many prior works fail to adequately address class imbalance, proper feature extraction, and overfitting risks, limiting their robustness.

3. METHODOLOGY

The application involved deploying an optimized CNN model for identifying RA in knee images. The CNN model was trained using a refined set of parameters, as illustrated in Figure 1. The model can classify diseases into five distinct classes: normal, doubtful, mild, severe, and moderate.

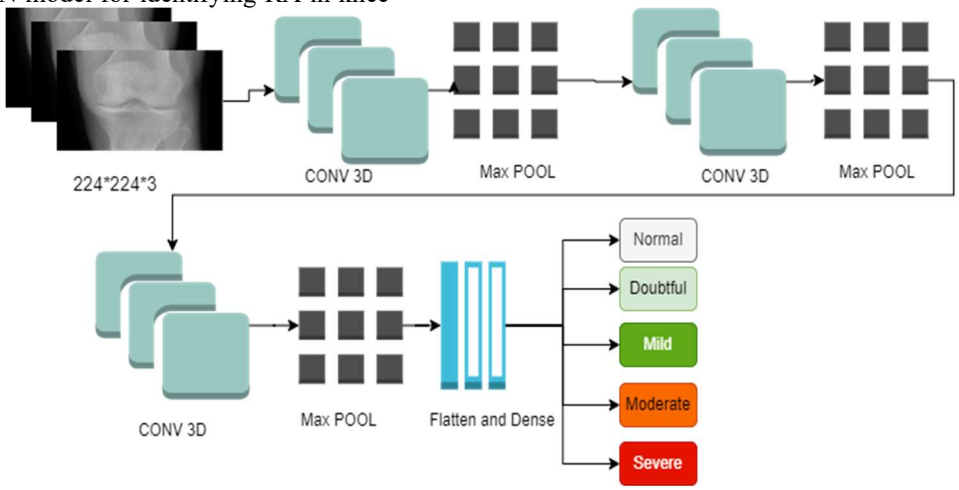

Figure 1: Implemented CNN model

3.1 Data Augmentation

For our approach, we utilized knee medical X-ray images sourced from Kaggle [31], comprising a dataset of 1650 samples categorized into five classes: normal, doubtful, mild, moderate, and severe, as detailed in Table 1. Notably, the original dataset exhibited uneven distribution among the classes. To mitigate the risk of overfitting, we implemented data augmentation by employing the flipping method, which involved flipping the samples at various angles, as depicted in Figure 2. Subsequently, we balanced the dataset, resulting in an equal distribution of 514 images across all classes, as illustrated in Figure 3. This balancing step helps enhance the model's generalization performance.

Table 1: Actual Knee Samples And Augmented Samples.

	Normal	Doubtful	Mild	Moderate	Severe	Total
Actual number of samples	514	477	232	221	206	1650
Augmented samples	514	514	514	514	514	2570

Journal of Theoretical and Applied Information Technology
30th April 2025. Vol.103. No.8

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

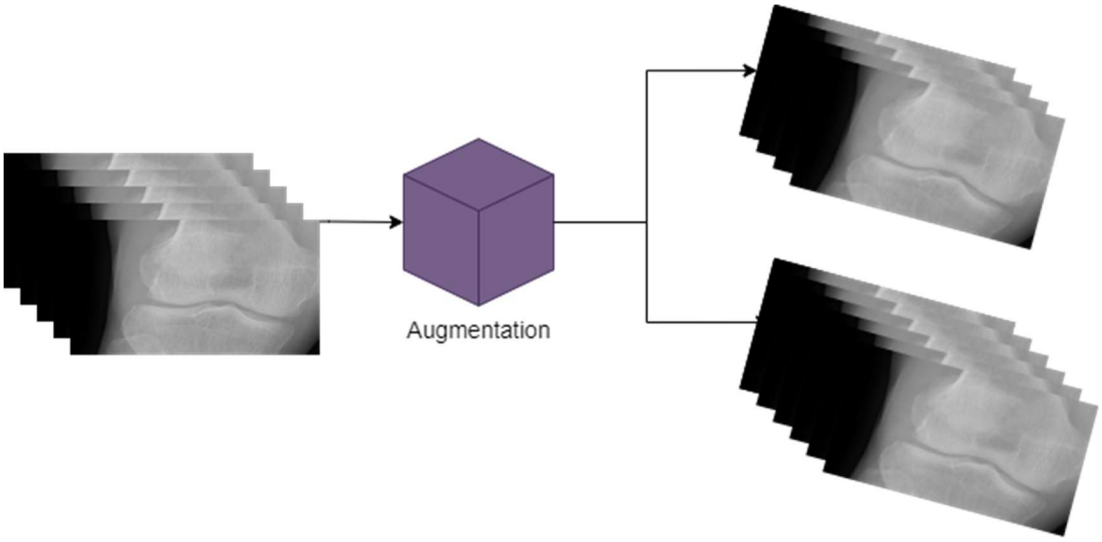

Figure 2 Generating New Sample By Flipping

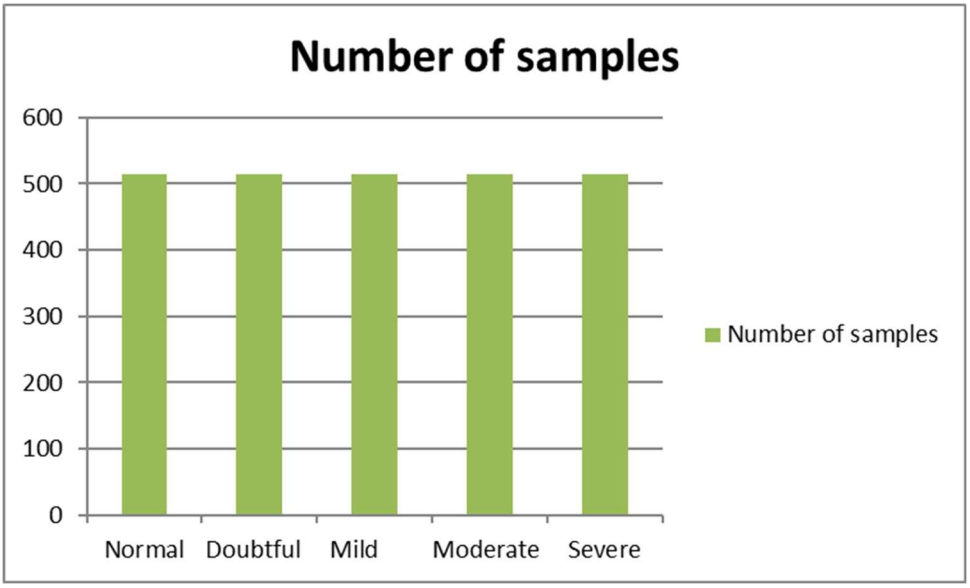

Figure 3 Balanced Number Of Samples Over All Classes

After augmenting all target classes to achieve a balanced distribution, all samples were resized to 224x224 pixels in grayscale. Subsequently, pixel values were scaled between -1 and 1 using Equation (1). The X-ray samples were then divided into training, testing, and validation sets in a 0.75, 0.15, and 0.10 ratio, respectively and randomly. This allocation resulted in 1925 samples for training, 260 for testing, and 385 for validation. Figure 4 visually depicts sample X-ray images.

