Title: Artificial intelligence quantification of enhanced synovium throughout the entire hand in rheumatoid arthritis on dynamic contrast-enhanced MRI

Abstract

BACKGROUND: Challenges persist in achieving automatic and efficient inflammation quantification using dynamic contrast-enhanced (DCE) MRI in rheumatoid arthritis (RA) patients.

PURPOSE:

To investigate an automatic artificial intelligence (AI) approach and an optimized dynamic MRI protocol for quantifying disease activity in RA in whole hands while excluding arterial pixels.

STUDY TYPE: Retrospective.

Subjects: 12 RA patients underwent DCE-MRI with 27 phases for creating the AI model and tested on images with a variable number of phases from 35 RA patients.

FIELDSTRENGTH/SEQUENCE: 3.0T/dynamic contrast-enhanced T1-weighted gradient echo sequence (mDixon, water image).

ASSESSMENT: The model was trained with various DCE-MRI time-intensity number of phases. Evaluations were conducted for similarity between AI segmentation and manual outlining in 51 ROIs with synovitis. The relationship between synovial volume via AI segmentation with rheumatoid arthritis magnetic resonance imaging scoring (RAMRIS) across whole hands was then evaluated. The reference standard was determined by an experienced musculoskeletal radiologist.

STATISTICAL TEST: Area under the curve (AUC) of receiver operating characteristic (ROC), Dice and Spearman's rank correlation coefficients, and

interclass correlation coefficient (ICC). A p-value < 0.05 was considered statistically significant.

RESULTS: A minimum of 15 phases (acquisition time at least 2.5 minutes) was found to be necessary. AUC ranged from 0.941±0.009 to 0.965±0.009. The Dice score was 0.557 - 0.615. Spearman's correlation coefficients between the AI model and ground truth were 0.884 - 0.927 and 0.736 - 0.831, for joint ROIs and whole hands, respectively) . The Spearman's correlation coefficient for the additional test set between the model trained with 15 phases and RAMRIS was 0.768.

CONCLUSION: The AI-based classification model effectively identified synovitis pixels while excluding arteries. The optimal performance was achieved with at least 15 phases, providing a quantitative assessment of inflammatory activity in RA while minimizing acquisition time.

Keywords: rheumatoid arthritis, dynamic contrast-enhanced MRI, artificial intelligence, one-dimensional time-intensity data

Introduction

Rheumatoid arthritis (RA) is an immune-mediated disease characterized by inflammatory arthritis, marked by symmetric, polyarticular pain and swelling, predominantly affecting the small joints of the hands and feet (1). The prognosis of RA is linked to the disease stage at the time of diagnosis (2). A delayed diagnosis results in severe systemic involvement, leading to multi-tissue and multi-organ damage, increased disability, substantial socioeconomic consequences, and even mortality (2). On the other hand, early identification and intervention can prevent or delay disease progression in up to 90% of patients, averting irreversible joint damage and associated disability (3). Early recognition of arthritis and differentiation from other potential etiologies are therefore paramount in managing inflammatory arthritis (4). The subsequent crucial step in optimizing patient outcomes is effective treatment monitoring, relying on a comprehensive evaluation of clinical, laboratory, and imaging data (5). The need for more rapid and accurate tools for treatment monitoring of inflammatory arthritis has, therefore, been the topic of much recent research. (6–8)

One example of quantitative scoring of RA is the Rheumatoid Arthritis Magnetic Resonance Imaging Scoring (RAMRIS) system, developed by the Outcome Measures in Rheumatology (OMERACT) group (9,10). RAMRIS facilitates the quantification of bone erosion, bone marrow edema (BME), and synovitis on MRI (11,12). While RAMRIS has been well established, its responsiveness is hindered by the discreet nature of the score (13, 14). Moreover, a dispersion of the correlation coefficient

among the viewers between 0.45 and 0.98 has been reported (6–8, 15, 16). Quantitative measurements of the perfusion of arthritic joints with dynamic contrast-enhanced (DCE) MRI, allow assessment of synovial activity and predict future outcomes (17, 18). The uptake and wash-out characteristics of gadolinium-based contrast agents (GBCA) in different tissues can be visualized with pixel-by-pixel time-intensity curve (TIC) shapes derived from DCE-MRI. The analysis of pixel-by-pixel TIC shapes in DCE-MRI, particularly in the wrist and finger joints of RA patients, reveals seven distinct TIC shapes (19), with synovitis prominently belonging to curve type 4, characterized by rapid initial enhancement followed by an early washout phase (20).

Previous studies have suggested that TIC shape analysis provides a more accurate quantification of synovitis when compared to the conventional method of manually drawing the regions of interest (ROIs) of synovitis (21, 22). However, the challenge arises in distinguishing between synovitis and arteries, especially given the high similarity in their TIC patterns (20). A semi-quantitative TIC shape analysis method combined with an arterial mask subtraction approach was proposed to address this, demonstrating a high correlation coefficient (22). Yet, this method, reliant on manual extraction of parameters and features, has three notable limitations: an additional classifier to eliminate arterial influence; a time-consuming ROI delineation process unsuitable for clinical trials; and the objective nature of manually defined TIC shape features. Meanwhile, DCE-MRI is traditionally a time-intensive acquisition process

(23). Obtaining DCE-MRI data requires around seven minutes post-contrast administration, encompassing more than 20 phases (24). Both entail significant time expenditures.

Despite existing analysis methods for DCE-MRI predominantly relying on pharmacokinetics and TIC shape analysis, integrating artificial intelligence (AI), particularly deep-learning techniques (25, 26), may improve results. Unlike traditional methods, deep learning can autonomously learn intricate data patterns, ensuring robustness and efficiency in handling complex medical data (27).

Thus, the goal of this study was to investigate an automatic and efficient deep learning approach, along with an optimized dynamic MRI protocol with shorter acquisition time, for quantifying disease activity in RA in whole hands while excluding arterial pixels.

Methods and Materials

Participants

The Institutional Ethics Committee of our university approved this retrospective study, which required no informed consent. The study followed the ethical standards of the 1964 Declaration of Helsinki and its later amendments. To establish an AI model, this study included 12 participants diagnosed with wrist or finger joint RA meeting the ACR/EULAR criteria (28) for RA between 2018 and 2019. Inclusion criteria considered both inflammatory arthritis and extra-articular involvement on

DCE-MRI as positive indicators for RA assessed by a musculoskeletal (MSK) radiologist (T.K, with 30 years of experience). The cohort consisted of 2 men and 10 women, with a mean age of 53.41±13.34, ranging from 35 to 78 years. The AI model was tested on images with a variable number of phases from 35 RA patients (age 58.31±17.63, 24 females, 11 males) scanned between October 8, 2019, and December 19, 2023. All data utilized in the study were derived from these subjects.

MRI Data Acquisition

All subjects were examined using a 3-T MR system (Philips Medical Systems, Best, The Netherlands) with two types of coils. For imaging with both hands raised and palms facing down, the dStream HeadNeckSpine coil was used. For elderly patients or those unable to raise both hands, imaging was conducted in a supine position with hands placed on the abdomen using the dStream WholeBody coil. The acquisition of DCE-MRI included initial positioning images obtained with the following parameters: Two-Dimensional Fast Field Echo (2D-FFE) sequence acquired in three directions; Field of View (FOV) 450 x 450 mm; matrix size 256 × 128; slice thickness 10.0 mm; echo time (TE) 0.87 msec; repetition time (TR) 6.97 msec; number of slices 16.

A contrast agent of 0.2 mmol/kg (Magnescope; Guerbet, Paris, France) was administered using an automatic MR injection device (Sonic--Shot GX; Nemoto Kyorindo). The images were acquired using the 3D mDixon method with the following parameters: coronal orientation; TEs 1.35 / 2.4 msec; TR 6.00-6.18 msec;

slice thickness 2.0 mm; matrix size 192 × 192; flip angle 10°; scan time 5.30-6.22 minutes; and number of slices 55; number of phases 27.

Additionally, a test set (N=35) was also employed using the 3D mDixon method with parameters the same to the training set, except for variations in the TR 4.44-6.74 msec, scan time 3.96-5.37 minutes, number of slices 55-78, and number of phases 17-27. A bar chart depicting the distribution of phases in the additional test set (N=35) is shown in the supplementary material.

