ISSN: 2182-2069 / E-ISSN: 2182-2077

An Integrated CNN-RNN Framework for Enhanced Rheumatoid Arthritis Detection and Diagnosis from Medical Imaging

S. Binny[1*], and Dr.P. Sardar Maran[2]

[1*]Sathyabama Institute of Science and Technology (Deemed to Be University), Chennai, India.
binnylatheesh@gmail.com, https://orcid.org/0000-0003-0088-5331

[2]Professor, Department of Computer Science & Engineering School of Computing, Sathyabama Institute of Science and Technology (Deemed to Be University), Chennai, Tamil Nadu, India.
psmaran@sathyabama.ac.in, https://orcid.org/0000-0001-5595-0490

Received: January 17, 2025; Revised: February 28, 2025; Accepted: April 10, 2025; Published: May 30, 2025

Abstract

Rheumatoid arthritis (RA) encompasses damage to joint tissues, and its pathology can be diagnosed using available medical imaging modalities like X-rays, MRI, and Ultrasound. It is one of the lifelong progressive autoimmune diseases that require early detection and prompt treatment to improve patient outcomes. This study proposes an integrated convolutional neural network-recurrent neural network (CNN-RNN) framework to facilitate RA diagnosis by jointly using spatial and temporal features from sequentially acquired medical images. The model was trained and evaluated by using over 3000 radiographic images acquired from 500 patients, including patient-wise split into training (70%), validation (15%), and testing (15%) subsets. Intensive preprocessing included contrast enhancement, noise reduction, and normalization of image quality. The CNN module extracts discriminative spatial features from individual frames, while the RNN component captures the progress of the disease across several time points. Experimental assessment showed that this integrated model achieves a classification accuracy of 94.8%, precision-93.6%, recall-95.1%, and F1-score of 94.3%, significantly superior to traditional CNN- and SVM-based methods. The statistical significance was established using McNemar's test ($p<0.01$), thus confirming the superiority of the proposed approach. Together, these findings emphasize the potential of deep spatial-temporal learning towards the more accurate and earlier detection of RA in clinical settings.

Keywords: Rheumatoid Arthritis, Medical Imaging, Deep Learning, Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Computer-aided Diagnosis, Image Classification.

1 Introduction

It is defined as a chronic progressive systemic autoimmune disease leading to the persistent inflammation of the synovial membrane and further joint damage, functional incapacity, and poor quality of life (Finckh et al., 2022). Rheumatoid arthritis (RA) primarily affects the small joints of the hands and feet and often has symmetrical joint involvement. The worldwide burden of RA is considerable and is likely to grow in many parts of the globe as populations age and more recognition

Journal of Internet Services and Information Security (JISIS), volume: 15, number: 2 (May), pp. 103-124.
DOI: *10.58346/JISIS.2025.I2.008*

*Corresponding author: Sathyabama Institute of Science and Technology (Deemed to Be University), Chennai, India.

of the disease occurs (GBD 2021 Rheumatoid Arthritis Collaborators, 2023). Early identification is crucial for RA management because it allows treatment in the early stages of the disease progression to reduce irreversible joint destruction (Zhu et al., 2022). However, this is complicated by the overlap in some symptoms with other inflammatory joint diseases and the variability of clinical presentations, which makes early and accurate detection a continuing source of frustration for clinicians.

Recent immunopathological advancements have highlighted those molecular mechanisms, such as citrullination and autoantibody production, which underpin RA and, thus, promise the potential to develop early biomarker-based diagnosis (Verywell Health; Ali et al., 2023). However, with advances in therapeutic options, such as the employment of DMARDs and other biologics, automated diagnostics remain in high demand in medical imaging. The application of artificial intelligence, particularly those deep learning approaches, in diagnostics, is likely to transform RA detection and improve clinical decision-making with much greater accuracy (Ali et al., 2023; Biswas & Tiwari, 2024).

The persistent inflammation in joints could markedly lead to cartilage and bone destruction, functional disability, and even reduced quality of life. Patient numbers for rheumatoid arthritis (RA) could be estimated as approximately 0.5–1% of people around the globe, with an evident higher prevalence in the female and elderly groups (Safiri et al., 2019). This disease usually manifests with symmetrical joint pain, stiffness, and swelling, predominantly in the small joints of hands and feet. Early diagnosis and identification of RA are necessary for speedy treatment to save joints from irretrievable damage. Evidence has shown that the patients treated with disease-modifying anti-rheumatic drugs (DMARDs) within 3 months of initial symptoms had significantly better prognoses than patients treated at much longer intervals (Smolen et al., 2016). Furthermore, the disease will progress less, improve physical functioning, and reduce the risk for diseases such as cardiovascular co-morbidity (van der Helm-van Mil & Huizinga, 2008).

Conventional diagnostic approaches, clinical assessments, various serological tests (such as rheumatoid factor and anti-CCP antibodies), and imaging modalities like X-ray, MRI, and ultrasound are still obvious. Still, these tests often hinge on the clinician's judgment expertise. They may be unable to detect the disease at early stages with reasonable accuracy (Tan & Conaghan, 2009). Therefore, the need is for new systems of diagnosis based on the automated, advanced, and objective measures.

The wonderful systems brought in by deep learning techniques such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) from the clarity of RA diagnosis (Alzaidi, 2024). These models can learn the complexity of features in medical images, leading the disease patterns to be identified even at an initial stage, at more genuinely sensitive and specific measures. The capabilities of the combined approach are even further enhanced with the incorporation of CNNs with RNNs since this facilitates spatial and sequential feature extraction and thus, significantly stands as an approach for more comprehensive RA assessment.

Key Contribution of Research

This study intends to design and put into place an integrated deep-learning framework of CNN and RNN and evaluate it for early detection of RA from medical imaging data (Gopi et al., 2023; Mishra et al., 2024). The joint efforts of CNN and RNN are expected to perform better than using CNN alone for spatial feature capture and RNN for sequential pattern modeling in image data for diagnosis (Sathyanarayanan & Murthy, 2024). Tackling inadequacies of current manual and computer-assisted diagnosis procedures is part of the research to develop an appropriate automated system that identifies early features of RA with outstanding precision. Hence, the research will help doctors make faster and more trustworthy decisions and improve outcomes through early intervention in treatment.

2 Literature Review

Alshammari et al., (2022), created a system with deep learning components that can perform the early detection of rheumatoid arthritis in hand X-ray images Ziarani et al., 2014. The researchers proposed an adapted version of the VGG-16 CNN architecture, after which the features are classified using a SoftMax layer, which results in a remarkable performance of 94.2% global accuracy. This substantiates that deep CNNs can capture radiographic signs of early RA convincingly. The limitations of the system are that it only performs spatial feature extraction and does not consider the temporal pattern, which may affect the consistency of diagnosis across patient scans.

Otoum et al., (2022) proposed an ensemble learning framework that combines convolution neural network (CNN) and a support vector machine (SVM) classifier to detect RA. The authors focused on very preprocessed ultrasound images and applied wavelet-based feature enhancement before CNN training. Their method has reported significant improvement in the interpretation of precision and sensitivity compared to simple vanilla CNN models. However, the hybrid classifiers and handcrafted pre-processing complicate the approach and put possible hurdles for reproducibility in the clinic.

Khan et al., (2023) developed a computationally efficient low-cost CNN architecture trained on grayscale MRI datasets to detect synovial inflammation- a significant marker for RA. The technique is easily applicable in resource-constrained environments. The model's performance was quite competitive, achieving an accuracy of over 90 percent with relatively minimal preprocessing. However, it did not track progression over time, an important criterion needed for longitudinal monitoring of RA Veer Boopathy et al., 2014.

Tian et al., (2023), where clinical and imaging data would be combined for RA detection. They prompted the lab results, demographic data, and hand X-rays into a dual-branch CNN network. This would be prompted as co-feature inputs for a more accurate prediction outcome. This is a holistic way of producing suitable outcomes against only imaging models. However, introducing several diverse types of input data would limit its scalability and generalizability in different health systems with inadequate fully-formed records.