$$Scale = \frac{\sum_{i=0}^{n} IMG_i}{127.5-1} \quad (1)$$

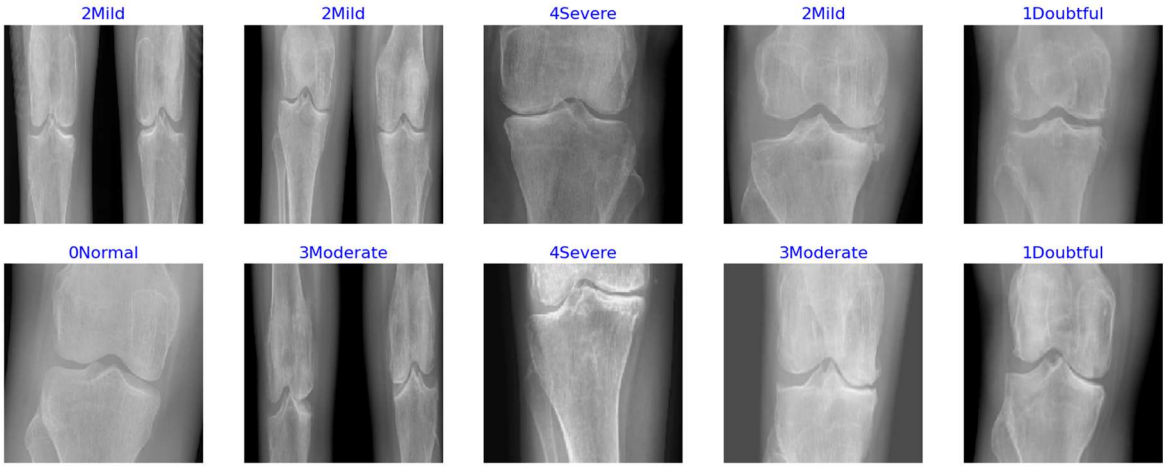

Figure 4 Sample Knee X-Ray Images

3.2 Model Implementation

We implemented a CNN model, illustrated in Figure 1, comprising three convolutional and max-pooling layers. A 3x3 convolutional filter is applied in each convolutional layer, with the filters set to 16, 32, and 64, respectively. A max-pooling layer with a filter size of 2x2 follows each convolutional layer. Regularization is applied to the first Conv2D layer with a penalty factor of 0.01, and ReLU activation functions are employed in each layer.

After three iterations of convolution and pooling, a fully connected layer is introduced to flatten the data. Subsequently, it is connected to a dense layer with a size of 5, allowing the model to map to one of the five classes. The Adam optimizer updates weights, and a softmax activation function is applied to the final dense layer.

During training, the model employs an initial learning rate of 0.001, and changed dynamically with accuracy, loss. Cross-entropy is the loss metric, and accuracy is recorded after each iteration. The model undergoes training for 50 epochs, and both loss and validation loss are monitored to evaluate its performance. This architecture aims to effectively capture and classify features in the input data for accurate prediction across multiple classes.

4. RESULT ANALYSIS

We utilized a Kaggle knee X-ray dataset, augmenting it to address sample imbalances. The data was then trained on an optimized CNN model, exploring various batch sizes, learning rates, and epochs to identify the best hyperparameters. The model underwent training for 50 epochs, and performance metrics such as accuracy, precision, recall, and F1-score were evaluated.

Figure 5 illustrates the training and validation loss and accuracy trends of the model before augmentation. From it is clearly observed that the model under fitted, and due to unbalanced data it is biases from Figure 13. Figure 6 the loss and accuracy is after adding artificial data, the results are improved. But after augmentation and increasing the model complexity the model performed well as shown Figure 7. Initially starting at 23.0 and 6.5 for loss and accuracy, respectively, the model demonstrated notable improvement epoch by epoch. Training, validation loss, and accuracy indicated a well-balanced model that overfit and underfit the data.

The ROC curve, depicted in Figure 8 and 9, revealed that before augmentation the model is completely biased, and after augmentation and the model provided a higher true positive rate in balanced data than in unbalanced data, resulting in a final area under the curve (AUC) of 0.94. Table 2, presenting weighted and micro-average metrics, affirms the consistency and superior performance of the model. The model is performed consistently from Figure 10, 11 and 12 the precision is in between 0.93 to 0.95, and recall is fluctuated in between 0.93 to 0.96, and F1score is moved in between 0.93 to 0.95, so here it is observed a small difference like 0.02 from Figure 15. The confusion matrix in Figure 13 and 14 indicates optimal true positive rates for each class and a favorable false-negative outcome.