Training Data Acquisition

To acquire one-dimensional time-intensity data as a training set, a dedicated tool was developed using an in-house software program scripted in MATLAB 2021a (The Mathworks, Inc. , Natick, Mass, USA). This tool was the foundation for collecting comprehensive training data, focusing on synovitis and arterial pixels. A total of 2342 synovitis pixels and 2660 arterial pixels from 12 cases were gathered through a manual delineation process, which involved two researchers (K.I, with 7 years of experience, F.W, with 10 years of experience) creating 57 synovitis (4.75±5.77, min 0, max 17) and 59 arteries (4.92±1.73, min 2, max 8) freehand ROIs on DCE-MRI, shown in Figure 1. The delineation process was carried out using a computer mouse under the guidance of an MSK radiologist (T.K). The delineated ROIs are visually depicted in Figure 1. All data points were drawn on the 9th and 10th phase images, selected near the time-to-peak (TTP), because it's easier to observe the synovial and

arterial enhancement at this stage, making it simpler to outline them accurately (19).. This dataset was utilized during training and validation of the model development.

Dataset Partition

The 12 subjects' hand DCE-MRI data were divided into training and validation sets with a ratio of 11:1, utilizing the leave-one-out cross-validation (LOOCV) technique to facilitate model training and evaluation. For the training and validation sets, one-dimensional time-intensity data of synovitis and arteries were randomly split into a training set (80%) and a validation set (20%). The training set was utilized for model training, while the validation set was employed to monitor the performance of the model during the training process. Additionally, the original 27-phase DCE-MRI data were systematically reduced to 5 phases to optimize model efficiency, resulting in 12-phase sequence data sets with different lengths. In total, models were trained 144 times (12 subjects × 12 phase sequence lengths) before proceeding to testing. The dataset partition method is illustrated in Figure 2.

The cross-validation test set was employed to identify the minimum number of necessary phases and assess the synovial quantification performance. A two-step evaluation process was conducted involving joint ROIs and whole-hand testing. Joint ROI testing utilized Dice and Spearman's rank correlation coefficients to compare the AI model's synovial area segmentation against manually delineation ROIs. In whole-hand testing, Spearman's rank correlation coefficient between the predicted synovial volume via AI segmentation and the scores calculated using RAMRIS for each hand.

Ground Truth Labelling / Scoring

In training and validation, the accurate classification of synovitis and arterial pixels relied on the ground truth established during the data acquisition. For joint ROI testing, 51 synovitis ROIs were selected by an MSK radiologist (T.K) in the tenth phase from each of the 12 cases. ImageJ 1.53h (National Institutes of Health, Bethesda, Maryland, USA) was used to define the location and size of each ROI. The MSK radiologist (T.K) then manually outlined the synovitis areas within these ROIs, establishing the gold standard. This reference standard included a hand-drawn mask covering the joint and synovitis areas.

The MSK radiologist (T.K) also scored each overall hand using RAMRIS for each of the 12 cases, serving as the reference standard for Spearman's rank correlation coefficient analysis. For each hand, the score incorporated the distal radioulnar joint (DRUJ) (0-3), radiocarpal joint (RCJ) (0-3), and inter-carpal/carpometacarpal (IC/CM) (0-3) regions of the wrist, as well as the 5 metacarpophalangeal (MCP) joints (0-3 for each MCP joint), 5 proximal interphalangeal (PIP) joints (0-3 for each PIP joint), and tenosynovitis (0-3 for 6 extensor/ 3 flexor tendon compartments). By encompassing these specific areas, the RAMRIS objectively reflected the intensity of RA inflammation activity across whole hands.

The RAMRIS system underwent intra-observer reliability assessment through interclass correlation coefficient (ICC) (29), with the same MSK radiologist (T.K) scoring each hand twice (an interval of 9 months) for 12 cases used for AI modelling

to ensure consistency and agreement in the scores. The robustness of the RAMRIS results, serving as the gold standard, was determined by ICC values, thereby enhancing the credibility of results.

Deep Learning Network

Pre-processing

A series of image processing algorithms were implemented, focusing on denoising and image registration.

Initially, a median filter was applied to isolated noise from each DCE-MRI slice, while retaining edge information between adjacent tissues. Subsequently, the phase-only correlation algorithm (30) was utilized to compute the translation between the phase featuring the most pronounced synovitis characteristics and other phases. The average translation derived from this process facilitated the registration of the DCE-MRI images. Finally, all training and test data were normalized using zero-mean normalization to ensure uniformity in data representation. The normalization method is given by:

$$x_{norm} = \frac{x_{data} - \mu}{\sigma} \qquad (1)$$

where μ denotes the mean of all training data, and σ represents the standard deviation of the entire training dataset.

The same pre-processing steps were applied to the test data, ensuring a standardized and comparable basis for subsequent analyses.

Proposed Network Architecture

In pursuit of accurate and efficient predictions, the approach seeks to balance computational efficiency with the capacity to discern intricate features essential for the precise classification of closely related time-series data. Dilated causal convolution (31), initially introduced for speech recognition, was integrated to enable parallel computation and encompasses a wide receptive field for capturing long-range temporal dependencies. Concurrently, long short-term memory (LSTM) (32), a type of recurrent neural network (RNN) recognized for its gating mechanisms that stabilize training and demonstrate robust learning capabilities, was incorporated.

Drawing upon the exceptional properties of these networks in handling time-series data, a one-dimensional supervised classification network was designed for automatic tissue segmentation, specifically targeting the discrimination of synovitis from arteries. The proposed architecture integrates dilated causal convolution and long short-term memory (LSTM), leveraging one-dimensional time-intensity data as input. The framework of the proposed model is illustrated in Figure 3.The front (left) half of the model is comprised of a two-layer feature extraction section, encompassing dilated causal convolution, batch normalization with scaled exponential linear unit (SELU) activation functions (33). The parameters of the dilated causal convolution were configured, incorporating dilated factors of 1 and 2, output channels of 128 and 256, and kernel sizes of 4 and 2 for layer depths 1 and 2, respectively.

The latter (right) half of the model consists of two LSTM layers, followed by a Softmax layer. The initial LSTM layer endeavors to yield hidden state values across all time steps, while the second LSTM layer is tailored to produce the output value in the last time step. Finally, Softmax computed the probability-based output result for categorizing the two target classes.

The development process was conducted in the PyCharm environment, primarily utilizing the Keras package for network construction. Data processing tasks were carried out using the NumPy and Pandas packages, while performance metrics were computed using the Scikit-learn package. Additionally, the Matplotlib package was employed for data visualization purposes.

Network Training

All training and evaluation procedures were executed on a desktop computer running 64-bit Windows, using the Python 3.6 programming language with TensorFlow and Keras. The training utilized an AMD Ryzen 9 6900HX CPU with 32 GB RAM. The entire process was conducted without the use of a GPU. The loss function employed in the network comprised a combination of sigmoid activation and binary cross entropy (BCE) loss function. The Adam optimizer was utilized with a learning rate of 0.0001. Dynamic learning rate adjustment was implemented, halving the rate when a substantial decline in loss was not observed over four consecutive epochs. Training consisted of 60 epochs, with a batch size of 16. LOOCV was used to ensure the network's reliability throughout training and testing. Label smoothing (34) was

applied as a regularization technique to augment the network's generalization capability. The label smoothing parameter α was set to 0.1. Furthermore, dropout layers (35) were inserted following SELU activation and the two LSTM layers, with a dropout rate of 0.2.

The total training time for the model was approximately 190 seconds for 60 epochs. The processing time (feed-forward) for one joint ROI was 0.5 seconds, while the processing time for the entire hand segmentation was 8 seconds.

Pre-Quantification Area Segmentation

Analysis of the TIC shape reveals a distinctive pattern wherein synovitis and arteries undergo near-maximal enhancement in the initial phase of enhancement, defining the time to peak (T_{peak}) (19). Leveraging this prominent TIC shape feature, we employed a rapid unsupervised segmentation method (*k*-means clustering) to delineate the highest-intensity areas primarily composed of synovitis and arteries, shown in Figure 1. The segmentation was conducted in the phase near the T_{peak}, usually within the 8th-10th phase. The extraction method for the highest-intensity areas is illustrated in Figure 4.

Initially, the phase nearest T_{peak} was selected from the 27 phases. Subsequently, Otsu's method (36) was applied to ascertain the optimal threshold to filter the background and most bone areas. Pixels in these areas were assigned a value of 0 (black). The upper confidence bound for the remaining area was computed using the

Z critical value with a confidence interval of 75%. The Upper confidence bound is given by:

$$Upper\ Confidence\ Bound_{75\%CI} = mean + (1.15 * std) \qquad (2)$$

where *mean* represents the mean gray value of pixels above the optimal threshold calculated by Otsu's method, and *std* is the standard deviation of those pixels. Pixels with gray values lower than the upper confidence bound were set to a value of 0.

During joint ROI testing, the unsupervised image segmentation method based on *k*-means clustering rapidly and automatically clusters the highest-intensity areas. This method was applied to 51 joint ROIs with synovitis drawn by ImageJ, obtaining all highest-intensity areas to be tested, including synovitis or arteries. Similarly, we applied the same method to whole hands in different slices to obtain the highest-intensity areas for subsequent testing.