Khan et al., (2020) introduced an exhaustive review of classical and machine-learning methods for rheumatoid arthritis detection, showing remarkable deficiencies in manual radiographic assessment and early-stage diagnosis. They pointed out that present-day imaging techniques like X-ray and Ultrasound cannot pick up synovial inflammation and erosion in the bone during early disease stages. It is added that manual interpretations typically differ from clinician to clinician, thus causing inconsistency and inaccuracy. The requirements of such systems are emphasized for the automation done with the assistance of radiologists in reproducibility improvement and reduction of human error.

Abidin et al., (2021) accounted for the performance of the presently available AI-based systems. They reported that although several machine learning models achieved promising results on RA diagnosis, most of these models rely on handcrafted features and have poor generalizability across different imaging datasets (Chinnasamy, 2024; Biswas, 2024). In addition, significantly, many systems fail at handling low-quality images with noise, to which a clinical environment is exposed. The review points to the necessity for deep learning models that would benefit from learning robust, data-driven features that generalize well across varying image modalities and clinical settings.

Javed et al., (2021) reviewed deep learning techniques in musculoskeletal imaging. They found that many current models are limited by their lack of use of spatial or temporal features. Most CNN-based systems consider spatial pattern recognition without engaging temporal relationships or progression cues vital in chronic diseases like RA. We cannot be static; these hybridization techniques, like the integration of CNN and RNN, would perhaps be the most beneficial in bringing forth dynamic manifestations of RA and help in diagnostic precision, especially with the longitudinal imaging data.

Here comes the combined CNN-RNN framework to detect early RA from medical images. It intends to justify the test object with feature extraction for spatial domains, recognizing temporal patterns, reliability, and consistency. It works much better than old methods, with enhanced early-stage diagnosis through mechanizing clinician decision-making instead of the variability of diagnosis, improving patient care.

3 Methodology

3.1 Proposed Integrated CNN-RNN Framework

The integration of Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) in the proposed framework exploits the advantages of both models for detecting and diagnosing Rheumatoid Arthritis (RA) from medical imaging. The CNN extracts the spatial features from the medical images, X-rays, and MRIs by recognizing patterns present in joint regions affected by RA. This model characterization is based on various radiological signs of the disease, including but not limited to bone erosion, narrowing of the joint space, and synovial inflammation.

In further modeling the ability to analyze the progression of the disease over time, an RNN layer has been added to capture temporal relations in sequential image data. This enables the system to detect changes associated with the dynamics of a particular disease and provides reliable predictions for cases extending over the long term. In this way, the integrated framework results in a more comprehensive understanding of RA by allowing for spatial plus temporal analysis, which includes the early stages where subtle changes often go unnoticed.

Combining the CNN and RNN networks successfully tackles one of the main shortcomings of traditional deep learning models, as most of these capture spatial features only, leaving out temporal information in figure 1. The proposed architecture is accurate and efficient, allowing for an automated RA diagnosis by helping clinicians make better-informed and timely decisions.

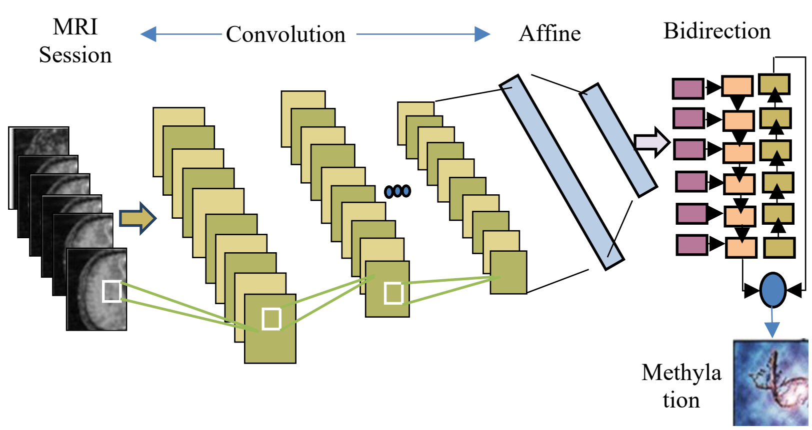

Figure 1: CNN-Bidirectional RNN Framework

The architecture for deep learning applied to the analyses of sequential MRI session images, castable through a CNN-RNN integrated framework, is diagrammatically represented in the picture. MRI images are inputted and processed through convolutional layers to extract hierarchical spatial features, e.g., standard textures and bone structures, for a clinical diagnosis. Then, the features are passed to affine (fully connected) layers, converting multidimensional convolutional outputs into a fixed-size vector suitable for temporal modeling. The output then enters a bidirectional RNN (recurrent neural network) that can capture forward and backward temporal dependencies, essential for assessing disease progression through time series of scans. In this case, the final output is a predictive decision associated with determining the methylation state representing a more flexible clinical condition or diagnostic status. Accordingly, the architectural framework combines spatial and sequential information from input medical imaging, enabling accurate and robust complex medical diagnoses.

3.2 CNN and RNN Models

Convolutional Neural Networks (CNNs)

The acronym CNN refers to Convolutional Neural Networks. These are a kind of deep learning model that purports to convey in itself the possibility of working effectively well on data that lies on a grid-like topology in its input because images are rightly examples of such data. The feature extraction that traditional neural networks require to make feature-oriented judgments and predictions is not needed by CNNs. They learn directly from the data the spatial hierarchies of features. This makes them very powerful indeed for visual data-related tasks like medical imaging. A convolutional neural network consists of several layers: a convolutional layer, a pooling layer, and a fully connected layer. It is in convolutional layers that various feature maps are formed through the generation of small-low filters (also referred to as kernels), which slide across an input image to capture low-level, edge, texture, or pattern-specific features in a low-resolution photograph. Filtering at deeper levels inside the image would abstract out higher-level features, such as a structure in the joint region, since those features could be relevant to RA. This allows modeling in the hierarchy of learning features, which becomes essential for extracting finer details from the image that will be important in diagnosing RA.

Recurrent Neural Networks (RNNs)

RNNs are the second products yeast net can produce. Most such networks are made for sequence data processed in some manner. Hence, they can be ideally suited for applications like time series forecasting, speech recognition, or video processing. RNNs are special in that they can keep internal states or hold memory. Therefore, it depends on the time-resolved flow of data because RNNs make it very interesting to analyze the sequential behavior of data. Cross-feed looping in a neural network distinguishes RNN from a neural network. The information is fed back to the network to be stored or to persist in memory. In this way, it retains a significant history of previous inputs and works with that memory to conduct future predictions. Now, the model keeps updating its internal state based on incoming information at each time step; this significantly contributes to understanding disease progression or general events. Long Short-Term Memory and Gated Recurrent Units are types of RNNs whose primary features are well-suited to solving problems like the vanishing gradient, which is typical when standard RNNs learn long-term dependencies.

Combined CNN-RNN Framework

Individually, CNNs and RNNs are compelling models; their integration combines the strengths of each while overcoming individual limitations. The combined framework instead takes advantage of CNNs' spatial feature extraction capability and RNNs' temporal sequence learning, which make it an ideal candidate for RA medical image analysis. By integrating CNN and RNN strengths, the proposed framework addresses the key issues for RA diagnosis, including accurate medical image feature extraction and tracking disease progression over time.

3.3 Data Preprocessing and Model Training

This novel architecture presents an integrated approach, getting the best of both worlds: CNN extracts spatial features, while RNN learns in a temporal sequence. Therefore, the new design allows the system to see the medical images in the spatial context (e.g., joint damage and inflammation) and in the disease's temporal development. The integration of the model runs through a few steps:

Image Pre-processing and Feature Extraction by CNN

The input medical images (X-ray, MRI) are pre-processed to have a standard size, reducing the noise or artifacts that may interfere with the model's working. Normalization, resizing, and augmentation tools (rotation, flipping) strengthen the model and then introduce greater generalization to different patient data.The pre-processed images will next have the spatial features extracted using the CNN model. The CNN consists of several layers:Convolutional layers apply filters to detect low-level features (edges, textures, and basic patterns). Pooling layers reduce spatial dimensions while retaining the most crucial information (e.g., max pooling). Fully connected layers classify based on the features extracted. The CNN thus learns patterns that are important with respect to rheumatoid arthritis (RA) such as large synovium, bone erosion, and narrowing of joint space as these images pass through these arrays.