$$F1\ score = \frac{2 * precision\ * recall}{precision + recall} \quad (2)$$

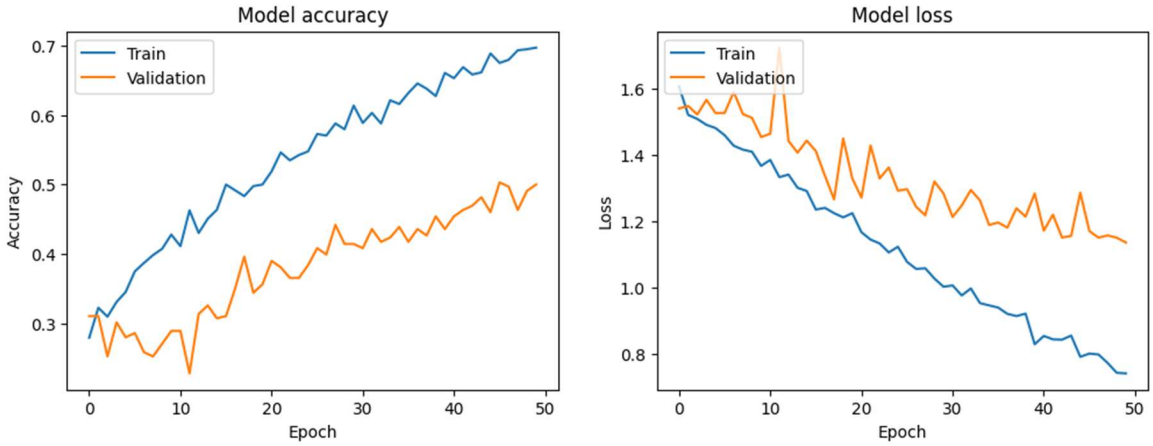

Figure 5: Training And Validation Loss, Accuracy Before Augmentation

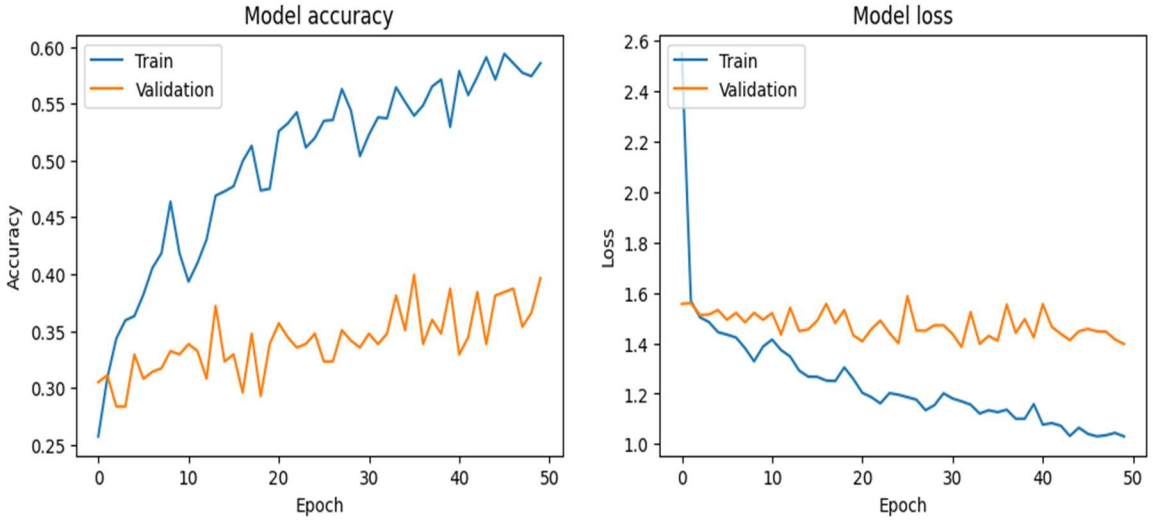

Figure 6: Training And Validation Loss, Accuracy After Augmentation

Journal of Theoretical and Applied Information Technology
30th April 2025. Vol.103. No.8

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

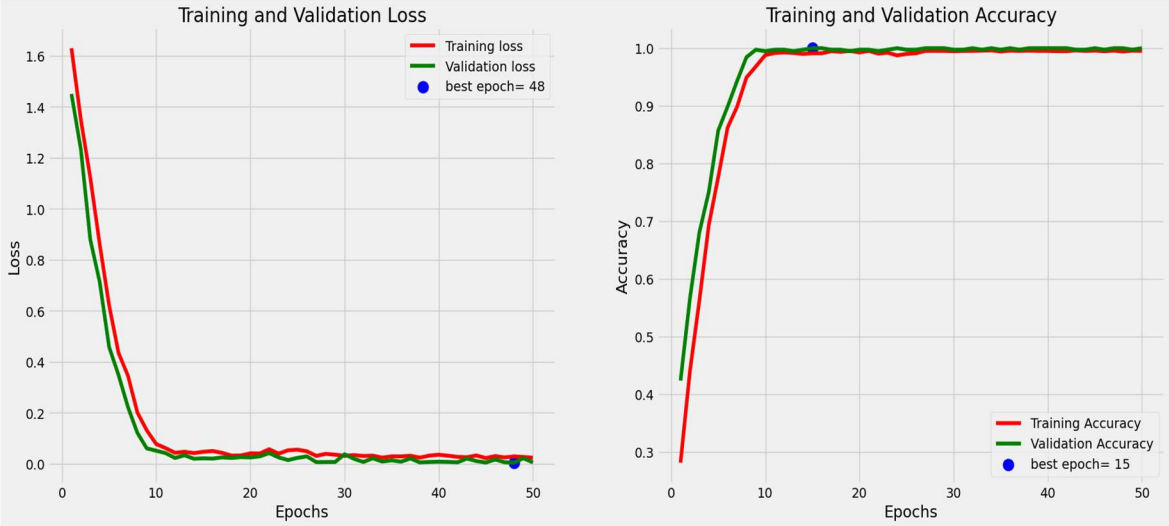

Figure 7: Loss And Accuracy Of Training And Validation For Proposed Model

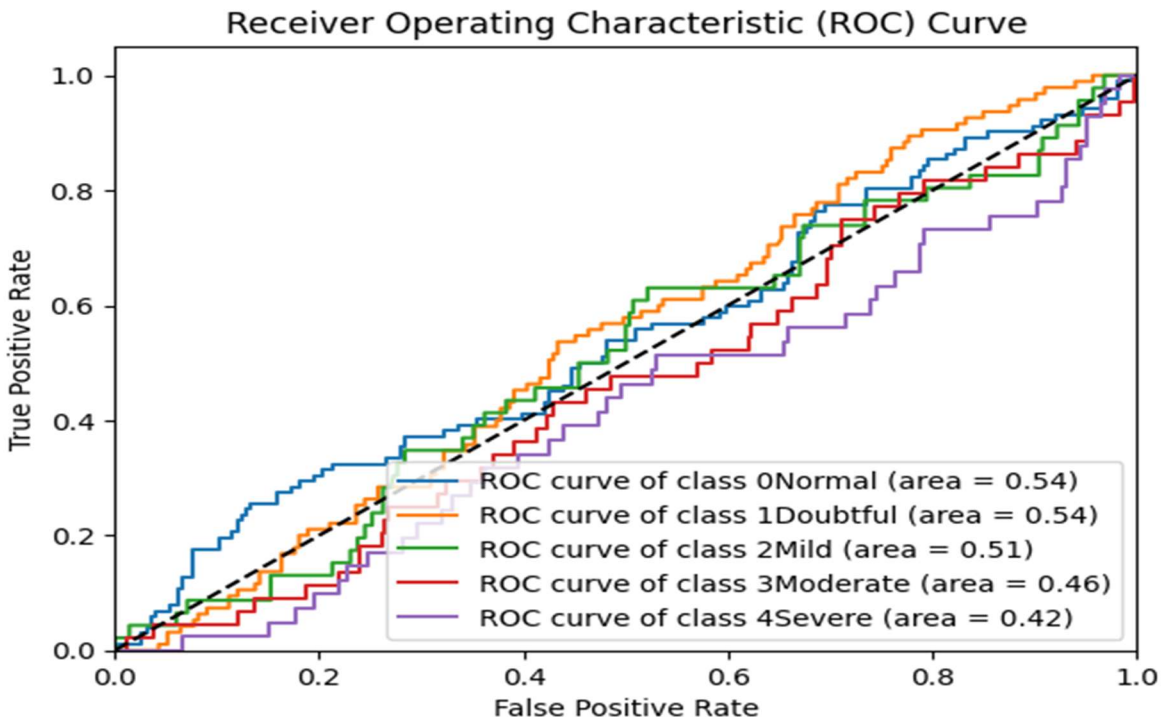

Figure 8: ROC Curves Of All Classes Before Augmentation

Journal of Theoretical and Applied Information Technology
30th April 2025. Vol.103. No.8