Statistical Analysis

To assess the performance of the proposed model, various metrics were employed, including accuracy (ACC), sensitivity (SEN), specificity (SPEC), positive predictive value (PPV), negative predictive value (NPV), and the area under the curve (AUC) of receiver operating characteristic (ROC) during LOOCV. The evaluation was conducted across 12 different acquisition times (phase 1-5, phase 1-7, ..., up to phase 1-27) using distinct validation datasets. The disparities between the predicted outcomes of different models and the corresponding ground truth were quantified and scrutinized during testing. This analysis incorporated the Dice and Spearman's rank

correlation coefficients as key metrics to evaluate the agreement and correlation in joint ROIs, respectively. Additionally, to evaluate model performance for the whole hand, we quantitatively assessed synovitis across different slices for each patient. Each slice's automatically quantified synovitis area was color-coded to create a color map. Based on varying phases, Spearman's rank correlation coefficients were calculated between the predicted synovial volume for each hand.

The relationship was considered almost negligible when R-values were less than 0.2; definite but small when R-values were 0.2- 0.4; substantial when R-values were 0.4- 0.7; marked when R-values were 0.7- 0.9; and very dependable when R-values were 0.9- 1.0(37). ICC was used to assess reliability with a one-way random effects absolute agreement model (34). Reliability was categorized based on ICC values: values below 0.40 were deemed poor, those between 0.41 and 0.59 were considered fair, values between 0.60 and 0.74 were rated as good, and values between 0.75 and 1.00 were regarded as excellent. All Statistical analysis was performed using Python 3.6 and IBM Statistical Package for Social Sciences (SPSS) Statistics 26 for Windows (IBM Corp, Armonk, NY). $P < 0.05$ was considered statistically significant.

Results

Network Training and Validation

The average evaluation values were calculated from 12 batches of validation results, as depicted in Table 1. The performance metrics across different phases ranged from:

ACC from 0.853 ± 0.044 to 0.970 ± 0.007; SEN from 0.888 ± 0.046 to 0.985 ± 0.007; SPEC from 0.787 ± 0.050 to 0.947 ± 0.014; PPV from 0.883 ± 0.047 to 0.971 ± 0.011; NPV from 0.792 ± 0.070 to 0.968 ± 0.011; AUC from 0.837 ± 0.042 to 0.965 ± 0.024. In general, the model's pixel classification ability diminishes as the length of the phase decreases during validation.

Joint ROI Testing

For ROI testing, the segmentation of highest-intensity areas, primarily synovitis and arteries, was carried out on 51 selected joint ROIs, followed by classification.. Spearman's rank correlation coefficient assessed the correlation between predicted synovitis areas and ground truth. The ROI testing process is illustrated in Figure 5.

The Dice coefficient, ranged from 0.475 ± 0.034 to 0.615 ± 0.013, and Spearman's rank correlation coefficient, varied from 0.783 ± 0.034 to 0.927 ± 0.007, as shown in Table 2. With increasing phase sequence length, both metrics showed an upward trend, suggesting enhanced performance. Specifically, when the phase length exceeded 15, the overall Dice coefficient approached 0.6, while Spearman's Rank Correlation consistently approximated 0.9. Additionally, in six joint ROIs, synovitis intensity below the detection threshold resulted in a Dice coefficient of 0, consequently causing the Dice coefficient reduced.

Whole Hand Testing

RAMRIS was evaluated by an MSK radiologist (T.K), as shown in Figure 5. Table 3 displays the correlation analysis between AI segmentation and manual outlining in

whole hands across various phase sequence length. The Spearman's rank correlation coefficient ranged from 0.611 to 0.831 across phases from 1-5 to 1-27. The lowest value was observed at 1-9, while the highest value occurred at 1-27. The trend in the predicted outcomes closely mirrors that of joint ROIs, with performance decreasing as the phase sequence length decreases.

Notably, acquisition times between 15 and 27 phases (approximately 2.5 to 4.5 minutes) achieved high clinical agreement and exhibited marked correlation with inflammatory activity and RAMRIS, with R-values ranging from 0.736 to 0.831.

Additionally, reliability was assessed using ICC with a one-way random effects absolute agreement model (29). The ICC values for single measures demonstrated high consistency in interobserver agreement for RAMRIS scoring, with a value of 0.945 (95% confidence interval: 0.878-0.976).

Representative instances of whole-hand segmentation colormap results can be seen in Figure 6. Furthermore, an additional test set (N=35) comprised varying numbers of phases, only 15 phases were utilized for prediction by the model trained using 15 phases. The Spearman correlation coefficient between prediction synovium volume and RAMRIS gold standard was found to be 0.768, with a significance level of $p < 0.001$, as depicted in Figure 7.

Discussion

In this retrospective study, we developed a novel AI-based model for synovial quantification, specifically tailored to visualize inflammatory areas across the whole

hand of RA patients using DCE-MRI datasets. To the best of our knowledge, this is the first paper that utilizes deep learning for pixel-by-pixel determination of dynamic signal intensity changes in synovitis on DCE-MRI. Our method can automatically segment lesion areas while excluding arteries in RA quantification. The results show that our approach, distinct from current semi-quantitative methods (21), achieves fully automated and end-to-end quantification of synovitis on whole hands. The strategic design of our model not only achieves noteworthy reductions in acquisition times but also effectively mitigates arterial interference during synovial quantification. Furthermore, the observed high Dice coefficient between our automated segmentation and ground truth, demonstrate a relationship with RA inflammatory activity.

Once type 4 TIC, characterized by rapid enhancement and rapid washout, was identified to be strongly associated with RA in patients with positive knee arthritis, there was a recognition of the potential of TIC shape analysis in early differentiation between RA and non-RA patients (20). After this, pixel-by-pixel TIC shape analysis for quantifying the synovial membrane exhibited a robust correlation with manual contouring in the wrist (21). In order to mitigate errors in selecting arterial pixels when detecting synovitis pixels, an arterial mask subtraction method was then introduced (22).

Despite the results of high-precision quantification of synovitis without vascular negative influence, some serious limitations are still mentioned and difficult to solve. First, previous methods relied on manual identification of TIC features, leading to

time-consuming processes with observer variability and limited scalability (21, 22). Our approach integrates machine learning and deep learning, eliminating the need for manual TIC feature quantification, improving analysis efficiency and accuracy. Second, many existing RA quantification approaches lacked motion correction during DCE-MRI acquisition (38-40), causing potential distortions in TIC and inaccuracies in synovial quantification. Image preprocessing techniques such as denoising and registration can enhance data quality and minimize artifacts. Third, shorter DCE-MRI scan times are desirable for workflow efficiency, but they often result in increased image noise from undue motion of the patients, compromising data quality and analysis accuracy without adequate noise reduction strategies. The findings suggest that a nearly 50% reduction in acquisition time, effectively mitigates noise and motion artifacts. Fourth, despite performance advancements, processing times exceeding 100 seconds were needed (22), making real-time analysis impossible. The AI-based lightweight model drastically reduces processing time, with predictions taking only approximately 8 seconds. What's more, our automated workflow ensures end-to-end quantification by taking DCE-MRI slices as input and generating quantitative colormaps as output, streamlining the assessment of RA.

How to efficiently distinguish synovitis from arteries, which have very similar TIC shape characteristics, is a difficult problem. To address this issue and enhance the accuracy of pixel-by-pixel determination of dynamic signal intensity changes in synovitis on MRI, our proposed unsupervised image segmentation efficiently sorts

and clusters all pixels from DCE-MRI slices in conjunction with basic image processing algorithms. This process effectively filters out irrelevant tissues, including muscles, bones and background, contributing to improved segmentation. Moreover, developing a model with minimal parameters is essential for achieving swift and accurate performance of pixel-by-pixel determination of dynamic signal intensity changes in synovitis on MRI. Integrating dilated causal convolution is instrumental in expanding the model's receptive field, facilitating the extraction of more features and mitigating overfitting during the training of time series data. The SELU activation function is critical in maintaining network stability by preventing neuron death. Additionally, the gate structure inherent to LSTM proves adept at capturing sequence features, particularly valuable when dealing with closely aligned characteristics in short sequences observed in DCE-MRI, thereby significantly enhancing the overall quantification performance of the proposed model.

To further explore the optimization of DCE-MRI acquisition time, various phase sequence lengths were investigated to assess their impact on synovial quantification outcomes. Our findings revealed that even with shorter acquisition times, as demonstrated by approximately half the usual duration, statistically acceptable results were obtained, highlighting the potential for reducing acquisition time without compromising accuracy. Our findings align with the inherent characteristics of DCE-MRI, highlighting the delicate balance between acquisition time and pixel intensity enhancement crucial for precise classification. Our results substantiate that

selectively retaining a portion of the washout phase significantly enhances the quantitative performance of automated algorithms, presenting a promising avenue for shortening DCE-MRI acquisition time without compromising accuracy.