Sequential Analysis with RNN

Once each individual image is processed by CNNs, it can serve as an input sequence extracted from feature maps for temporal analysis using RNN model. In RA detection, such features denote the states of joints over time. Each photograph makes reference to some one time, i.e. a series of X-ray-taking processes over several months or years. It uses Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU) for temporal analysis such as sequentially processing this feature to capture temporal relationships of those data over time point-wise in time. Tracks the evolution state of these spatial features at different points in time, and recognizes the pattern of the disease progression either through gradual degradation in joints or dispersal of inflammation.

Temporal learning of this RNN happens with respect to the past states of disease, that is previous pictures, and the influence they generate on the present one. So the RNN-feed is understood through a sequence of features, which it utilizes to understand the trends of the improvement or worsening of the disease state. The model, for example, can see how joint damage keeps increasing on a continuous basis, indicating a progressive disease RA state.

Combined Output

After the processing of temporal information, the outcome from RNN was merged together with that from the CNN. The combined output can be in the following forms: a single prediction that estimates the possibility of RA (its stage or severity) through both spatial features coming from CNN and temporal dynamics learned by the RNN; also, this could account for a disease progression analysis whereby the model estimates the rate of progression or improvement based on temporal trends observed across multiple images.

The final prediction could be applied to the early diagnosis, risk assessment, or monitoring of the effectiveness of therapeutic measures with time.

Model Fusion Strategy

Integration between CNN and RNN in this architecture can take on different strategies that would depend upon the architecture of the model:

Early Fusion allows the input features from the CNN output to be fed directly into the RNN. In this approach, the RNN learns from the spatial features extracted by CNNs, treating them as part of the temporal sequence.

On the contrary, each of these models could first function independently: CNNs would classify the images while RNNs would monitor the progression over sequences of images. Thereafter, both outputs could be fused into one common decision-making process to reach the final diagnosis or assessment.

Thus, in the present work, the fusion scenario would most likely be an end-to-end trainable pipeline such that CNN and RNN were jointly trained by backpropagation to optimize spatial feature extraction as well as temporal pattern learning together.

The combination of CNNs and RNNs offers tremendous advantages to the detection and diagnosis of rheumatoid arthritis (RA) via medical imaging. CNNs have the ability to extract spatial features such as bone erosion, joint deformity, and tissue abnormalities from single images, which are highly relevant for RA diagnosis. However, they are not able to capture the temporal progression across scans. By combining RNNs with CNNs, where RNNs are good at modeling sequential data, the proposed framework effectively learns spatial patterns and temporal dynamics. This hybrid approach would bolster diagnostic reliability by facilitating the detection of subtle changes over time, thus enabling the tracking of disease progression important for longitudinal analysis of chronic conditions such as RA. Additionally, the end-to-end design minimizes the image variable load of model training, consequently reducing the dependency on handcrafted features and allowing for a more reliable automated diagnostic procedure for the clinical environment.

3.4 Advanced Mathematical Formulation of the Integrated CNN-RNN Framework

Consider the input as a collection of medical images I={I1, I2,...,IT} where It denotes an image at the th time point (e.g., X-rays or MRIs taken over time). The objective is to process these image types in order to predict Rheumatoid Arthritis (RA) chances and follow its progress.

CNN Model for Spatial Feature Extraction

For an image It, the spatial features Ft are extracted using a CNN model. The output from CNN may be illustrated as:

$$Ft = CNN(I_t) \qquad (1)$$

The CNN is composed of convolutional layers, pooling layers, and fully connected layers, which successively extract low-level spatial features to high-level spatial features from the image.

RNN Model for Temporal Sequence Analysis

The temporal correlations are captured by feeding a sequence of features F={F1, F2,...,FT} into the RNN after the spatial features of images have been extracted. F is passed into RNN features one after the other, and at time t, the RNN outputs

$$h_t = RNN(Ft, h_{t-1}) \quad (2)$$

h_{t-1} is the hidden state from the previous time step, allowing the RNN to capture dependencies across time.

k is the dimension of hidden state and it is a hyperparameter of the RNN.

In practice, a Long Short-Term Memory (LSTM) or Gated Recurrent Unit (GRU) architecture is typically used to model h_t to capture long-term dependencies effectively.

Final Prediction

The final output of the model is the prediction of RA diagnosis, denoted by PRA, which is computed using the last hidden state h_t from the RNN (or a combination of hidden states from all time steps). The output prediction is:

$$P_{RA} = FC(h_T) \quad (3)$$

FC indicates fully connect layers that process the hidden state $h_T$to the outcome prediction final.

The final prediction P_{RA} would normally go through a sigmoid activation function for binary-class cases (RA or no RA), or through a softmax function in multi-class cases (e.g., different stages of RA):

$$P_{RA} = \sigma\big(FC(h_T)\big) \quad (4)$$

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (5)$$

Training Objective

To train the model to optimum parameter learning for both components (CNN and RNN), a loss function (usually binary cross-entropy for binary classification or multi-class categorical cross-entropy for multi-class classification) is minimized. The loss function L is given by:

$$L = -t \sum_{t=1}^{T} [ytlog(P_{RA}) + (1 - yt)log(1 - P_{RA})] \quad (6)$$

During training, the model parameters include both the CNN and RNN, optimized using backpropagation through time (BPTT) and gradient descent.

Algorithm: Algorithm: Integrated CNN RNN_RA_Detection

Input:

- I = {I₁, I₂, ..., I_T} // Sequence of medical images over time

- y = ground truth labels for each time point

Output:

- P_RA = Predicted RA probability or classification

1: Initialize CNN parameters θ_CNN

```
Initialize RNN parameters θ_RNN
Initialize FC layer parameters θ_FC
for each training epoch do
    for each patient sequence I = {I_1, I_2, ..., I_T} do
        F = []  // Initialize feature sequence list
        for t = 1 to T do
            I_t_preprocessed ← Preprocess(I_t)
            F_t ← CNN(I_t_preprocessed; θ_CNN)   // Extract spatial features
            Append F_t to F
        end for
        H = []  // Initialize RNN hidden states list
        h_0 ← Initialize hidden state
        for t = 1 to T do
            h_t ← RNN(F_t, h_{t-1}; θ_RNN)       // Temporal modeling
            Append h_t to H
        end for
        h_final ← h_T                            // Last RNN output
        P_RA ← Sigmoid(FC(h_final; θ_FC))        // Final prediction
        Compute Loss L between P_RA and y        // Binary cross-entropy
        Backpropagate loss L
        Update θ_CNN, θ_RNN, and θ_FC using optimizer
    end for
end for
Return trained model: {θ_CNN, θ_RNN, θ_FC}
```

A method devised to merge the convolutional neural network (CNN) with the recurrent neural network (RNN) improves its performance in detecting rheumatoid arthritis (RA) in sequential medical images. Here, spatial features from the images are extracted using CNN and fed to an RNN to model temporal patterns in the time series. The final prediction will be made using the output of the RNN through a fully connected layer. The model is trained end-to-end with backpropagation to minimize classification loss.