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

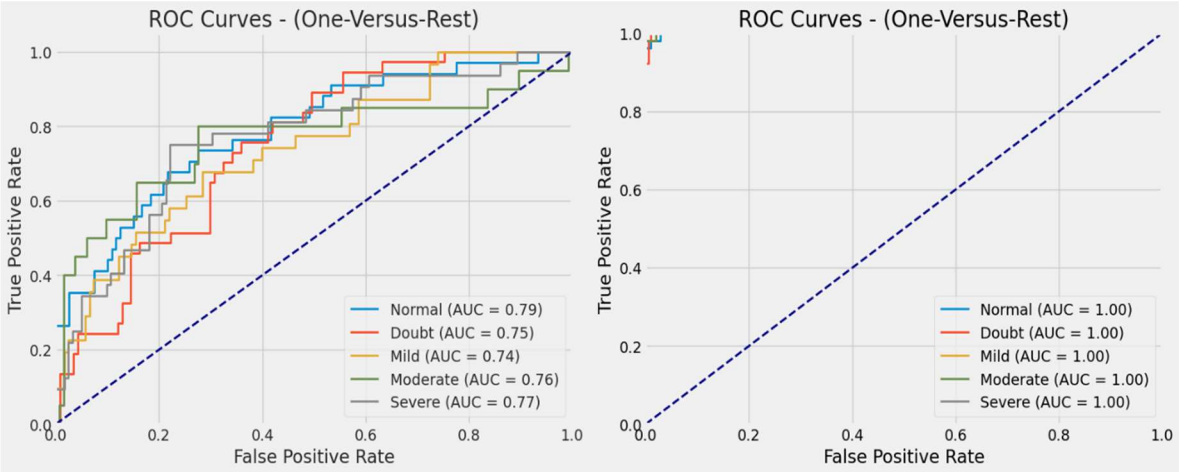

Figure 9: ROC Curve For Proposed Model Before And After Augmentation

In Table 2, the accuracy, representing the actual and predicted label values ratio, is approximately 0.94 for this model. This value indicates a high level of correctness in the model's predictions. The support count, which illustrates the consistency of the model in making predictions, demonstrates that the model has a sufficient number of instances to substantiate its performance.

Beyond accuracy, precision and recall are valuable metrics that signify the ratio of samples correctly predicted to the total number of samples that should be predicted. In Table 2, these ratios are observed to be high, further reinforcing the model's efficacy. Precision and recall provide insights into the model's ability to make accurate positive predictions and its sensitivity to correctly identifying positive instances.

Table 2: Precision And Recall And F1-Score Of Proposed Models

	P	R	F1	Support
Normal	0.94	0.95	0.95	52
Doubtful	0.93	0.94	0.93	52
Mild	0.95	0.95	0.96	52
Moderate	0.93	0.93	0.93	52
Server	0.94	0.95	0.94	52
Acc	-	-	0.94	260
M-AVG	0.94	0.95	0.95	260
W-AVG	0.95	0.95	0.95	260