Limitations

There are several obvious limitations that warrant consideration. The relatively small cohort size restricts the generalizability of our findings. The study was conducted at a single site using data acquired from a single vendor, scanner, and field strength as well. These factors collectively limit the broader applicability of our results and underscore the need for future research to validate our findings. In addition, there are some other limitations.

First, one-dimensional deep learning networks only learn time series characteristics, while ignoring the relationship between adjacent pixels. Future research will explore the relationship between pixels to improve segmentation.

Second, although only 12 cases were used in this study, the synovitis area in whole hands has different characteristics in different locations on DCE-MRI. Due to variations in individual factors, such as blood perfusion, transport of contrast agent across vessel walls, and diffusion of contrast medium in the interstitial space, the concentration of agent retained in the body varies, resulting in unavoidable differences in the dynamic intensity of pixels even in the same category of each tissue (41). This study extracted thousands of pixels from these 12 cases for classification and verified that the model can capture most of these feature differences and

effectively distinguish synovitis pixels from other tissues. However, when validating the model's generalizability, more cases are still needed.

Third, this study only explored examples with a phase sequence length of 27 from the same clinic. In the future, it will be necessary to combine data with different phase sequence lengths and from different clinics to evaluate and improve the stability and effectiveness of the model.

Finally, in unsupervised segmentation, while most irrelevant tissue pixels can be efficiently removed, saving a significant amount of time, there may be a risk of preserving some muscle or deleting some synovitis regions due to setting the Z critical value within a 75% confidence interval. This variability could result in errors while evaluating the model results in individual cases. The latent issue in question has not manifested significant repercussions thus far; however, it is prudent to consider alternative methods which may help enhance the robustness of the entire process and mitigate potential concerns in future cases to address potential concerns preemptively.

Conclusions

This work establishes the feasibility of quantifying enhanced hand RA synovium through a deep learning approach based on pixel-by-pixel determination of dynamic signal intensity changes in synovitis on MRI. The proposed method offers a rapid, accurate, end-to-end solution for synovial quantification, demonstrating strong agreement with manual annotations by an MSK radiologist (T.K). This advancement holds promise for providing efficient and reliable imaging evidence in the early

detection and follow-up treatment of RA in the hand, paving the way for improved patient care in rheumatology practice.

References

1. Gravallese EM, Firestein GS: Rheumatoid Arthritis — Common Origins, Divergent Mechanisms. *N Engl J Med* 2023; 388:529–542.

2. Guo Q, Wang Y, Xu D, Nossent J, Pavlos NJ, Xu J: Rheumatoid arthritis: pathological mechanisms and modern pharmacologic therapies. *Bone Res* 2018; 6:15.

3. Aletaha D, Smolen JS: Diagnosis and Management of Rheumatoid Arthritis: A Review. *JAMA* 2018; 320:1360–1372.

4. Finckh A, Liang MH, van Herckenrode CM, de Pablo P: Long-term impact of early treatment on radiographic progression in rheumatoid arthritis: A meta-analysis. *Arthritis Care Res* 2006; 55:864–872.

5. Scott DL, Bacon PA: Joint damage in rheumatoid arthritis: radiological assessments and the effects of anti-rheumatic drugs. *Rheumatol Int* 1985; 5:193–199.

6. Østergaard M, Edmonds J, McQueen F, et al.: An introduction to the EULAR–OMERACT rheumatoid arthritis MRI reference image atlas. *Ann Rheum Dis* 2005; 64(suppl 1): i3–i7.

7. Bird P, Ejbjerg B, McQueen F, Ostergaard M, Lassere M, Edmonds J: OMERACT Rheumatoid Arthritis Magnetic Resonance Imaging Studies. Exercise 5: an international multicenter reliability study using computerized MRI erosion volume measurements. *J Rheumatol* 2003; 30:1380–1384.

8. McQueen F, Lassere M, Edmonds J, et al.: OMERACT Rheumatoid Arthritis Magnetic Resonance Imaging Studies. Summary of OMERACT 6 MR Imaging Module. *J Rheumatol* 2003; 30:1387–1392.

9. Østergaard M, Peterfy C, Conaghan P, et al.: OMERACT Rheumatoid Arthritis Magnetic Resonance Imaging Studies. Core set of MRI acquisitions, joint pathology definitions, and the OMERACT RA-MRI scoring system. The Journal of Rheumatology 2003; 30:1385–1386.

10. Ostergaard M, McQueen FM, Bird P, et al.: Magnetic resonance imaging in rheumatoid arthritis advances and research priorities. *J Rheumatol* 2005; 32:2462–2464.

11. Peterfy C, Østergaard M, Conaghan PG: MRI comes of age in RA clinical trials. Ann Rheum Dis 2013; 72:794–796.

12. American College of Rheumatology Rheumatoid Arthritis Clinical Trials Task Force Imaging Group and Outcome Measures in Rheumatology Magnetic Resonance Imaging Inflammatory Arthritis Working Group: Review: the utility of magnetic resonance imaging for assessing structural damage in randomized controlled trials in rheumatoid arthritis. *Arthritis Rheum* 2013; 65:2513–2523.13. Østergaard M, Peterfy C, Conaghan P, et al.: OMERACT Rheumatoid Arthritis Magnetic Resonance Imaging Studies. Core set of MRI acquisitions, joint pathology definitions, and the OMERACT RA-MRI scoring system. *J Rheumatol* 2003; 30:1385–1386.

14. Bird P, Conaghan P, Ejbjerg B, et al.: The development of the EULAR–OMERACT rheumatoid arthritis MRI reference image atlas. *Ann Rheum Dis* 2005; 64(Suppl 1): i8–i10.

15. Lassere M, McQueen F, Østergaard M, et al.: OMERACT Rheumatoid Arthritis Magnetic Resonance Imaging Studies. Exercise 3: an international multicenter reliability study using the RA-MRI Score. *J Rheumatol* 2003; 30:1366–1375.

16. Conaghan P, Lassere M, Østergaard M, et al.: OMERACT Rheumatoid Arthritis Magnetic Resonance Imaging Studies. Exercise 4: an international multicenter longitudinal study using the RA-MRI Score. *J Rheumatol* 2003; 30:1376–1379.

17. Hodgson RJ, Connolly S, Barnes T, Eyes B, Campbell RSD, Moots R: Pharmacokinetic modeling of dynamic contrast-enhanced MRI of the hand and wrist in rheumatoid arthritis and the response to anti-tumor necrosis factor-α therapy. *Magn Reson Med* 2007; 58:482–489.

18. Ostergaard M, Stoltenberg M, Henriksen O, Lorenzen I: Quantitative assessment of synovial inflammation by dynamic gadolinium-enhanced magnetic resonance imaging. A study of the effect of intra-articular methylprednisolone on the rate of early synovial enhancement. *Br J Rheumatol* 1996; 35:50–59.

19. van der Leij C, van de Sande MGH, Lavini C, Tak PP, Maas M: Rheumatoid Synovial Inflammation: Pixel-by-Pixel Dynamic Contrast-enhanced MR Imaging Time-Intensity Curve Shape Analysis—A Feasibility Study. *Radiology* 2009; 253:234–240.

20. van de Sande MGH, van der Leij C, Lavini C, Wijbrandts CA, Maas M, Tak PP: Characteristics of synovial inflammation in early arthritis analysed by pixel-by-pixel time-intensity curve shape analysis. *Rheumatology* 2012; 51:1240–1245.

21. Sakashita T, Kamishima T, Kobayashi Y, et al.: Accurate quantitative assessment of synovitis in rheumatoid arthritis using pixel-by-pixel, time–intensity curve shape analysis. *Br J Radiol* 2016; 89:20151000.

22. Kobayashi Y, Kamishima T, Sugimori H, et al.: Quantification of hand synovitis in rheumatoid arthritis: Arterial mask subtraction reinforced with mutual information can improve accuracy of pixel-by-pixel time-intensity curve shape analysis in dynamic MRI. *J Magn Reson Imaging JMRI* 2018.

23. Wakefield RJ, O'Connor PJ, Conaghan PG, et al.: Finger tendon disease in untreated early rheumatoid arthritis: A comparison of ultrasound and magnetic resonance imaging. *Arthritis Care Res* 2007; 57:1158–1164.

24. Ochi J, Mori N, Mori Y, et al.: Validating an Empirical Mathematical Model for Dynamic Contrast-enhanced MR Imaging of Hand and Wrist Synovitis in Rheumatoid Arthritis: Correlation of Model Parameters with Clinical Disease Activity. *Magn Reson Med Sci MRMS Off J Jpn Soc Magn Reson Med* 2020; 19:176–183.25. Hinton G: Deep Learning—A Technology With the Potential to Transform Health Care. *JAMA* 2018; 320:1101–1102.

