Highlights

A Deep Registration Method for Accurate Quantification of Joint Space Narrowing Progression in Rheumatoid Arthritis

Haolin Wang,Yafei Ou,Wanxuan Fang,Prasoon Ambalathankandy,Naoto Goto,Gen Ota,Taichi Okino,Jun Fukae,Kenneth Sutherland,Masayuki Ikebe,Tamotsu Kamishima

- An intra-subject rigid registration network for bone displacement measurement.
- Achieved sub-pixel accuracy monitoring of joint space in rheumatoid arthritis.
- Significantly improved robustness for scaling, rotation, and noise.
- Misalignment visualization for reliability assessment.

A Deep Registration Method for Accurate Quantification of Joint Space Narrowing Progression in Rheumatoid Arthritis

Haolin Wang[a,**], Yafei Ou[b,c,*,**], Wanxuan Fang[a], Prasoon Ambalathankandy[b,c], Naoto Goto[b,c], Gen Ota[b,c], Taichi Okino[d], Jun Fukae[e], Kenneth Sutherland[f], Masayuki Ikebe[b,c] and Tamotsu Kamishima[g]

[a]*Graduate School of Health Sciences, Hokkaido University, Sapporo, 060-0812, Hokkaido, Japan*
[b]*Research Center For Integrated Quantum Electronics, Hokkaido University, Sapporo, 060-0813, Hokkaido, Japan*
[c]*Graduate School of Information Science and Technology, Hokkaido University, , Sapporo, 060-0813, Hokkaido, Japan*
[d]*Department of Radiological Technology, Sapporo City General Hospital, Sapporo, 060-8604, Hokkaido, Japan*
[e]*Kuriyama Red Cross Hospital, Yubari, 069-1513, Hokkaido, Japan*
[f]*Global Center for Biomedical Science and Engineering, Hokkaido University, Sapporo, 060-8638, Hokkaido, Japan*
[g]*Faculty of Health Sciences, Hokkaido University, Sapporo, 060-0812, Hokkaido, Japan*

ARTICLE INFO

Keywords:
Rheumatoid Arthritis
Joint Space Narrowing
Image Registration
Deep Learning
Radiology
Computer-aided Diagnosis

ABSTRACT

Rheumatoid arthritis (RA) is a chronic autoimmune inflammatory disease that leads to progressive articular destruction and severe disability. Joint space narrowing (JSN) has been regarded as an important indicator for RA progression and has received significant attention. Radiology plays a crucial role in the diagnosis and monitoring of RA through the assessment of joint space. A new framework for monitoring joint space by quantifying joint space narrowing (JSN) progression through image registration in radiographic images has emerged as a promising research direction. This framework offers the advantage of high accuracy; however, challenges still exist in reducing mismatches and improving reliability. In this work, we utilize a deep intra-subject rigid registration network to automatically quantify JSN progression in the early stages of RA. In our experiments, the mean-square error of the Euclidean distance between the moving and fixed images was 0.0031, the standard deviation was 0.0661 mm and the mismatching rate was 0.48%. Our method achieves sub-pixel level accuracy, surpassing manual measurements significantly. The proposed method is robust to noise, rotation and scaling of joints. Moreover, it provides misalignment visualization, which can assist radiologists and rheumatologists in assessing the reliability of quantification, exhibiting potential for future clinical applications. As a result, we are optimistic that our proposed method will make a significant contribution to the automatic quantification of JSN progression in RA. Code is available at https://github.com/pokeblow/Deep-Registration-QJSN-Finger.git.

1. Introduction

Rheumatoid arthritis (RA) is a chronic autoimmune inflammatory disease marked by joint swelling and tenderness, resulting in progressive articular destruction combined with severe disability. Joint space narrowing (JSN) caused by cartilage destruction can have a significant impact on functional status (Pfeil et al., 2013). Early diagnosis and treatment with disease-modifying antirheumatic drug (DMARD) therapy can prevent irreversible disability by halting RA before damage occurs to the joints, thereby avoiding or significantly slowing the progression of joint damage in 90% of patients (Platten et al., 2017). Therefore, using inexpensive, convenient, widely available imaging techniques with high sensitivity and specificity is essential for early diagnosis of RA and early intervention and management (Aletaha and Smolen, 2018).