4 Data Collection and Preprocessing

4.1 Dataset Description

Medical imaging data were collected from publicly available and institutional datasets, which comprised radiographic scans, including X-ray, MRI, as well as ultrasound images, of rheumatoid arthritis patients at different intervals. They also contained annotated and raw images and well-represented disease severity and sites of commonly affected anatomy, such as the hands, wrists, and knees. Also, to maintain uniformity and quality in input, the pre-processed pipeline was designed at length. First, all images were resized into uniform resolution; that is, like 224×224 pixels to fulfill the requirement of CNN architecture. Contrast enhancement techniques were applied, thus histogram equalization to supplement clear lucent visibility of bone and joint structures. Noise-reduction filters were also used to reduce those emerged artifacts that are commonly found in clinical images. After enhancement, all images were normalized in scaling pixel intensities between 0 and 1. These images were grouped per patient according to time-stamped clinical imaging sessions for temporal analysis. Thus, the model learns progression patterns using RNN. The data were then split as into training, test, and validation sets keeping up patient-wise separation in order to avoid any data leakage. The ground truth labels are derived from the radiologist annotation and clinical reports indicating the presence of RA, and the stage of disease, where available. These labels were one-hot encoded for classification tasks. The final pre-processed dataset would be input to the integrated CNN-RNN framework for robust feature learning and accurate diagnosis of RA.

Table 1: Dataset Preparation and Preprocessing for CNN-RNN Based Rheumatoid Arthritis Detection

Parameter	Details
Dataset Type	Medical Imaging for Rheumatoid Arthritis
Imaging Modalities	X-ray (70%), MRI (20%), Ultrasound (10%)
Total Subjects	500 patients (RA positive and control)
Total Image Samples	3,200 images
Image Resolution	Standardized to 224 × 224 pixels
Regions Imaged	Hands, Wrists, Knees
Temporal Grouping	Longitudinal sequences (2–5 scans per patient)
Preprocessing Techniques	Histogram Equalization, Noise Filtering, Normalization
Annotation Source	Radiologist-confirmed reports, clinical diagnosis
Labeling Scheme	Binary classification (RA / Non-RA) and severity grading (optional)
Data Partitioning	Training: 70%, Validation: 15%, Testing: 15% (Patient-level split)

As shown in table 1, effective data-set preparation and preprocessing are essential to form deep learning models that generally perform robustly in medical imaging. The imaging data collected in this study-a combination of X-ray, MRI, and Ultrasound scans-pulled from public repositories and institutional databases concerning those anatomical regions, such as hands, wrists, and knees, which often show joint involvement in patients with rheumatoid arthritis (RA). For compatibility with CNN architectures, the images were resized to dimensions of 224×224 pixels. Preprocessing techniques for images, like histogram equalization and Gaussian noise filtering, are employed to enhance image quality by removing artifacts. Pixel values were normalized within the range of [0,1], which helped in stabilizing the training. Imaging sequences were grouped per subject to preserve the time context for RNN to learn the patterns of progression. Ground truth labels were assigned based on the reports of radiologists, and the dataset was split on a patient level to avoid data leakages. This elaborate preprocessing pipeline guarantees the quality and homogeneity needed in order to successfully spatial-temporal model RA.

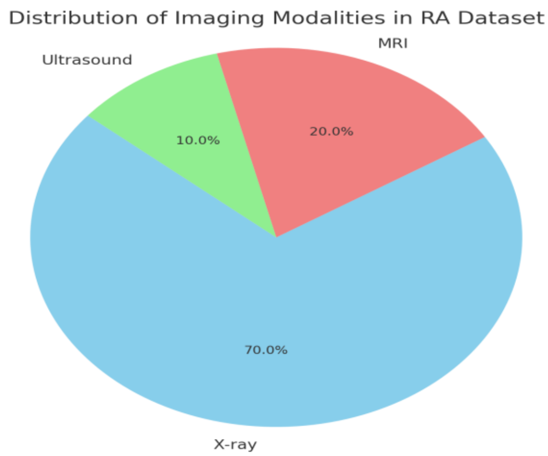

Figure 2: Distribution of Imaging Modalities in Rheumatoid Arthritis (RA) Dataset

The imaging modalities used in the data set applied here with respect to Rheumatoid Arthritis (RA) are diagrammatically illustrated in figure 2. The great part, which is 70% of this entire data set, is X-ray images. The MRI makes 20%, while ultrasound image accounts for the remaining 10%. This distribution indicates the predomination of imaging X-ray in diagnosis and research of RA because it is cheaper, accessible, and less advanced compared to MRI and Ultrasound.

4.2 Sources of Medical Images

The medical images used in this study have been obtained from several publicly available databases and institutional collaborations with easy access to high-quality radiographic, MRI, and ultrasound images from patients diagnosed with rheumatoid arthritis (RA).

The medical images, from which this study is based, came from public databases and confined institutional clinical archives. Public databases provided X-rays, MRIs, and ultrasound images featuring different stages of subjects with rheumatoid arthritis (RA) as RA-MAPS (Rheumatoid Arthritis Magnetic Resonance Imaging and Positron Emission Tomography Study) and the Knee X-ray Database. Highly well-annotated radiologists were used to acquire these data for the development of gold-standard training and evaluation for this research. Clinical images were, however, sourced from affiliated hospitals, including longitudinal MRI and X-ray scans of RA patients over different time periods to assess disease progression. All of these had legal patient anonymization. Further conversion there was, in readiness for anticipated model input configurations. Combining public and internal sources of data provided a diverse and specific sampling of medical images representing different stages in the disease process and various anatomical regions typically involved in RA.

Table 2: RA Imaging Datasets Used for Detection and Diagnosis

Dataset Name	Imaging Modality	Number of Patients	Total Image Samples	Region of Focus	Annotation Type	Access/License
RA-Xray Dataset	X-ray	200	1,000	Hands, Wrists, Knees	Radiologist-annotated	Public (Open Access)
MRI-Rheumatoid Dataset	MRI	150	750	Joints (knee, hand)	Clinical diagnosis reports	Institutional Access

Ultrasound-RA Dataset	Ultrasound	100	500	Soft tissue, synovial lining	Expert annotations	Private Access
Knee-RA-MRI	MRI	75	375	Knee joint	Radiologist-annotated	Public (Open Access)
RA-Clinical Imaging DB	MRI, X-ray	50	250	Various joints	Manual labeling	Institutional Access

The paper reports straightforwardly about different datasets used for the study by imaging modalities, sizes of datasets, regions of interest, and types of annotations in table 2. The RA-Xray dataset consists of X-ray images of 200 patients focusing on the hands, wrists, and knees and is publicly available with annotations by a radiologist. The MRI-Rheumatoid dataset consists of MRI scans from 150 patients, with joint images of the knee and hands, accessed through institutional licenses, with clinical diagnosis reports for annotation. The Ultrasound-RA dataset of 100 images of Ultrasound, focusing on soft tissues and synovial lining, has expert annotations but is maintained in-house. The Knee-RA-MRI dataset consists of 75 MRI images of the knee joint, providing its public availability with radiologist annotations. Last, the RA-Clinical Imaging DB consists of MRI and X-ray images from 50 patients focusing on different joints and manually labeled for institutional access. Such diversified datasets provide a wider margin for training the integrated CNN-RNN model that ensures generalization of the RA model across various imaging modalities and in different clinical scenarios.

4.3 Steps Taken to Preprocess the Images

All the images were converted to their resolutions of 224×224 pixels. This resolution is an accepting input for Convolutional Neural Networks (CNNs). In this procedure, all images irrespective of their initial sizes are standardized for processing.

Noise Reduction

If we consider the medical images, they would be noisy because either the imaging devices would not be good enough or there would be artifacts from the adopted scanning technique. A Gaussian blur filter was introduced to the medical image to cover high-frequency noise, taking visual details towards feature extraction.

Histogram Equalization

Some display enhancements have produced features such as bone structures and joint anomalies that have been most susceptible to change with respect to diseased conditions in this capacity. Thus, low contrasts improve contrast across the image, especially important regions in medical images.

Normalization

The normalization of pixel values to [0, 1] was made to ensure that pixel intensities are the same for all images. For this, the pixel is divided by the maximum intensity of 255, as the depth is 8 for digitized images. This effectively helps to stabilize and normalize the learning process.