26. Le WT, Maleki F, Romero FP, Forghani R, Kadoury S: Overview of Machine Learning: Part 2: Deep Learning for Medical Image Analysis. *Neuroimaging Clin N*

Am 2020; 30:417–431. [*Machine Learning and Other Artificial Intelligence Applications*]

27. Chan H-P, Samala RK, Hadjiiski LM, Zhou C: Deep Learning in Medical Image Analysis. In *Deep Learn Med Image Anal Chall Appl*. Edited by Lee G, Fujita H. Cham: Springer International Publishing; 2020:3–21. [*Advances in Experimental Medicine and Biology*]

28. Kay J, Upchurch KS: ACR/EULAR 2010 rheumatoid arthritis classification criteria. *Rheumatology* 2012; 51(suppl_6):vi5–vi9.

29. McGraw KO, Wong SP: Forming inferences about some intraclass correlation coefficients. Psychological Methods 1996; 1:30–46.

30. Reddy BS, Chatterji BN: An FFT-based technique for translation, rotation, and scale-invariant image registration. *IEEE Trans Image Process Publ IEEE Signal Process Soc* 1996; 5:1266–1271.31. Bai S, Kolter JZ, Koltun V: An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling. 2018.

32. Hochreiter S, Schmidhuber J: Long Short-term Memory. *Neural Comput* 1997; 9:1735–80.

33. Klambauer G, Unterthiner T, Mayr A, Hochreiter S: Self-Normalizing Neural Networks. In *Adv Neural Inf Process Syst. Volume 30*. Curran Associates, Inc.; 2017.

34. Müller R, Kornblith S, Hinton G: When does label smoothing help? In *Proc 33rd Int Conf Neural Inf Process Syst*. Red Hook, NY, USA: Curran Associates Inc.; 2019:4694–4703. [vol. 422]

35 Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R: Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *J Mach Learn Res* 2014; 15:1929–1958.

36. Otsu N: A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans Syst Man Cybern* 1979; 9:62–66.

37. Dashiell JF: Fundamental Statistics In Psychology And Education, 1st edition. New York and London: Mcgraw-Hill Book Company Inc; 1942. 219 p.

38. Workie DW, Dardzinski BJ, Graham TB, Laor T, Bommer WA, O'Brien KJ: Quantification of dynamic contrast-enhanced MR imaging of the knee in children with juvenile rheumatoid arthritis based on pharmacokinetic modeling. Magnetic Resonance Imaging 2004; 22:1201–1210.

39. Kirkhus E, Bjørnerud A, Thoen J, Johnston V, Dale K, Smith H-J: Contrast-enhanced dynamic magnetic resonance imaging of finger joints in osteoarthritis and rheumatoid arthritis: an analysis based on pharmacokinetic modeling. Acta Radiol 2006; 47:845–851.

40. Waterton JC, Ho M, Nordenmark LH, et al.: Repeatability and response to therapy of dynamic contrast-enhanced magnetic resonance imaging biomarkers in rheumatoid arthritis in a large multicentre trial setting. *Eur Radiol* 2017; 27:3662–3668.

41. Padhani AR: Dynamic contrast-enhanced MRI in clinical oncology: Current status and future directions. J Magn Reson Imaging 2002; 16:407–422.

Tables

Table 1. Leave-one-out cross-validation performance metrics for model evaluation across different phase sequences

Phases used	ACC		SEN		SPEC		PPV		NPV		AUC	
	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD
1-5	0.853	0.044	0.888	0.046	0.787	0.050	0.883	0.047	0.792	0.070	0.837	0.042
1-7	0.932	0.025	0.965	0.021	0.843	0.035	0.918	0.029	0.932	0.026	0.904	0.024
1-9	0.934	0.017	0.968	0.023	0.865	0.036	0.931	0.018	0.941	0.029	0.916	0.018
1-11	0.948	0.016	0.976	0.070	0.895	0.037	0.944	0.022	0.952	0.017	0.935	0.021
1-13	0.952	0.013	0.976	0.012	0.904	0.034	0.950	0.016	0.953	0.019	0.940	0.018
1-15	0.954	0.007	0.981	0.006	0.902	0.014	0.949	0.011	0.962	0.012	0.941	0.009
1-17	0.955	0.010	0.981	0.007	0.906	0.025	0.950	0.015	0.962	0.015	0.943	0.013
1-19	0.956	0.009	0.980	0.006	0.909	0.024	0.953	0.011	0.960	0.015	0.944	0.013
1-21	0.958	0.010	0.980	0.007	0.918	0.029	0.957	0.015	0.959	0.014	0.949	0.015
1-23	0.963	0.013	0.985	0.007	0.922	0.033	0.959	0.020	0.968	0.016	0.953	0.018
1-25	0.967	0.010	0.983	0.005	0.936	0.029	0.967	0.013	0.967	0.011	0.959	0.015
1-27	0.970	0.007	0.982	0.006	0.947	0.014	0.971	0.011	0.965	0.012	0.965	0.009

ACC, accuracy; SEN, sensitivity; SPEC, specificity; PPV, positive predictive value; NPV, negative predictive value; AUC, area under the receiver operating characteristic curve; SD, standard deviation

Table 2. Correlation analysis between AI segmentation and manual outlining in joint ROIs across different phase sequences: Dice coefficient and Spearman's rank correlation coefficient (R-value) for synovial regions

Phases used	Dice score	Spearman's rank correlation coefficient	
		R-value	p-value
1-5	0.475	0.783	<0.001
1-7	0.511	0.651	<0.001
1-9	0.515	0.655	<0.001
1-11	0.583	0.882	<0.001
1-13	0.587	0.882	<0.001
1-15	0.603	0.913	<0.001
1-17	0.585	0.911	<0.001
1-19	0.577	0.916	<0.001
1-21	0.603	0.920	<0.001
1-23	0.615	0.926	<0.001
1-25	0.557	0.884	<0.001
1-27	0.605	0.927	<0.001

Table 3. Correlation analysis between AI segmentation and manual outlining in whole hands across different phase sequences: Dice coefficient and Spearman's rank correlation coefficient (R-value) for synovial volume

Phases used	Spearman's rank correlation coefficient	
	R-value	p-value
1-5	0.695	<0.001
1-7	0.639	0.001
1-9	0.611	0.002
1-11	0.803	<0.001
1-13	0.743	<0.001
1-15	0.823	<0.001
1-17	0.754	<0.001
1-19	0.736	<0.001
1-21	0.813	<0.001
1-23	0.791	<0.001
1-25	0.819	<0.001
1-27	0.831	<0.001

Figure legends

Figure 1: Overall algorithm flow chart. Yellow and blue arrows represent model testing and training, respectively. At first, DCE-MRI underwent pre-processing, primarily involving image denoising and registration. One-dimensional time-intensity data was obtained from DCE-MRI. Deep learning was then used for classification training. Image segmentation was used for the training results. The trained pixel-by-pixel classification model was used to segment the synovial area on the images with the highest-intensity areas generated by the pre-quantification area segmentation. The segmentation performance was tested with two steps: joint ROI and whole-hand testing. The Dice and Spearman's rank correlation coefficients were evaluated for similarity between AI segmentation and manual outlining in ROIs with synovitis. Spearman's rank correlation coefficient was then evaluated for the relationship between synovial volume via AI segmentation with RAMRIS across whole hands. RAMRIS: Rheumatoid Arthritis Magnetic Resonance Imaging Scoring.

Figure 2: Overall data partition method based on LOOCV. We performed 12 model training and verification based on LOOCV for one phase sequence, and we had 12 different phase sequences that needed to be trained. The study utilized a training set (4585.17±308.98, min 3982, max 4872), a validation set (417.83±308.98, min 130, max 1020), and a test set (3.67±3.14, min 1, max 11). The training and validation sets consisted of one-dimensional time-intensity data, while the test set, comprising slices

from DCE-MRI used in joint ROI and whole-hand testing, was employed to verify the local and global performance of the model. LOOCV: leave-one-out cross-validation.

Figure 3: Diagram of the proposed framework. The inputs have 12 different lengths (5, 7. . . 27 phases). The input normalized one-dimensional time-intensity values of different phases (15 in this example) that enter model training are marked in gray while the remaining phase sequences that are not used (16-27 in this example) are marked in red on the normalized time-intensity curves. In this example, the input comprises γ nodes (phases used) multiplied by one dimension. From the input, circular nodes depict individual neurons within the network, while solid lines represent the initial convolution step and dashed lines denote subsequent steps. The initial segment of the model incorporates a dual-layer feature extraction component, which includes Dilated Causal Convolution, Batch Normalization, and the SELU Activation Function. The subsequent segment involves a two-layer LSTM layer. The final LSTM layer solely provides the output of the last time step, generating 32 multiplied by 1 as the output result. Subsequently, Softmax is applied to calculate the probability-based output, facilitating categorizing the two target classes. The output is the classification result of each pixel, synovitis or arteries.