Radiography has proven to be effective in identifying RA patients at a higher risk of further damage progression. Detecting early joint damage through radiography holds significant prognostic value both in clinical practice and clinical trials. The gold standard for assessing radiographic joint destruction in RA is the Sharp/van der Heijde scoring method (SvdH) (Van der Heijde, 2000) or the Genant-modified Sharp score (GSS) (Genant et al., 1998), in which radiologists and rheumatologists visually assess and score JSN progression and bone erosion of the hands, wrists and feet (Rydell et al., 2021). Rheumatologists traditionally measure JSN progression visually, relying on their training and expertise. However, even experienced rheumatologists face significant challenges in visually measuring JSN progression of 0.30 mm or less (Kato et al., 2019; Minh et al., 2022). Computer-aided diagnosis (CAD) has shown great potential to handle the challenges by quantitatively extracting joint space features in RA (Peloschek et al., 2007; Langs et al., 2008; Huo et al., 2015; Ou et al., 2023), which is widely used in clinical settings for rapid diagnosis (Hirano et al., 2019) or high-precision monitoring of RA progression (Okino et al., 2023).

1.1. Medical image analysis in joint space of RA

According to the nature of the methods and their output metrics, previous works on joint space quantification for

*Corresponding author
**Equal Contribution
oygrallen@gmail.com (Y. Ou)
ORCID(s): 0000-0002-5767-0019 (H. Wang); 0000-0001-7510-0813 (Y. Ou); 0000-0002-9855-5527 (P. Ambalathankandy); 0000-0002-6770-8260 (M. Ikebe); 0000-0002-7293-7471 (T. Kamishima)

Table 1
Feature comparison of mainstream frameworks of joint space quantification in RA.

Method	Output	Sensitivity	Detectable	Purpose
Margin detection	JSW	Medium	Early stage	Both
Classification	SvdH	Low	All stages	Qualitative
Registration	JSN	High	Early stage	Quantitative

RA can be divided into three frameworks: edge detection based joint space width (JSW) quantification, classification based SvdH scoring, and registration based JSN progression quantification. As shown in Table 2, each of these three frameworks have their advantages and disadvantages under different evaluation standards.

The margin detection based JSW quantification framework was the earliest computer-aided methodology in RA. It can be performed as follows: (i) Detect the bone margin by using a supervised machine learning (ML) network (Langs et al., 2008) or image features (Huo et al., 2015) such as intensity, gradient and derivative. (ii) Fit polynomial functions to bone margin curves. (iii) Quantify JSW according to the distance between polynomial functions. This framework has wide application. With the quantified absolute JSW, the SvdH score can be determined for qualitative diagnosis, while JSN progression (relative JSW) can be calculated for quantitative monitoring. However, this framework has some limitations: (i) Since this framework relies on margin information to determine the JSW, it is only suitable for use in the early stages of RA when there is a clear bone margin. (ii) This framework can only achieve pixel-level accuracy, limiting the sensitivity of joint space monitoring.

To overcome these limitations, a ML classification based SvdH scoring framework was proposed for rapid diagnosis at any stage of progression. In this framework, a supervised trained classifier, such as CNN (Hirano et al., 2019) or SVM (Nakatsu et al., 2020) is used to classify hand joint images from level 0 to level 4. This framework can be used in any stage of RA and can help rheumatologists make qualitative assessments. However, due to the inherent characteristics of classification, this framework is not suitable for quantitative analysis and has low sensitivity.

JSN progression is an important indicator for drug management in RA and has received widespread attention. However, over the course of one year, JSN progression can be less than one pixel, making it difficult to detect, as shown in Fig. 1. To quantitatively monitor JSN progression with high accuracy, a registration based JSN progression quantification framework was proposed (Ou et al., 2023, 2019). Taking the metacarpophalangeal (MCP) joint as an example, this method can be performed as follows: (i) Segment the proximal phalanx bone and metacarpal bone. (ii) Measure the displacements of the proximal phalanx bone and metacarpal bone between the baseline and follow-up finger joint images respectively by using an image registration algorithm. (iii) Calculate the displacement difference between the proximal phalanx bone and metacarpal bone to measure JSN progression.

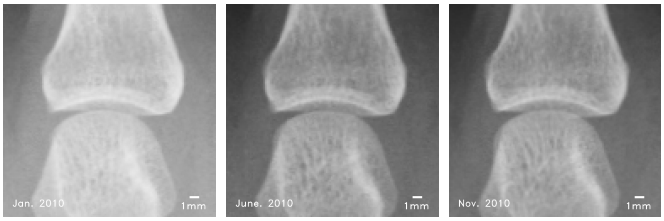

Figure 1: JSN progression of a MCP joint for the little finger over a period of 10 months. From left to right the images are baseline, five-month, and ten-month images (spatial resolution: 0.175 mm/pixel). JSN progression is usually less than one pixel per year, making it difficult for radiologists and rheumatologists to detect. Operating with an algorithm with pixel level accuracy to quantify JSN progression over a period of one year may not be possible. JSN progression measured at five and ten months relative to baseline using our method are -0.197 pixel and 0.174 pixel, respectively.