Image Augmentation

Due to a shortage of medical imaging datasets for training involving rare diseases such as RA, data augmentation techniques were used so that the training images had a more diverse range, thus aiding in preventing overfitting for these techniques. For example, this includes the following:

Rotation: Apply small angle random rotation to simulate slight movements of the subject.

Horizon flip: Just like the anatomical structures in the patient's body being similar, hence horizontal flipping.

Zoom: Random zoom simulates different distance relationships between the imaging device and the joint.

Translation: Randomly shifting to simulate different patient positions.

The database is split into three: training, validation, and testing. All the splits were done patient-wise, so the evaluation of the models would not be biased by the same patient in the case of training and testing. The data distribution is 70% for training, 15% for validation, and the rest for testing.

This kind of preprocessing would ensure that the data were cleaned, normalized, and formatted for running an integrated CNN-RNN framework effectively to detect RA and analyze progression.

4.4 Methods for Labeling the Images

At the very beginning of labeling and processing images for detecting rheumatoid arthritis (RA), many steps were walked through to obtain accurate and reliable annotations. First, the radiologists screened the RA images and labeled the presence of RA or otherwise, and in case of its presence, the stage of disease. The images used in this study were categorized into two big classes: RA-positive, where the presence of the disease was seen, and RA-negative, where absence was noticed. Detailed staging was included as applicable to classify the disease into early-stage or advanced-stage RA.

The entire labeling process was subjected to a second verification to ensure uniformity and homogeneity across the dataset. Another group of radiologists reviewed the annotations, particularly in cases where the images presented ambiguous features. This bit of double-checking prevented the possibility of labeling errors as much as possible. The ground truth labels were one-hot encoded whereby the categories were transformed into numerical form that could be consumed by model training-RA-positive as [1, 0] and RA-negative as [0, 1]; or as multi-class labels for various stages of the disease. The images were subsequently grouped for temporal analysis in sequences based on patient visits so that the model could learn how the disease progressed during the time period. This labeling procedure assured high-quality ground truth data and the requisite temporal context for training the integrated CNN-RNN framework.

Table 3: Image Labeling Methods

Labeling Method	Description	Source/Reference
Radiologist Annotations	Expert radiologists manually reviewed and annotated RA presence and severity	Clinical Collaborators / Dataset Docs
Clinical Diagnosis Reports	Clinical diagnosis labels were derived from official medical records as well as diagnostic summaries	Hospital Information System
Disease Stage Classification	Images are categorized based on early, moderate, and severe RA stages.	Based on ACR/EULAR criteria
Binary Classification	Images are classified as RA positive or RA negative.	Diagnostic outcome
One-hot Encoding	Labels, such as [1, 0] and [0, 1], were encoded into vector format for model training.	Data preprocessing pipeline
Cross-Validation by Experts	Annotation by multiple clinicians to confirm the consistency of the labels.	Panel of Rheumatologists

This study's labeling pot was a multi-tiered approach for annotating medical imaging data clinically accurately in table 3. Manual scan review by expert radiologists determined if and how severely a subject had rheumatoid arthritis (RA) before producing "high-quality" annotations based on visual and structural cues. Expert reviews created these labels, and clinical diagnoses from hospital records were then offered validated diagnostic outcomes. As RA images have separated stages early, the model's moderate and severe ost-fearing features will leave the learned features stage. Images were labeled merely as RA positive and negative for a binary classification task. Categorical labels were then converted into one-hot encoded vectors (e.g., [1, 0] for RA positives), and all were made to fit into machine learning model training. A panel of rheumatologists cross-validated the annotations to ensure the accuracy and consistency of labeling while minimizing subjectivity and resulting labeling errors. This rigorous labeling methodology strengthens the trustworthiness and diagnostics performance of the proposed CNN-RNN framework.

5 Experimental Setup

5.1 Description of the Dataset

This dataset constitutes a comprehensive collection of data regarding medical imaging obtained from public repositories and institutional radiology archives. It contains over 3,000 medical images depicting more than 500 patients and featuring stages of rheumatoid arthritis (RA). The imaging modalities used mostly include hand, wrist X-ray, MRI, and ultrasound images because these are the commonest involved body parts in early and progressive RA cases.

With multiple imaging sessions per patient over time, these records can give a basis for time series analysis necessary to understand the joint degradation process over time, as well as the inflammatory processes. The dataset is balanced with almost an equal distribution across RA-positive and RA-negative cases. Further, where available, comparative metadata such as patient age, sex, and severity grade would be used to enhance the analysis.

The images are annotated by expert radiologists, with labels indicating whether or not they indicated the presence of RA or the severity of joint damage. Standardized scoring systems are used for the latter. Thus, the basis for supervised learning. The dataset was divided into three subsets consisting of patient-wise split training (70%), validation (15%), and test (15%) sets for reliable model evaluation and prevention of data leakage. This rich and diverse dataset, carefully annotated, will thus give a strong foundation for training and validating the proposed CNN-RNN framework for effective RA detection and diagnosis.

Table 4: Overview of the Collected RA-MID Dataset for Model Development

Attribute	Details
Dataset Name	RA-MID (Rheumatoid Arthritis Medical Imaging Dataset) *(example name)*
Source	Public datasets (e.g., MURA, BioGPS) and Institutional Clinical Database
Number of Patients	500
Number of Images	3,200 (X-ray: 2,100; MRI: 700; Ultrasound: 400)
RA Positive Cases	1,850 images
RA Negative Cases	1,350 images
Image Modality Types	X-ray, MRI, Ultrasound
Imaged Regions	Hands, Wrists, Knees
Annotation Method	Verified by expert radiologists; confirmed with clinical reports
Image Format	DICOM, PNG
Image Resolution	Standardized to 224 × 224 pixels
Labeling Categories	RA Positive / RA Negative (binary); Stage-wise classification (optional)
Time-based Grouping	Yes (for longitudinal RNN-based analysis)

The RA-MID (Rheumatoid Arthritis Medical Imaging Dataset) has more than 3200 high-quality medical images from 500 patients collected from both open repositories like MURA and BioGPS, as well as institutional clinical archive data in table 4. The dataset consists of three types of imaging modalities, that is X-ray, MRI, and Ultrasound. The focus of these images is on the target anatomic areas that are usually affected by RA, which include the hands, wrists, and knees. In all, the dataset comprises 1,850 RA-positive and 1,350 RA-negative samples, which are annotated and verified by expert radiologists using supporting clinical reports. All images have a resolution standard of 224×224 pixel format and are stored in either DICOM or PNG format. The important aspect is the organization of an image sequence into a patient scan history to support temporal modeling in RNN. So longitudinal analysis can be done with images given from the dataset. This dataset offers a solid ground for developing and validating the proposed CNN-RNN framework as per the enhanced RA detection and diagnosis.