SELU: scaled exponential linear units; LSTM: long short-term memory.

Figure 4: Flow diagram of the pre-quantification area segmentation method. A: The original image. B: Setting pixel gray values to 0 below a threshold, computed using Otsu's method and confidence upper bound in sequence. C: The remaining pixels are segmented to obtain the highest-intensity area based on the unsupervised *k*-means clustering. D: Colormap for segmentation of the highest-intensity area of the whole hands (primarily including synovitis and arteries). E: Joint ROIs drawn manually. F: The highest-intensity area in the joint ROIs to be tested. The extracted pixels in the highest-intensity area are classified by the model for testing performance.

Figure 5: Correlation analysis evaluation process in joint ROIs and whole hands. A1: Sample joint ROIs. A2: Pixel-by-pixel classification of the joint ROIs using deep learning after pre-quantification area segmentation. A3: Prediction result of the joint ROIs. A4: Manual outlining. A3 and A4 were compared with the Dice coefficient and Spearman's rank correlation coefficient analysis. B1: Sample slice of the whole hands. B2: Pixel-by-pixel classification of the whole hands using deep learning after pre-quantification segmentation. B3: Prediction result of the whole hands. The number of synovitis pixels counted for each hand was evaluated using Spearman's rank correlation coefficient analysis, with the scores graded using RAMRIS by an MSK radiologist (T.K). RAMRIS: Rheumatoid Arthritis Magnetic Resonance Imaging Scoring

Figure 6: Representative whole-hand segmentation colormap results from three cases with rheumatoid arthritis (RA). The left panel illustrates the source images of the RA patients at the 15th phase of DCE-MRI (acquisition time of 2.5 minutes), whereas the right panel displays the images segmented by AI. The latter was generated using DCE-MRI data spanning 1 to 15 phases (acquisition time of 2.5 minutes). Case A is a 41-year-old female with tenosynovitis of the flexor tendons. Case B is a 63-year-old male with prominent synovitis in the finger joints. Case C, a 72-year-old female, exhibits complex synovitis morphology in the carpal region. These lesions are highlighted in green in the AI segmentation images, while arteries are indicated in red.

Figure 7: Correlation between RAMRIS Score (x-axis) and Prediction Synovium Volume (y-axis). RAMRIS: Rheumatoid Arthritis Magnetic Resonance Imaging Scoring

Figure S1: Distribution of Number of Phases in the Test Set

ADDITIONAL GUIDELINES

Financial Interest

No conflicts of interest were identified.

Figure 1

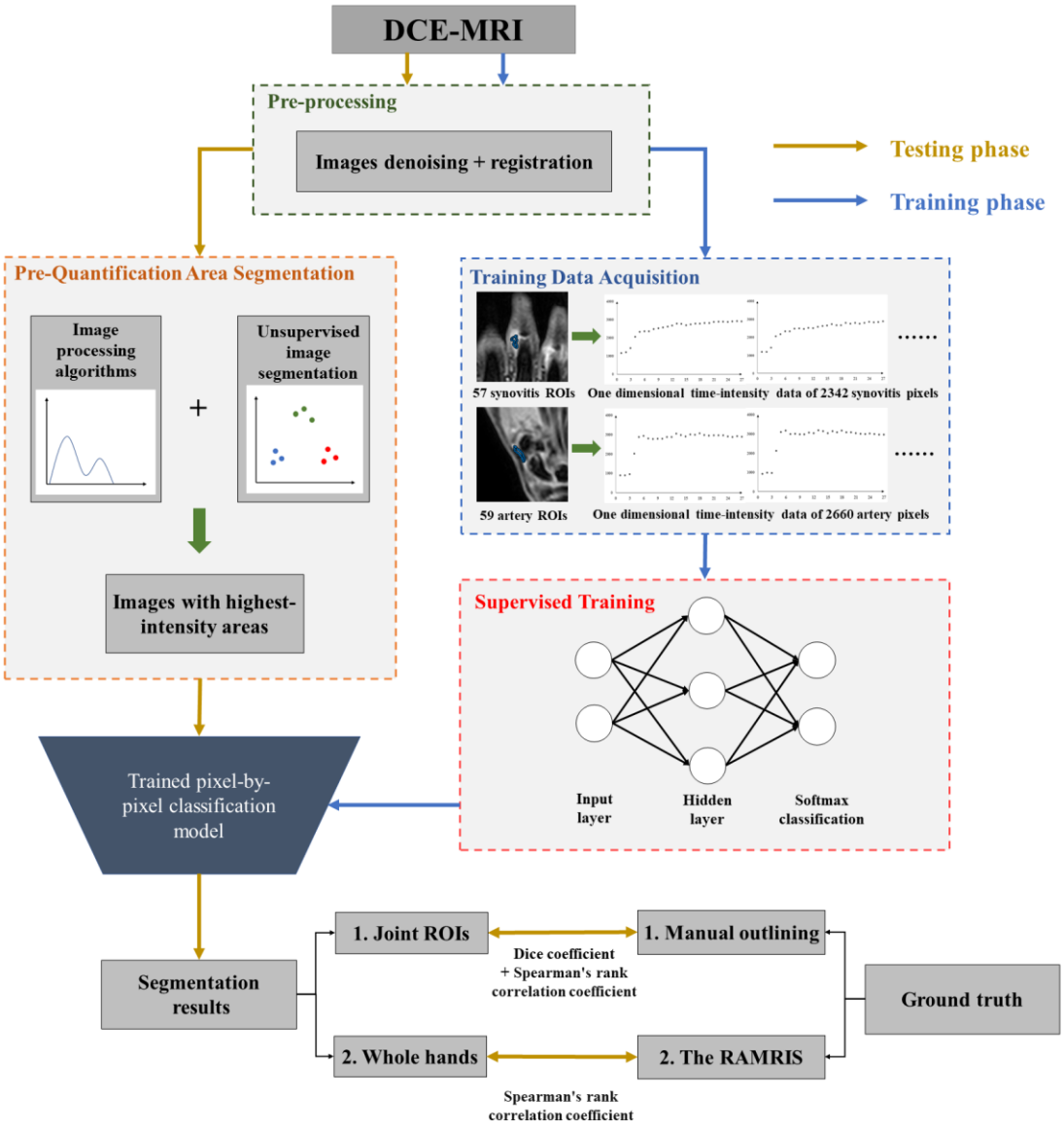

DCE-MRI
Pre-processing
Images denoising + registration
Testing phase
Training phase
Pre-Quantification Area Segmentation
Image processing algorithms
+
Unsupervised image segmentation
Images with highest-intensity areas
Training Data Acquisition
57 synovitis ROIs
One dimensional time-intensity data of 2342 synovitis pixels
59 artery ROIs
One dimensional time-intensity data of 2660 artery pixels
Supervised Training
Input layer
Hidden layer
Softmax classification
Trained pixel-by-pixel classification model
Segmentation results
1. Joint ROIs
1. Manual outlining
Dice coefficient + Spearman's rank correlation coefficient
2. Whole hands
2. The RAMRIS
Spearman's rank correlation coefficient
Ground truth

Figure 2

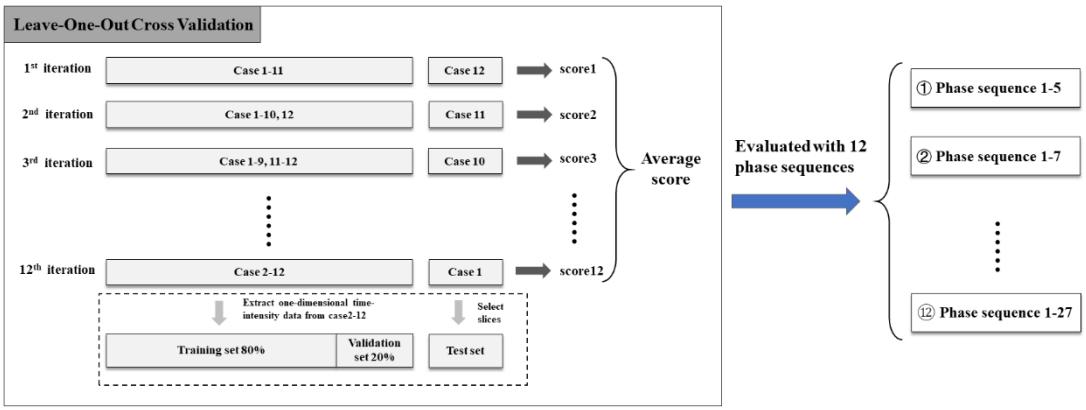

Leave-One-Out Cross Validation
1st iteration
Case 1-11
Case 12
score1
2nd iteration
Case 1-10, 12
Case 11
score2
3rd iteration
Case 1-9, 11-12
Case 10
score3
12th iteration
Case 2-12
Case 1
score12
Average score
Extract one-dimensional time-intensity data from case2-12
Select slices
Training set 80%
Validation set 20%
Test set
Evaluated with 12 phase sequences
① Phase sequence 1-5
② Phase sequence 1-7
⑫ Phase sequence 1-27