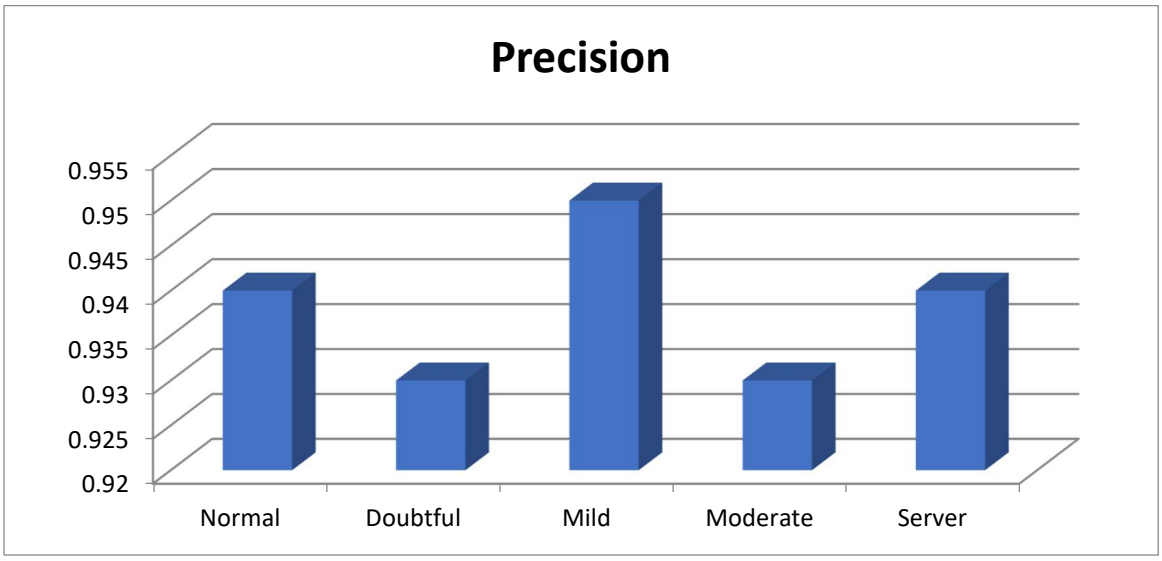

Figure 10: Comparison Of Precision Of All Classes

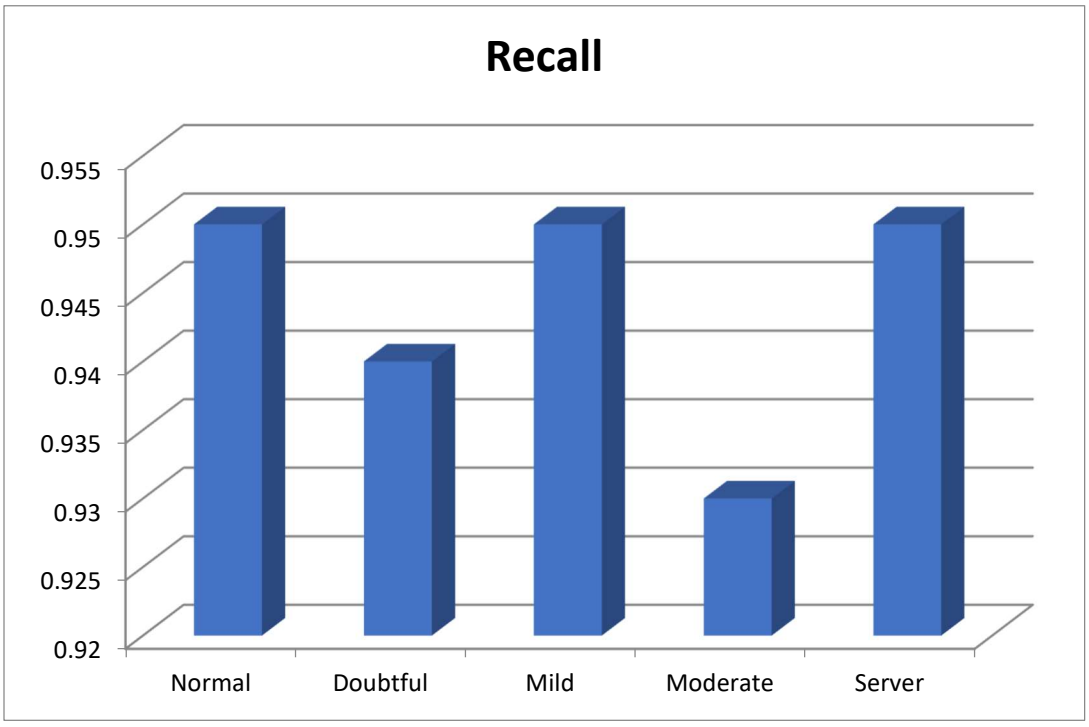

Figure 11: Comparison Of Recall Of All Classes

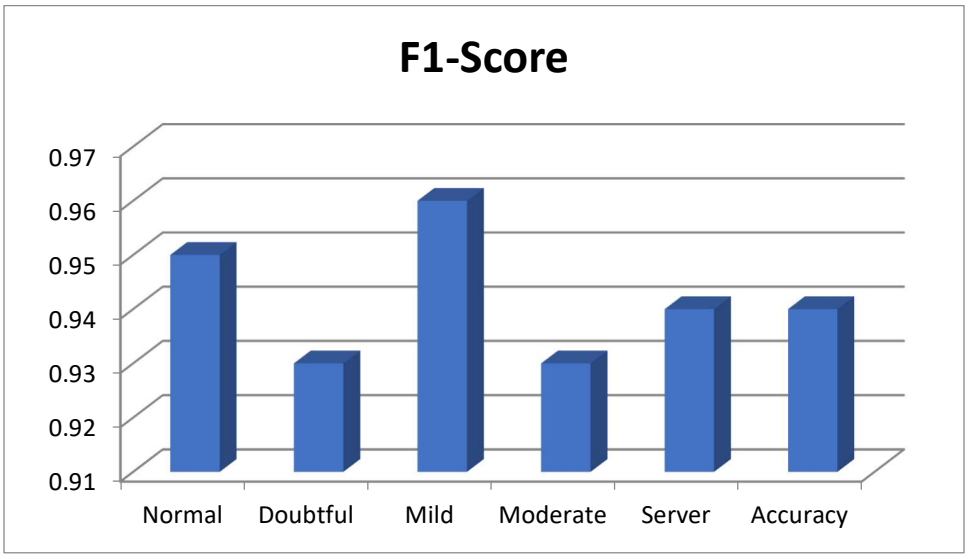

Figure 12: Comparison Of F1-Score Of All Classes

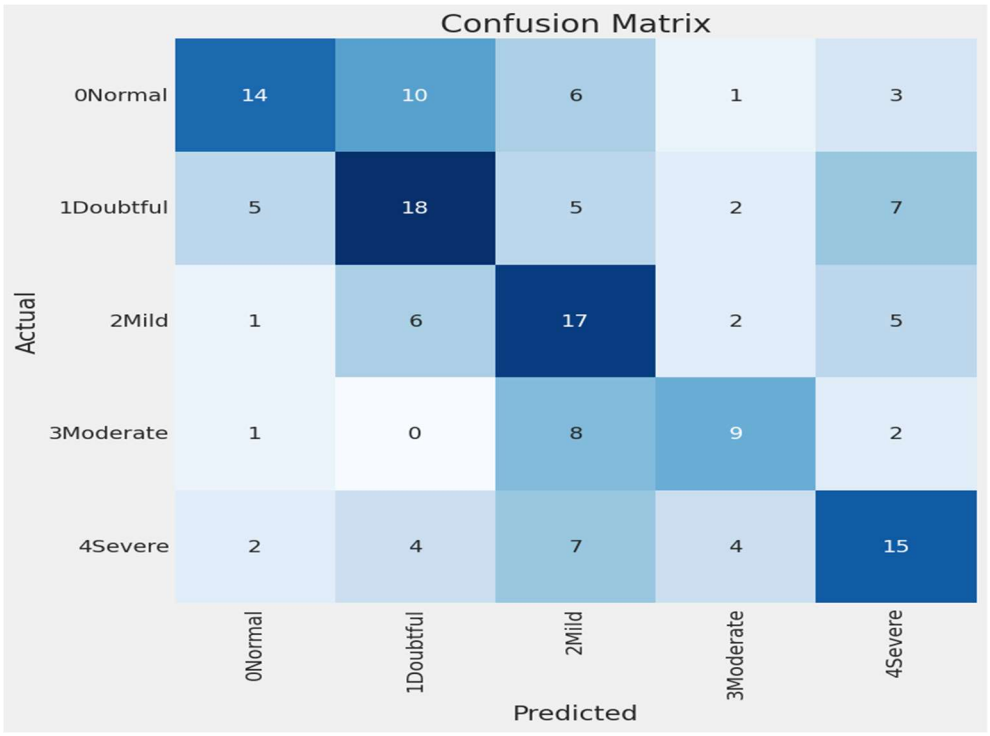

Figure 13: Confusion Matrix Before Augmentation

Journal of Theoretical and Applied Information Technology
30th April 2025. Vol.103. No.8

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

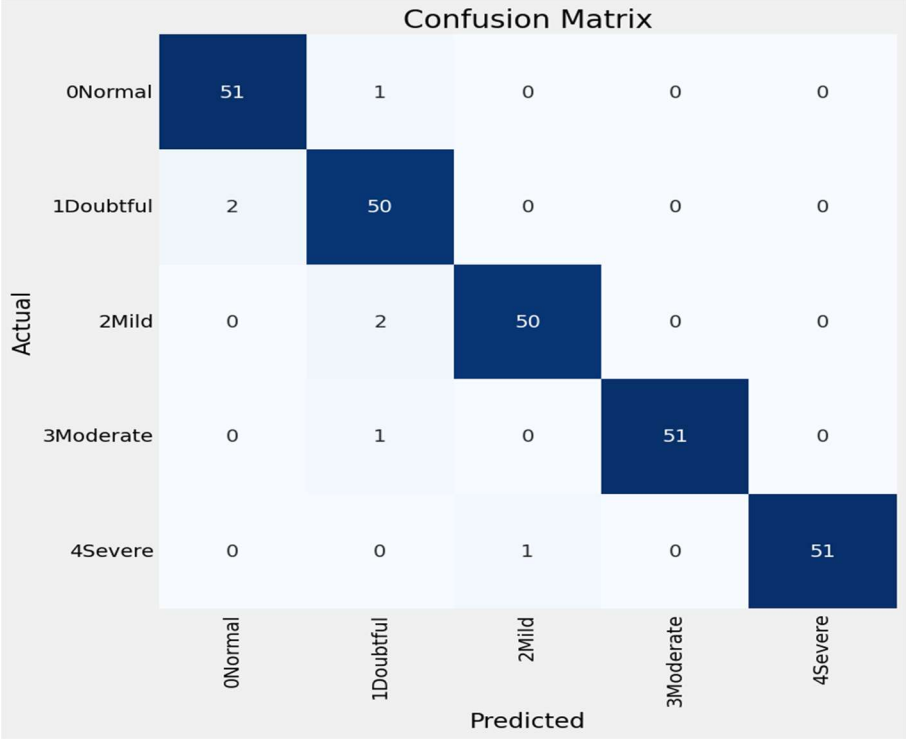

Figure 14: Confusion Matrixes For All 5 Classes Of Proposed Model After Augmentation