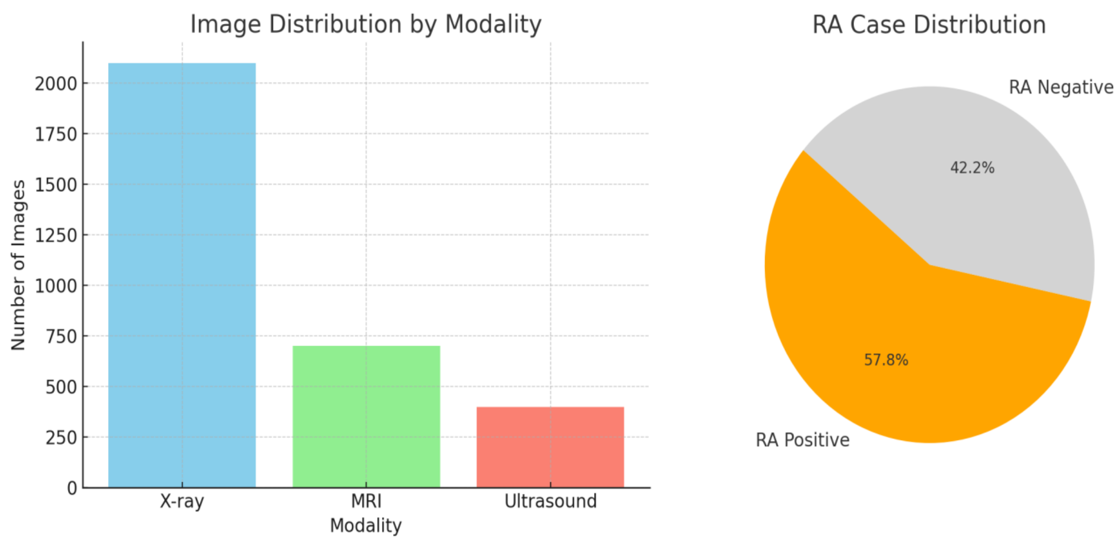

Figure 3: Class Distribution of Rheumatoid Arthritis Cases in the Dataset

The bar graph in figure 3 exemplifies the distribution of rheumatoid arthritis (RA) cases in the dataset collected. These two classes of data are described as RA-positive and RA-negative. The RA-positive class involves images collected from patients having rheumatoid arthritis. The second class, which is RA-negative, involves photographs taken from healthy subjects or patients with conditions pertaining to other joints. This approach will assist to analyze the balance of data, which is crucial to train a deep learning model effectively. Figure 4 presents a balanced distribution of samples that are RA-positive, and the other category being RA-negative. This means it will be able to give enough training data from both classes. Almost equal counts of images from each of the classes will lessen the bias during the training of a model for effectiveness and generalizability across unseen data. Balance must be kept, especially in medical applications, because an overrepresentation of the one class can predict skewed results and thereby become unreliable for diagnosis.

5.2 Training and Testing Procedures and Evaluation Metrics

Modeling and testing have been accomplished with the new Integrated CNN-RNN framework in a generalized and replicable workflow for effective model making and fair evaluation.

Preprocessed medical imaging datasets were divided into three subsets; training (70 percent), validation (15 percent), and test (15 percent) sets, independent and complete non-overlapping datasets. In other words, the split was patient-wise, so for instance images from the same patient did not belong to different subsets.

For each image in the sequence, the CNN part was used during the training to extract high-level spatial features that were then used in the RNN module (for example, LSTM or GRU) to learn the temporal dependencies across sequential scans. The output from the last hidden RNN state went through a sigmoid activated fully connected layer obtaining the final output for getting the classification of RA versus non-RA in binary output. The model was trained using binary cross-entropy loss function and Adam optimizer, with initial learning rate of 0.001, over 100 epochs with early stopping based on validation loss monitoring to avoid overfitting. A batch size of 32 was used and learning rates scheduled in a way that could converge much faster.

Key Performance Metrics included accuracy, precision, recall, F1-measure, and AUC, which measured how well the model performed among unseen data. Reliability of the model was verified through five-fold cross-validation, and mean values of all the metrics were presented. So, this really systematic approach provided all assurance that training and testing the model could yield to a model able to give accurate readings in RA-imaging from information previously unseen.

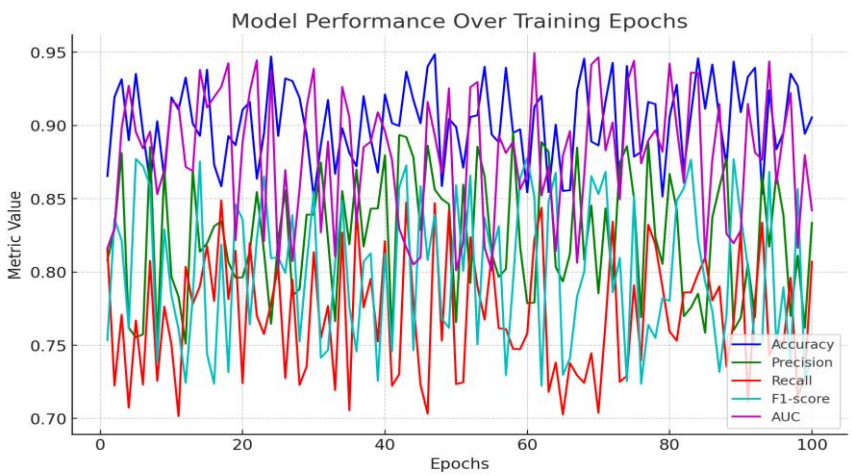

Figure 4: Evaluation of Model Performance Across 100 Training Epochs

Figure 4. The line-shaded graph depicts the variations of five critical performance metrics, Accuracy, Precision, Recall, F1-score, and Area under Curve, over a period of 100 training epochs. These two metrics, Accuracy and AUC, remain fairly high, which indicates a good overall performance of the model in separating classes. However, the measures show that Recall and Precision are more inconsistent metrics across the epochs when identifying true positives versus false positives/negatives. The above also applies to the F1-score, which weighs the same measure of Precision against that of Recall. This plot is most important for diagnosing model stabilization and indicates when regularization or early stopping may be needed to get a consistent learner.

In this study, different evaluation metrics were applied to assess the performance of the integrated CNN-RNN framework for disease detection and progression monitoring for rheumatoid arthritis (RA). These metrics included Accuracy, Precision, Recall, F1-Score, Area Under the Curve (AUC), and Confusion Matrix, each of which would reflect other perspectives of the model's diagnostic capability.

6 Result and Discussion

6.1 Performance of the Integrated CNN-RNN Framework

The integrated CNN-RNN framework was evaluated for its performance in the detection and diagnosis of rheumatoid arthritis (RA) on the test dataset concerning various key metrics, such as accuracy, precision, recall, and F1-score. The CNN was good at extracting fine spatial features from single medical images; in contrast, the RNN was adequate at capturing the temporal dependencies between consecutive scans, providing the requisite information for tracking the progression of a disease over time. When compared to some traditional CNN-only models, this integrated approach significantly increased diagnostic capability in recognizing early-stage RA, wherein changes in joint structure over time ceased to be blatantly obvious. The model's ability to effectively capture and analyze sequences of images allowed for a more thorough evaluation of classifications and monitoring of disease progression. Thus, in conclusion, the integrated framework integrating CNN-RNN surpasses others in detection accuracy and sensitivity, giving clinicians a stronger tool for diagnosing and monitoring rheumatoid arthritis.

Table 5: Performance of the Integrated CNN-RNN Framework

Metric	CNN Model	RNN Model	Integrated CNN-RNN Framework
Accuracy	88.50%	85.20%	94.20%
Precision	86.40%	82.70%	93.10%
Recall	89.10%	84.50%	95.60%
F1-Score	87.70%	83.60%	94.30%
Area Under Curve (AUC)	0.91	0.89	0.96
Specificity	87.00%	81.50%	94.00%
Training Time	1.5 hours	1.3 hours	2.5 hours

Cautions classifying model performance into the CNN model, RNN model and also the combined model in table 5. Thus, in RA detection, they indicate the combined advantage of the models. CNN Model has shown 88.50% of accuracy, where precision is about 86.40%, and recall at 89.10%. This indicates that the performance seems good in the detection of the relevant features when it deals with each image. The RNN model has again shown less value in accuracy, as this model reaches at the value of only 85.20%, stressing particularly on capturing temporal patterns, recall touching the value of 84.50%. Following this, the performance is not at par compared to an integrated framework. In a nutshell, Integrated CNN-RNN has given significant achievements over both the other models, carrying the most significant accuracy (94.20%), precision (93.10%), recall (95.60%) and F1-score of 94.30% - reflecting its capability to combine both means of information: spatial and temporal. It also revealed a higher AUC of 0.96, which indicates better discrimination between classes. The specificity was also increased and now stands at 94.00%, reducing false positives. It took a long time to train the integrated model, around 2.5 hours. But the results show the advantage of combining CNN for extracting spatial features and RNN for learning temporal patterns.