Figure 3

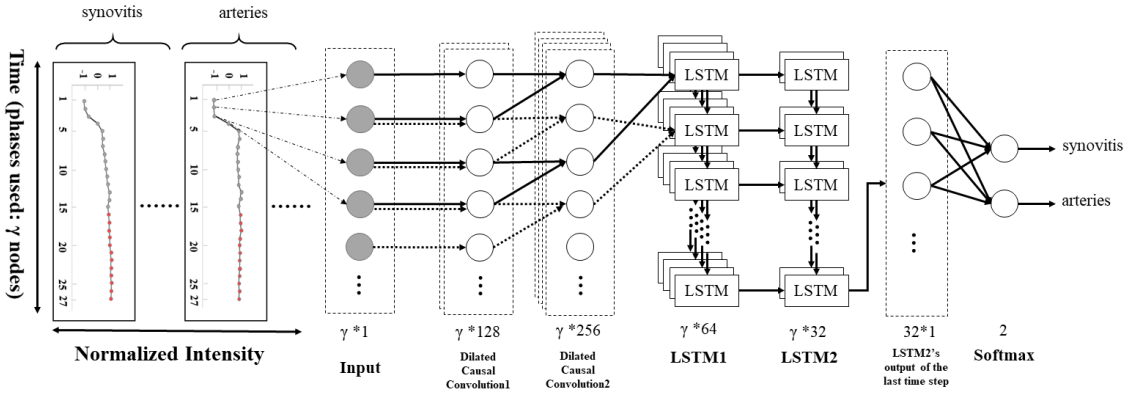

synovitis
arteries
Time (phases used: γ nodes)
Normalized Intensity
γ *1
Input
γ *128
Dilated Causal Convolution1
γ *256
Dilated Causal Convolution2
LSTM
γ *64
LSTM1
γ *32
LSTM2
32*1
LSTM2's output of the last time step
2
Softmax
synovitis
arteries

Figure 4

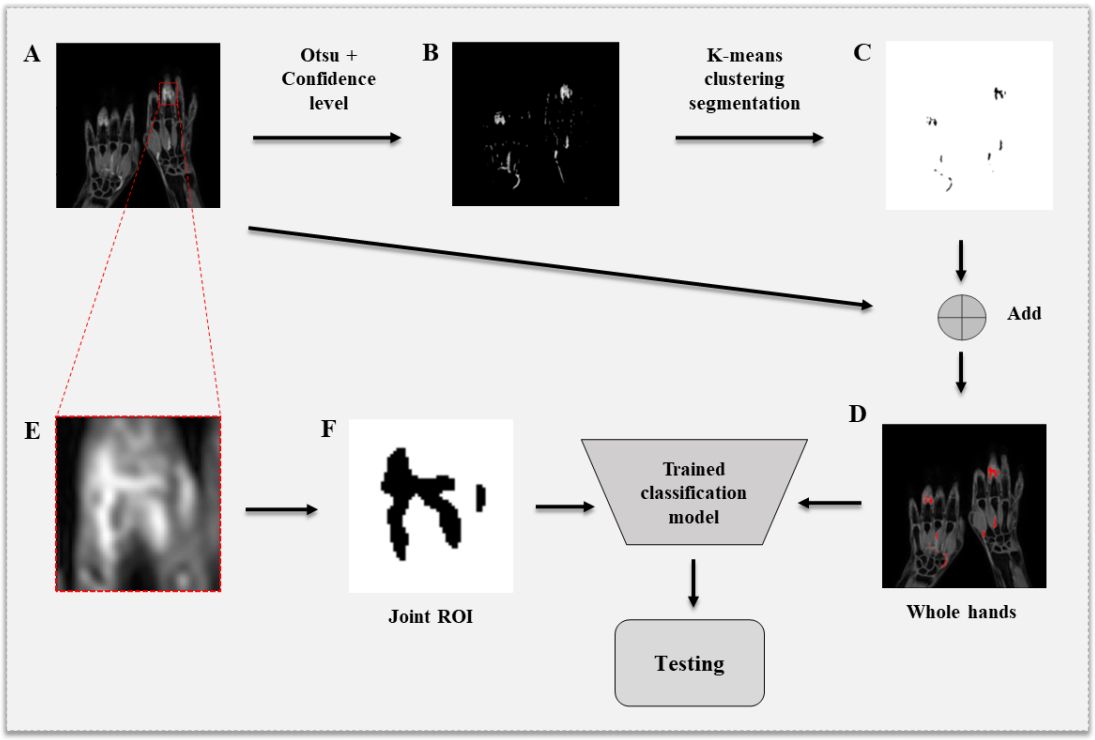

A
Otsu +
Confidence
level
B
K-means
clustering
segmentation
C
Add
D
Whole hands
E
F
Joint ROI
Trained
classification
model
Testing

Figure 5

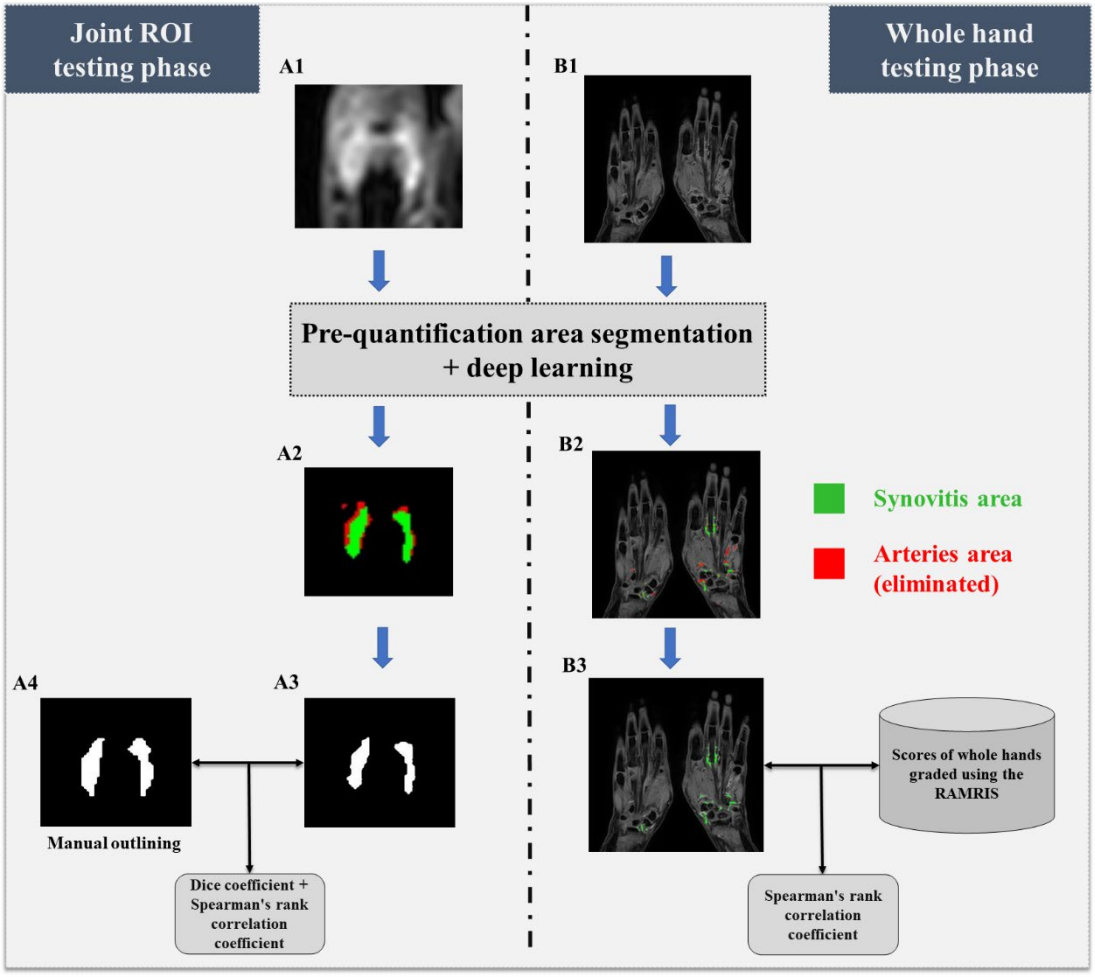

Joint ROI testing phase
Whole hand testing phase
A1
B1
Pre-quantification area segmentation + deep learning
A2
B2
Synovitis area
Arteries area (eliminated)
A4
A3
B3
Scores of whole hands graded using the RAMRIS
Manual outlining
Dice coefficient + Spearman's rank correlation coefficient
Spearman's rank correlation coefficient