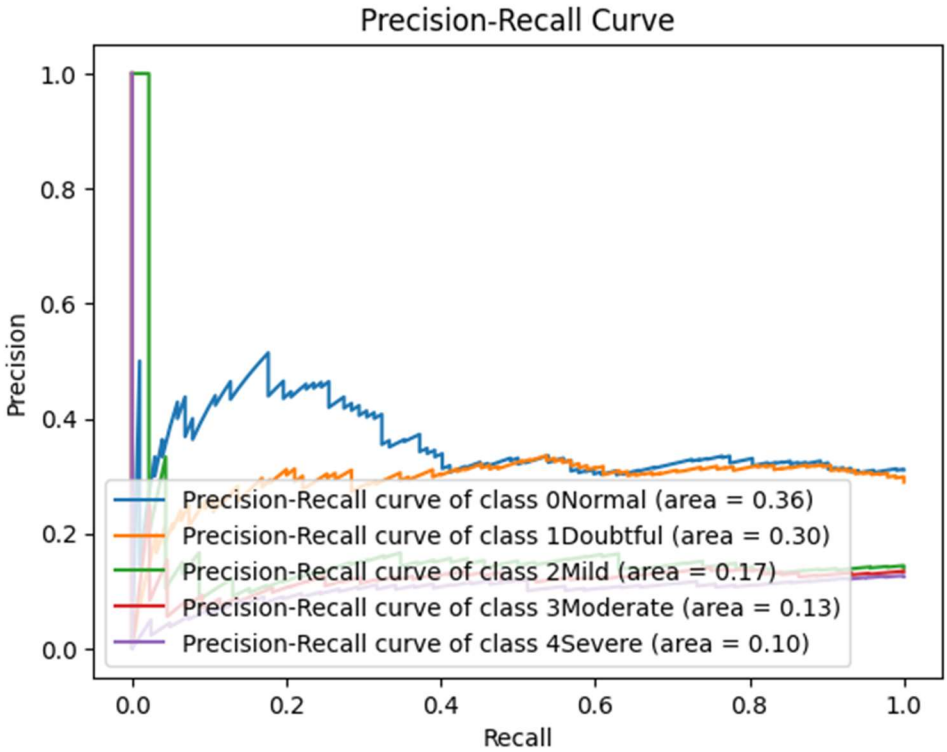

Figure 15 Precision And Recall Curves Before Augmentation

4.1. Comparison of proposed model with existing models:

Table 3 provides a comparative analysis of various deep learning and machine learning methodologies applied to RA diagnosis and prediction using diverse datasets. Several approaches have been employed, including CNNs, hybrid segmentation models, and AI-driven clinical decision support systems. For instance, RA-XTNet, a novel CNN model utilizing hand radiographs and thermal images, achieved an accuracy of 80.60% [7], whereas a standard CNN model for detecting RA from hand radiographs obtained 73.3% accuracy [14]. Similarly, deep learning methods leveraging MRI scans, such as consistency-based learning [15] and AI-based decision support systems [16], reported accuracy levels of 67.80% and 92.81%, respectively. Furthermore, radiograph-based RA detection using VGG and GoogleNet demonstrated a high accuracy of 96.15% [17], while Bayesian optimization for osteoporotic fracture prediction achieved 91.20% accuracy [18]. Additionally, hybrid segmentation models, such as RANET for RA diagnosis using X-ray images, reported an accuracy of 80.56% [19]. Other notable models include RA2-DREAM, which applied deep learning to assess joint damage progression in RA, yielding RMSEs of 35.0 [23], and a chronic disease-based osteoporosis risk prediction model that achieved an AUC of 0.773 [26]. Advanced deep learning methods, including AI quantification of synovium using dynamic contrast-enhanced MRI, exhibited AUC values as high as 0.941 [27], while discrimination of RA from osteoarthritis on hand radiographs reached an accuracy of 87.2% [29]. A HarDNet-based model for bone mineral density inference from hand radiographs achieved 80% accuracy [30]. Additionally, machine learning techniques have been used to predict inadequate responders to methotrexate in RA patients, yielding an AUROC of 0.72 [32], while multilevel modeling of joint damage in RA demonstrated a Pearson's correlation of 0.711 [33]. Automated software designed to detect joint space narrowing progression in RA showed significant results with $p = 0.004$ [34].

Table 3 Comparison Of Proposed Model With Existing Models

Ref. No.	Methodology	Dataset Used	Accuracy Percentage
[7]	RA-XTNet: A novel CNN model to predict RA from hand radiographs and thermal images.	Hand radiographs and thermal images.	80.60%
[14]	CNN model for detecting RA from hand radiographs.	Hand radiographs.	73.3%
[15]	Consistency-based deep learning using extremity MRI scans for RA classification and prediction.	MRI scans.	67.80%
[16]	AI-enabled clinical decision support system for diagnosing RA using X-ray images.	X-ray images from RA patients.	92.81%
[17]	Radiograph-based RA diagnosis via VGG, GoogleNET	Radiograph images.	96.15%
[18]	Bayesian optimization for enhanced osteoporotic fracture prediction in postmenopausal women.	Genetic and clinical risk data.	91.20%
[19]	Hybrid segmentation algorithm for RA diagnosis using X-ray images. RANET model	X-ray images.	80.56%
[23]	Deep learning to automatically detect joint damage progression in RA. RA2-DREAM algorithm	Radiographs from RA patients.	RMSEs 35.0
[26]	Prediction of osteoporosis risk using nationwide chronic disease data with machine learning.	Chronic disease database.	AUC: 0.773
[27]	AI quantification of synovium in RA using dynamic contrast-enhanced MRI.	DCE-MRI scans.	AUC ranged from 0.941%

[29]	Deep learning discrimination of RA from osteoarthritis on hand radiography.	RA or NOT	87.2%
[30]	HarDNet-based deep learning model for bone mineral density inference from hand radiographs.	Hand radiographs.	80%
[32]	Machine learning to identify profiles of inadequate responders to methotrexate in RA.	Clinical data on methotrexate-treated patients.	AUROC 0.72
[33]	Multilevel modeling of joint damage in RA using machine learning.	Clinical and imaging data.	Pearson's correlation 0.711
[34]	Automated software for detecting radiographic joint space narrowing progression in RA.	Radiographs and phantom study data.	GSS progression (p = 0.004)
*	Augmented CNN model	knee medical X-ray images	94%

The table 3 presents a comparison of prescribed model with proposed model for the detection, prediction, and outcome assessment of RA using different datasets. From all these models vein, Li, Y et al [15] with a CNN model to radiographs, enhanced with image pre-processing techniques, and attained an accuracy of 67.80%. Scientific, L. Let al [16] further explored DL models for RA detection using ultrasound images, resulting in 92.81% accuracy.