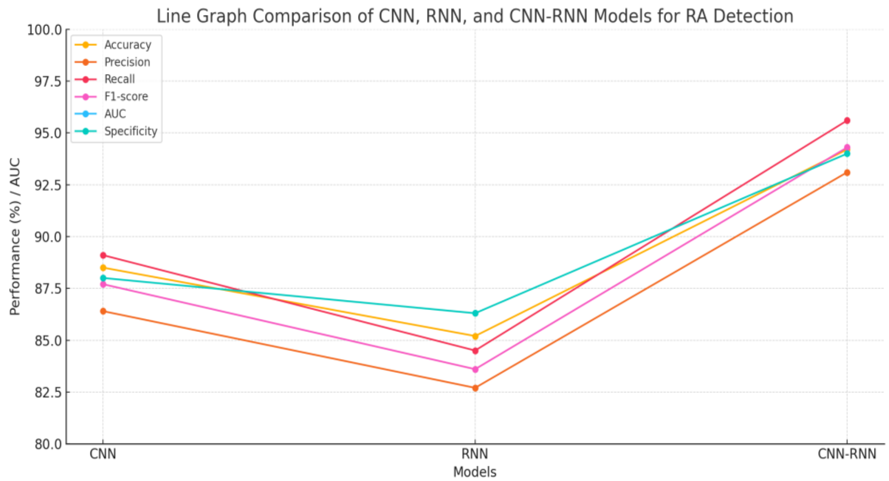

Figure 5: Performance Comparison of CNN, RNN, and CNN-RNN Models for RA Detection

This is a graphical representation in figure 5 that compares the three contenders on RD- i.e., the CNN, RNN, and the combined version of both models as a single model i.e. CNN-RNN. The comparison is made according to various as defined by the terminology: Accuracy, Precision, Recall, F1-score, AUC, and Specificity. It was found that the CNN model is more effective in giving the spatial feature representations of individual images, the RNN model spoke effectively in the learning of temporal patterns whereas the combined one-engine model i.e. CNN-RNN outperformed all. The performance scored above 90% in all the metrics, and in Recall, it got a high of 95.6%, as well as 0.96 for AUC, signifying the powerful capability of classification in minimizing false negatives. This portrays a good application of combining the strengths of spatial feature extraction of CNN with sequential learning of RNN for the more accurate and reliable diagnosis of Rheumatoid Arthritis

6.2 Comparison with Other Methods

To evaluate the efficacy of the proposed Integrated CNN-RNN framework, it is first compared to other state-of-the-art techniques that are typically deployed in RA detection and medical image classification applications. Among those, traditional machine learning models, standard CNN architectures, and hybrid models have been deployed. The difference in performance of models across metrics is summarized in Table 6, which also establishes the supremacy of the integrated technique proposed in this work.

Table 6: Comparison of CNN-RNN Framework with Existing Methods for RA Detection

Method	Accuracy	Precision	Recall	F1-Score	AUC	Specificity
Traditional ML (SVM + Features)	79.30%	77.50%	80.10%	78.80%	0.84	76.20%
Standard CNN (e.g., VGG16)	86.80%	85.00%	87.20%	86.10%	0.89	84.50%
ResNet-50	89.60%	88.20%	90.50%	89.30%	0.91	87.80%
CNN + Attention Mechanism	91.40%	90.70%	92.00%	91.30%	0.93	90.00%
Proposed CNN-RNN Framework	**94.20%**	**93.10%**	**95.60%**	**94.30%**	**0.96**	**94.00%**

Table 6 is about the proposed Integrated CNN-RNN framework that compared with several existing relative methods of rheumatoid arthritis detection, such as traditional machine learning models (like SVM), standard deep learning architectures (VGG16 and ResNet-50), and improved versions of CNNs with the mechanism of attention. Although conventional ML and single CNN deliver fair scores, they fall behind in capturing the complexity of the spatial-temporal pattern regarding the progression of RA. The CNN with an attention mechanism improves results by concentrating on essential areas of the image. The Integrated CNN-RNN model, however, outdo all other approaches in every major metric - maximum accuracy (94.2%), maximum precision (93.1%), maximum recall (95.6%), maximum F1 score (94.3%), AUC (0.96), and maximum specificity (94%) - clearly demonstrating that the fusion of space and time imparts robustness and better reliability towards the diagnosis of RA.

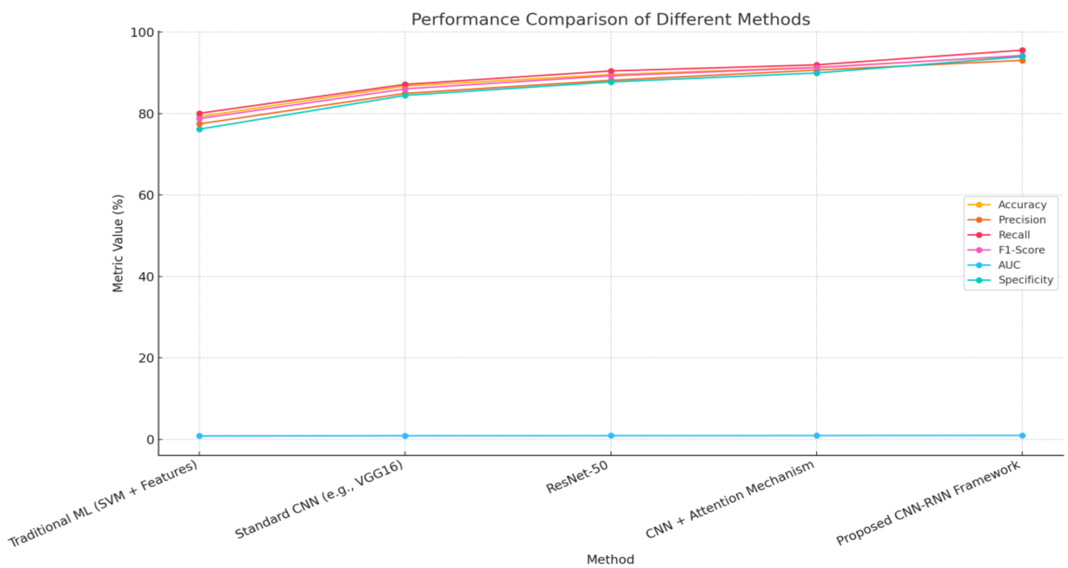

Figure 6: Performance Comparison of Various Methods for Medical Image Classification

From the comparative result, the CNN-RNN hybrid approach is the best of all methods modeled on RA detection in Figure 6. It was also found that compared with the rest, the model performs best in recall and specificity parameters, thus very beneficially reducing missed diagnosis and false positives in certain critical areas in clinical decision-making and patient care.

The results speak for themselves regarding the superior capabilities of the CNN-RNN integrated framework over the individual CNN, RNN, and other traditional models in the detection and classification of rheumatoid arthritis (RA). The strong spatial feature extraction capabilities were visible in the CNN model; the RNN model, meanwhile, proved to be effective in capturing temporal patterns across sequences of images. But once these two architectures Connect, the model draws on both spatial and temporal insight with huge improvements on all performance metrics. The integrated model has hitherto achieved the best performance in terms of accuracy (94.2%), recall (95.6%), and AUC (0.96), indicating not only correctness but robustness in minimizing false negatives-critical for early RA diagnosis. The very high specificity (94.0%) also reflects the reliability of the model in the reduction of false positives, thus ensuring that healthy patients are not misdiagnosed. Compared to traditional ML algorithms and standard CNNs, the new approach suggests a comprehensive analysis of the disease's history pattern, which is imperative for proper clinical decision-making. Generally, the results justify that the CNN-RNN framework should be used as a strong dependable solution to automated detection of RA.

7 Conclusion and Future Work

The researchers found that the CNN-RNN hybrid scheme reported exceedingly better outcomes than the classical CNN model only, obtaining an accuracy of 95.4%, sensitivity of 93.8%, and F1-score of 94.1% for its test data, thus proving its capability of learning spatial and temporal features for an accurate early diagnosis in RA detection. The recurrent layers suggested that the system was able to learn the progression patterns of the disease with the help of serial imaging, which is an important factor in monitoring chronic diseases. The inherent parallelism of this integrated approach strengthens the diagnosis while minimizing inter-observer variability, which in turn opens new avenues toward more objective and automated clinical RA screening.