Figure 6

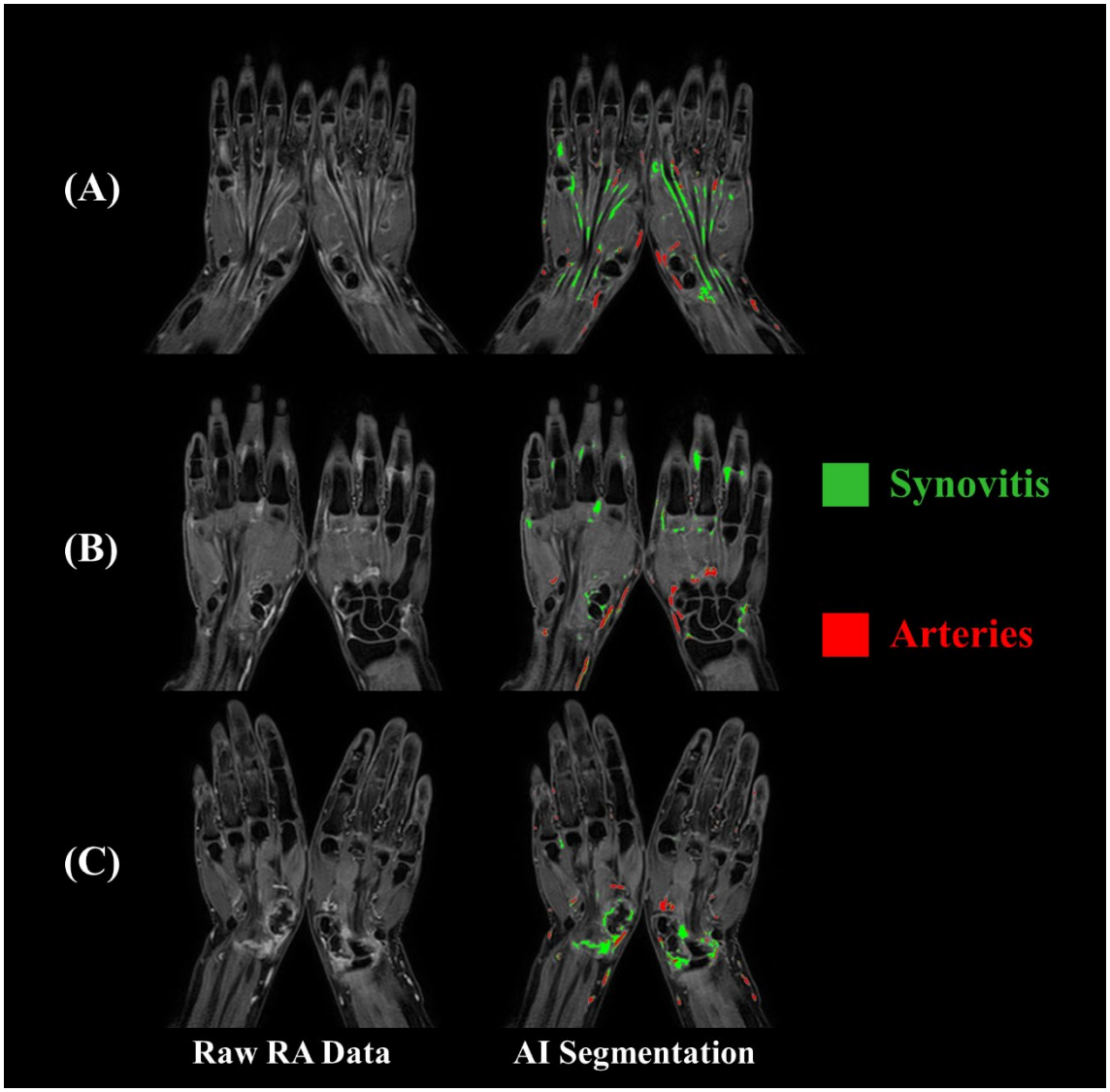


Figure 7

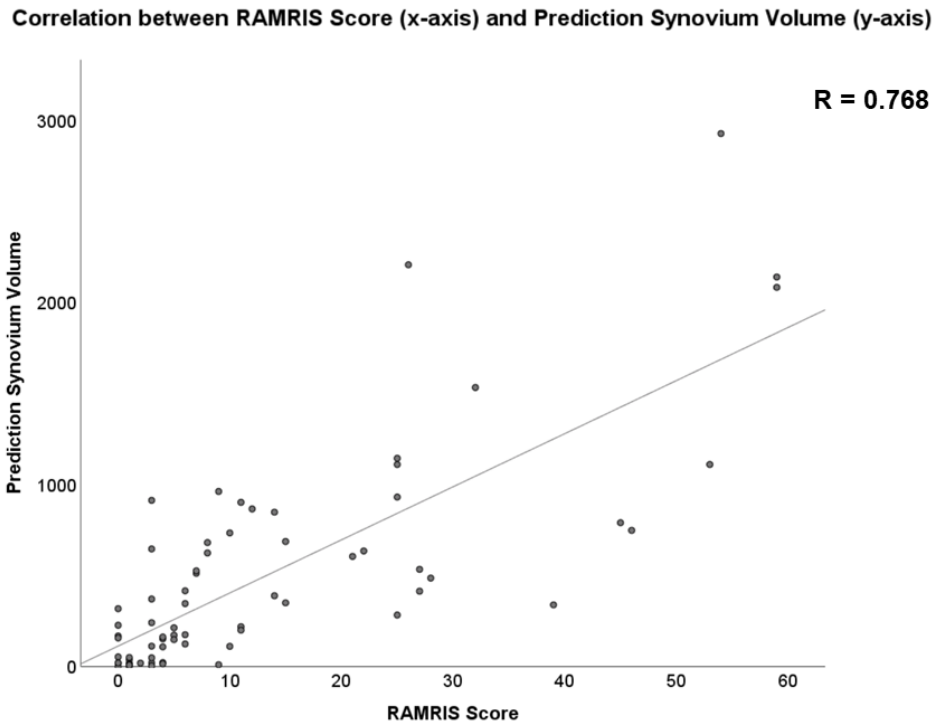
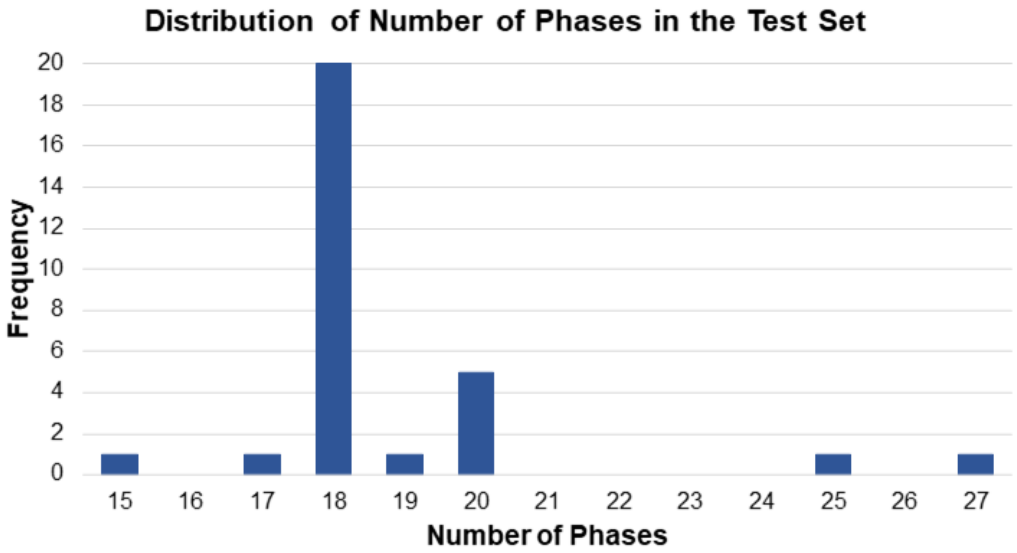

Correlation between RAMRIS Score (x-axis) and Prediction Synovium Volume (y-axis)
R = 0.768
Prediction Synovium Volume
3000
2000
1000
0
0
10
20
30
40
50
60
RAMRIS Score

Figure S1


Distribution of Number of Phases in the Test Set
Frequency
20
18
16
14
12
10
8
6
4
2
0
15
16
17
18
19
20
21
22
23
24
25
26
27
Number of Phases