Additionally, when the model have used clinical data for RA prediction, like Peng, Y. et al.[17] with simple ML models to clinical data, provided an accuracy of 96.16%, while Wu, Q., and Dai, J.[18] with same ML algorithms to predict RA disease activity based on clinical data and biomarkers, provided an accuracy of 91.2%. And [19] with DL methods on radiographic images got an accuracy of 80.56%. And [23] with an automated ML approach for evaluating radiographs, performed with RMSEs 35.0. The application of DL to MRI images like [29] and [30] performed well with an accuracy of 87.20% and 80.00%.

The proposed model an augmented CNN model specifically designed to analyze knee medical X-ray images, achieving a superior accuracy of 94%. This model outperforms several existing approaches, highlighting the effectiveness of combining CNN architecture with augmentation techniques to improve RA detection accuracy from X-ray images.

5. CONCLUSION

The proposed CNN-based knee X-ray classification model demonstrated high effectiveness in detecting rheumatoid arthritis. To address class imbalance in the Kaggle dataset (1,650 samples), we applied data augmentation techniques, including image flipping, resulting in a well-balanced dataset with 514 images per class. Our customized CNN model, optimized with fine-tuned hyperparameters, exhibited superior performance over 50 epochs, consistently outperforming other baseline models.

The ROC curve analysis validated the model's discriminative capability, achieving an AUC of 0.94, emphasizing its high true positive rate. Further, evaluation metrics reinforced the model's robustness, with an average precision of 0.95, recall of 0.94, and F1-score of 0.95. These findings confirm that data augmentation significantly enhances model performance, making it a valuable approach for improving AI-driven RA detection.

This study demonstrates the potential of deep learning in medical imaging for early RA diagnosis. Future work will focus on expanding the dataset, incorporating explainable AI techniques, and validating the model on multi-source medical images to enhance clinical applicability and trustworthiness.

LIMITATIONS & FUTURE WORK

- The model was trained on a single dataset, which may impact its generalizability to real-world clinical scenarios. Future work should validate the model using multi-source datasets, including different imaging modalities such as MRI and ultrasound.
- Despite achieving a high AUC of 0.94, the model remains a black-box system with limited

interpretability. The integration of Explainable AI (XAI) techniques, such as Grad-CAM, can enhance model transparency and clinical trustworthiness.

REFERENCES:

[1] Aletaha, D., Neogi, T., Silman, A. J., Funovits, J., Felson, D. T., Bingham, C. O., & Hawker, G. (2010). 2010 Rheumatoid arthritis classification criteria: An American College of Rheumatology/ European League Against Rheumatism collaborative initiative. Arthritis & Rheumatism,62(9),2569-2581. https://doi.org/10.1002/art.27584

[2] Momtazmanesh S, Nowroozi A, Rezaei N. Artificial Intelligence in Rheumatoid Arthritis: Current Status and Future Perspectives: A State-of-the-Art Review. RheumatolTher. 2022 Oct;9(5):1249-1304.Epub 2022 Jul 18. PMID: 35849321; PMCID: PMC9510088. https://doi:10.1007/s40744-022-00475-4.

[3] Kay, J., & Upchurch, K. S. (2012). ACR/EULAR 2010 rheumatoid arthritis classification criteria. Rheumatology,51(suppl_6),vi5i9. https://doi.org/:10.1093/rheumatology/kes279

[4] MengfangLi , Yuanyuan Jiang, Haisheng Zhu, Yanzhou Zhang Medical image analysis using deep learning algorithms 2023, https://doi.org/10.3389/fpubh.2023.1273253.

[5] Schaffert, D., Bibi, I., Blauth, M., Lull, C., von Ahnen, J. A., Gross, G., ... & Olsavszky, V. (2024). Using Automated Machine Learning to Predict Necessary Upcoming Therapy Changes in Patients With Psoriasis Vulgaris and Psoriatic Arthritis and Uncover New Influences on Disease Progression: Retrospective Study. JMIR Formative Research, 8, e55855. https://doi.org/10.2196/55855.

[6] Hügle, T. (2024). Advancing Rheumatology Care Through Machine Learning. Pharmaceutical Medicine, 38(2), 87-96. https://doi.org/10.1007/s40290-024-00515-0.

[7] Kesavapillai, A. R., Aslam, S. M., Umapathy, S., & Almutairi, F. (2024). RA-XTNet: A Novel CNN Model to Predict Rheumatoid Arthritis from Hand Radiographs and Thermal Images: A Comparison with CNN Transformer and Quantum Computing. Diagnostics, 14(17), 1911. https://doi.org/10.3390/diagnostics14171911.

[8] Betancourt-Hernández, M., Viera-López, G., & Serrano-Muñoz, A. (2018). Automatic diagnosis of rheumatoid arthritis from hand radiographs using convolutional neural networks. Revista Cubana de Fisica, 35(1), 39-43.

[9] Panwar, P., Chaurasia, S., Gangrade, J., Bilandi, A., & Pruthviraja, D. (2024). Optimizing knee osteoarthritis severity prediction on MRI images using deep stacking ensemble technique. Scientific Reports, 14(1), 26835. https://doi.org/10.1038/s41598-024-78203-x.

[10] Cush JJ. Rheumatoid Arthritis: Early Diagnosis and Treatment. Med Clin North Am. 2021 Mar;105(2):355-365. DOI: 10.1016/j.mcna.2020.10.006.

[11] Huo Y, Vincken KL, van der Heijde D, De Hair MJ, Lafeber FP, Viergever MA. Automatic Quantification of Radiographic Finger Joint Space Width of Patients With Early Rheumatoid Arthritis. IEEE Trans Biomed Eng. 2016 Oct;63(10):2177-86. Epub 2015 Dec 29. PMID: 26731521 .https://doi.org/10.1109/TBME.2015.2512941

[12] Takeda, R., Matsumoto, T., Maenohara, Y., Omata, Y., Inui, H., Nagase, Y., ... & Tanaka, S. (2022). Increasing trend of radiographic features of knee osteoarthritis in rheumatoid arthritis patients before total knee arthroplasty. Scientific Reports, 12(1), 10452. https://doi.org/10.1038/s41598-022-14440-2.

[13] Üreten, K., & Maraş, H. H. (2022). Automated classification of rheumatoid arthritis, osteoarthritis, and normal hand radiographs with deep learning methods. Journal of Digital Imaging,35(2),193-199. https://doi.org/10.1007/s10278-021-00564-w.

[14] Üreten K, Erbay H, Maraş HH. Detection of rheumatoid arthritis from hand radiographs using a convolutional neural network. ClinRheumatol. 2020 Apr;39(4):969-974. https://doi: 10.1007/s10067-019-04487-4. Epub 2019 Mar 8. PMID: 30850962 .DOI: 10.1007/s10067-019-04487-4

[15] Li, Y., Hassanzadeh, T., Shamonin, D. P., Reijnierse, M., van der Helm-van, A. H., & Stoel, B. C. (2024). Rheumatoid arthritis classification and prediction by consistency-based deep learning using extremity MRI scans. Biomedical Signal Processing and Control, 91, 105990. https://doi.org/10.1016/j.bspc.2024.105990

[16] Scientific, L. L. (2024). AIRADF: ARTIFICIAL INTELLIGENCE ENABLED CLINICAL DECISION SUPPORT SYSTEM FOR DIAGNOSING RHEUMATOID ARTHRITIS USING X-RAY IMAGES. Journal of Theoretical and Applied Information Technology, 102(19). https://www.jatit.org/volumes/Vol102No19/24Vol102No19.pdf.