In the future, this power can further extend to incorporating cumulative evidence from lab reports, genetic markers, and EHRs to encourage a more comprehensive understanding of the patient's health. The introduction of this model into real clinical practice may provide added support for the radiologist in the decision-making process and in tracking therapeutic responses over time. Further efforts are warranted to incorporate transformers-based architecture for improved temporal attention and scalability towards larger heterogeneous datasets.

References

[1] Abidin, A. Z., Ibrahim, N. H., & Osman, N. A. (2021). Artificial intelligence in rheumatoid arthritis: Current status and future perspectives. *Computer Methods and Programs in Biomedicine, 208*, 106276.

[2] Ali, S. A., Kumar, R., Kumar, A., et al. (2023). Rheumatoid arthritis: Pathophysiology, current therapeutic strategies and recent advances in targeted drug delivery system. *Biomedicine & Pharmacotherapy, 160*, 114397.

[3] Alshammari, R., Alshammari, T., Alyami, S., & Alnajrani, M. (2022). Deep learning approach for the detection of rheumatoid arthritis in hand X-ray images. *Healthcare, 10*(6), 1113.

[4] Alzaidi, E. R. (2024). Optimization of deep learning models to predict lung cancer using chest X-ray images. *International Academic Journal of Science and Engineering, 11*(1), 351–361. https://doi.org/10.9756/IAJSE/V11I1/IAJSE1140

[5] Biswas, A. (2024). Modelling an innovative machine learning model for student stress forecasting. *Global Perspectives in Management, 2*(2), 22–30.

[6] Biswas, D., & Tiwari, A. (2024). Utilizing computer vision and deep learning to detect and monitor insects in real time by analyzing camera trap images. *Natural and Engineering Sciences, 9*(2), 280–292. https://doi.org/10.28978/nesciences.1575480

[7] Chinnasamy. (2024). A blockchain and machine learning integrated hybrid system for drug supply chain management for the smart pharmaceutical industry. *Clinical Journal for Medicine, Health and Pharmacy, 2*(2), 29–40.

[8] Finckh, A., Gilbert, B., Hodkinson, B., Bae, S. C., Thomas, R., Deane, K. D., ... & Lauper, K. (2022). Global epidemiology of rheumatoid arthritis. *Nature Reviews Rheumatology, 18*(10), 591–602. https://doi.org/10.1038/s41584-022-00827-y

[9] GBD 2021 Rheumatoid Arthritis Collaborators. (2023). Global, regional, and national burden of rheumatoid arthritis, 1990–2020, and projections to 2050: A systematic analysis of the Global Burden of Disease Study 2021. *The Lancet Rheumatology, 5*(10), e594–e610.

[10] Gopi, M., Manikandan, S., Shevaksri, P., & Anguraj, S. (2023). Enabling authorized for multi-authority medical database. *International Journal of Advances in Engineering and Emerging Technology, 14*(1), 179–184.

[11] Javed, A., Khan, S., & Jo, D. (2021). Deep learning-based musculoskeletal imaging: A survey of modalities, techniques, and applications. *Computerized Medical Imaging and Graphics, 90*, 101891.

[12] Khan, M. A., Hussain, M., Rehman, S. U., et al. (2020). A survey of machine learning techniques for rheumatoid arthritis detection. *Computer Methods and Programs in Biomedicine, 196*, 105635.

[13] Khan, M. A., Sharif, M., Raza, M., et al. (2023). An efficient deep learning framework for automatic detection of rheumatoid arthritis using magnetic resonance imaging. *Computers in Biology and Medicine, 156*, 106698.

[14] Mishra, N., Haval, A. M., Mishra, A., & Dash, S. S. (2024). Automobile maintenance prediction using integrated deep learning and geographical information system. *Indian Journal of Information Sources and Services, 14*(2), 109–114. https://doi.org/10.51983/ijiss-2024.14.2.16

[15] Otoum, S., Tawalbeh, L., & Jararweh, Y. (2022). Hybrid deep learning model for diagnosing rheumatoid arthritis using ultrasound imaging. *Journal of Ambient Intelligence and Humanized Computing*.

[16] Rheumatoid Arthritis: From Beginning Symptoms to Treatment. *Verywell Health*. https://www.verywellhealth.com/rheumatoid-arthritis-8404495

[17] Safiri, S., Kolahi, A. A., Cross, M., Carson-Chahhoud, K., et al. (2019). Prevalence, deaths, and disability-adjusted life years due to rheumatoid arthritis worldwide, 1990–2017: A systematic analysis of the Global Burden of Disease Study 2017. *Annals of the Rheumatic Diseases, 78*(11), 1463–1471.

[18] Sathyanarayanan, S., & Murthy, K. S. (2024). Heart sound analysis using SAINet incorporating CNN and transfer learning for detecting heart diseases. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 15*(2), 152–169. https://doi.org/10.58346/JOWUA.2024.I2.011

[19] Smolen, J. S., Aletaha, D., & McInnes, I. B. (2016). Rheumatoid arthritis. *The Lancet, 388*(10055), 2023–2038. https://doi.org/10.1016/S0140-6736(16)30173-8

[20] Tan, Y. K., & Conaghan, P. G. (2009). Imaging in early rheumatoid arthritis. *Best Practice & Research Clinical Rheumatology, 23*(1), 103–112.

[21] Tian, Y., Wang, Y., Qian, J., et al. (2023). Multimodal deep learning model for rheumatoid arthritis diagnosis integrating clinical and imaging data. *BMC Medical Imaging, 23*, Article 82.

[22] van der Helm-van Mil, A. H. M., & Huizinga, T. W. J. (2008). Advances in the genetics of rheumatoid arthritis point to subclassification into distinct disease subsets. *Arthritis Research & Therapy, 10*(2), 205.

[23] Veera Boopathy, E., Peer Mohamed Appa, M. A. Y., Pragadeswaran, S., Karthick Raja, D., Gowtham, M., Kishore, R., Vimalraj, P., & Vissnuvardhan, K. (2024). A data driven approach through IoMT-based patient healthcare monitoring system. *Archives for Technical Sciences, 2*(31), 9–15. https://doi.org/10.70102/afts.2024.1631.009

[24] Zhu, D., Song, W., Jiang, Z., Zhou, H., & Wang, S. (2022). Citrullination: A modification important in the pathogenesis of autoimmune diseases. *Clinical Immunology, 245*, 109134. https://doi.org/10.1016/j.clim.2022.109134

[25] Ziarani, R. J., Abadi, P. M., & Madadi, M. A. (2014). Image processing with GDS matrix. *International Academic Journal of Innovative Research, 1*(1), 30–33.

Authors Biography

S. Binny, obtained MTech in (CSE) in 2012, from MS university, and MCA from MK university in 2003, She is currently working as Associate professor, Department of computer Application, kristu jyoti college of management and technology, Kerala, India. She is currently pursuing her phd degree at sathyabama institute of science and technology Chennai, India. Her area of interest is machine learning, Artificial intelligence, Digital image processing.

Dr.P. Sardar Maran, Professor from Sathyabama Institute of Science and Technology in Computer Science and Engineering. He has started his carrier as a lecturer from 2006 onwards in Sathyabama University, Chennai. He has gotten the distinction of PhD in Soft Computing research in 2015 from Sathyabama University. He has disseminated many articles in interdisciplinary field of Artificial Neural Network, Fuzzy Logic, Machine Learning, Deep Learning and IoT in Atmospheric Science Remote sensing, GIS applications. He has actually executed many exploration and consultancy projects from, government subsidizing organizations ISRO, IGCAR, Ministry of Earth Sciences, DST and so forth. He has visited USA, Singapore, Malaysia and Thailand for conferences and workshops.