[17] Peng, Y., Huang, X., Gan, M., Zhang, K., & Chen, Y. (2024). Radiograph-based rheumatoid arthritis diagnosis via convolutional neural network. BMC Medical Imaging, 24(1), 180. https://doi.org/10.1186/s12880-024-01362-w.

[18] Wu, Q., & Dai, J. (2024). Enhanced osteoporotic fracture prediction in postmenopausal women using Bayesian optimization of machine learning models with genetic risk score. Journal of Bone and Mineral Research, 39(4), 462-472. https://doi.org/10.1093/jbmr/zjae025.

[19] Rajesh, G., Malarvizhi, N., & Leung, M. F. (2024). A Hybrid Segmentation Algorithm for Rheumatoid Arthritis Diagnosis Using X-ray Images. Big Data and Cognitive Computing, 8(9), 104. https://doi.org/10.3390/bdcc8090104.

[20] Gandrup, J., Selby, D. A., & Dixon, W. G. (2024). Classifying Self-Reported Rheumatoid Arthritis Flares Using Daily Patient-Generated Data From a Smartphone App: Exploratory Analysis Applying Machine Learning Approaches. JMIR Formative Research, 8, e50679. https://doi.org/10.2196/50679.

[21] Chen, Y. J., Chen, D. Y., Lan, H. C., Huang, A. C., Chen, Y. H., Huang, W. N., & Chen, H. H. (2024). An optimal deep learning model for the scoring of radiographic damage in patients with ankylosing spondylitis. Therapeutic Advances in Musculoskeletal Disease, 16, 1759720X241285973. https://doi.org/10.1177/1759720X241285973

[22] Moradmand, H., & Ren, L. (2024). Automated Radiographic Total Sharp Score (ARTSS) in Rheumatoid Arthritis: A Solution to Reduce Inter-Intra Reader Variation and Enhancing Clinical Practice. https://doi.org/10.21203/rs.3.rs-4673875/v1.

[23] Venäläinen, M. S., Biehl, A., Holstila, M., Kuusalo, L., & Elo, L. L. (2024). Deep learning enables automatic detection of joint damage progression in rheumatoid arthritis—model development and external validation. Rheumatology, keae215.doi:10.1093/rheumatology/keae215.

[24] Likhith, R., Likith, M. C., Bhat, L. R., Kushal, U., & Priyanka, H. (2024, October). Machine Learning and Deep Learning for Knee Osteoarthritis Diagnosis: A Survey on Classification and Severity Grading. In 2024 IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER) (pp. 345-352). IEEE.https://doi.org/10.1109/DISCOVER62353.2024.10750563.

[25] Alam, A., Ahamad, M. K., Mohammed Aarif, K. O., & Anwar, T. (2024). Detection of rheumatoid arthritis using CNN by transfer learning. In Artificial Intelligence and Autoimmune Diseases: Applications in the Diagnosis, Prognosis, and Therapeutics (pp. 99-112). Singapore: Springer Nature Singapore. https://doi.org/:10.1007/978-981-99-9029-0_5.

[26] Tu, J. B., Liao, W. J., Liu, W. C., & Gao, X. H. (2024). Using machine learning techniques to predict the risk of osteoporosis based on nationwide chronic disease data. Scientific Reports, 14(1), 5245. https://doi.org/10.1038/s41598-024-56114-1.

[27] Mao, Y., Imahori, K., Fang, W., Sugimori, H., Kiuch, S., Sutherland, K., & Kamishima, T. (2024). Artificial Intelligence Quantification of Enhanced Synovium Throughout the Entire Hand in Rheumatoid Arthritis on Dynamic Contrast-Enhanced MRI. Journal of Magnetic Resonance Imaging.DOI: 10.1002/jmri.29463.

[28] Stoel, B. C., Staring, M., Reijnierse, M., & van der Helm-van Mil, A. H. (2024). Deep learning in rheumatological image interpretation. Nature Reviews Rheumatology, 20(3), 182-195. https://doi.org/10.1038/s41584-023-01074-5.

[29] Ma, Y., Pan, I., Kim, S. Y., Wieschhoff, G. G., Andriole, K. P., & Mandell, J. C. (2024). Deep learning discrimination of rheumatoid arthritis from osteoarthritis on hand

radiography. Skeletal Radiology, 53(2), 377-383. https://doi.org/10.1007/s00256-023-04408-2.

[30] Ho, C.S., Fan, T. Y., Kuo , C. F., Yen, T. Y., Chang, S. Y., Pei, Y. C., & Chen, Y. P. (2024). HarDNet-based deep learning model for osteoporosis screening and bone mineral density inference from hand radiographs. Bone, 117317. https://doi.org/10.1016/j.bone.2024.117317.

[31] Hassanzadeh, T., Shamonin, D. P., Li, Y., Krijbolder, D. I., Reijnierse, M., van der Helm-van, A. H., & Stoel, B. C. (2024). A deep learning-based comparative MRI model to detect inflammatory changes in rheumatoid arthritis. Biomedical Signal Processing and Control, 88, 105612. https://doi.org/10.1016/j.bspc.2023.105612.

[32] Duquesne, J., Bouget, V., Cournede, P. H., Fautrel, B., Guillemin, F., de Jong, P. H., ... & Bitoun, S. (2023). Machine learning identifies a profile of inadequate responder to methotrexate in rheumatoid arthritis. Rheumatology, 62(7), 2402-2409. https://doi.org/10.1093/rheumatology/keac645.

[33] Li H, Guan Y. Multilevel Modeling of Joint Damage in Rheumatoid Arthritis. Adv Intell Syst. 2022 Nov;4(11):2200184.Epub 2022 Oct 13. PMID: 37808948; PMCID: PMC10557461. https://doi.org/10.1002/aisy.202200184.

[34] Okino T, Ou Y, Ikebe M, Tamura K, Sutherland K, Fukae J, Tanimura K, Kamishima T. Fully automatic software for detecting radiographic joint space narrowing progression in rheumatoid arthritis: phantom study and comparison with visual assessment. Jpn J Radiol. 2023 May;41(5):510-520. Epub 2022 Dec 20. PMID: 36538163.https://doi.org/10.1007/s11604-022-01373-z.

[35] Kato K, Yasojima N, Tamura K, Ichikawa S, Sutherland K, Kato M, Fukae J, Tanimura K, Tanaka Y, Okino T, Lu Y, Kamishima T. Detection of Fine Radiographic Progression in Finger Joint Space Narrowing Beyond Human Eyes: Phantom Experiment and Clinical Study with Rheumatoid Arthritis Patients. Sci Rep. 2019 Jun 12;9(1):8526. PMID: 31189913; PMCID: PMC6561904. https://doi.org/10.1038/s41598-019-44747-6