Compared to other frameworks, this framework has potential for higher sensitivity and lower mean error. However, it also has some limitations: (i) Changes in bone characteristics caused by bone erosion in advanced RA can reduce the accuracy of the rigid registration algorithm and even cause mismatching. For the above reason, this framework is mostly used in the early stages of RA. (ii) Considering that the registration algorithm can only provide JSN progression, this limits its application for qualitative diagnosis.

1.2. ML-based image registration in medical image analysis

Registration in medical image processing refers to the process of aligning multiple medical images on a common coordinate system. It is an important step in many medical image analysis tasks (Chen et al., 2022). ML-based medical image registration algorithms have recently gained popularity (Chen et al., 2021). Registration algorithms can be categorized into rigid and non-rigid (deformable) registration based on the deformation model type. Non-rigid registration has broad applications in tissues or organs such as the brain, thorax, lung (Fu et al., 2020), and can be used to detect lung motion (Ehrhardt et al., 2010) and tumor regression (Neylon et al., 2017). For rigid registration, the transformation parameters are obtained based on the convolutional neural network to achieve image registration. A 2D/3D target registration for X-ray images using convolutional neural networks was presented to obtain transformation parameters (Miao et al., 2016). In our work, inspired by these advancements, we focus on the overall displacement of bones, enabling the rigid registration algorithm to achieve higher quantification accuracy. By utilizing the potential of deep learning, the problems of traditional computation-based image registration methods mentioned in the previous subsection can be addressed. This work is limited to detecting JSN progression of the same patient over consecutive time points (Stoel, 2020), that is, intra-subject registration.

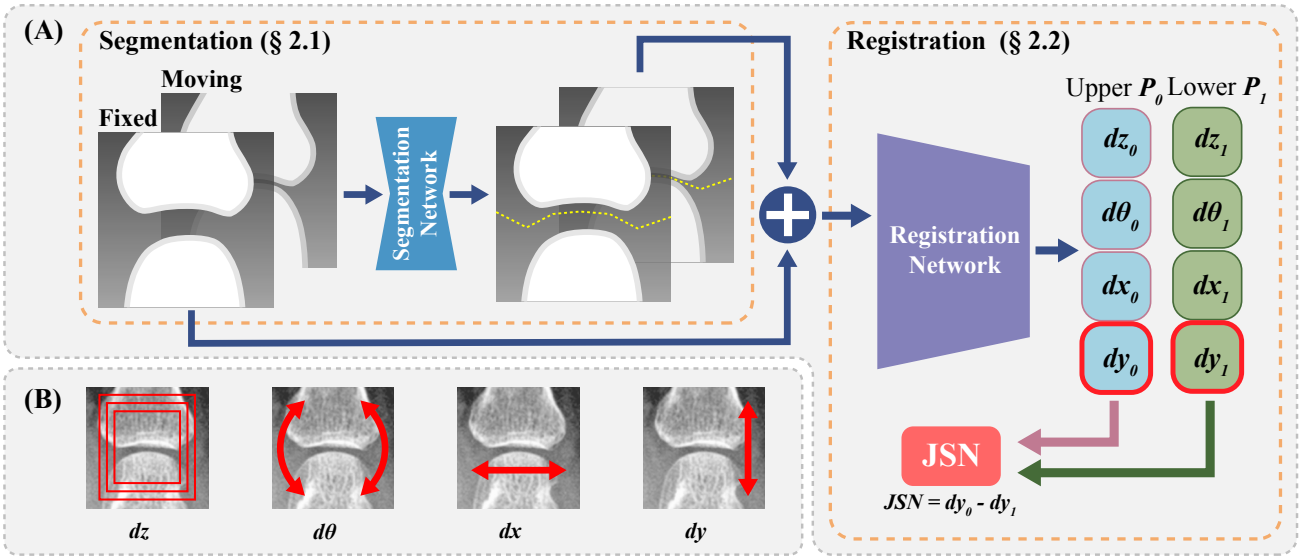

Figure 2: (A) Overview of our proposed deep learning image registration based JSN progression quantification methodology. This work can be divided into two steps: joint segmentation and JSN progression quantization. Using a MCP joint as an example, this work can be performed as follow: (i) A supervised U-net++ based network is implemented to segment the proximal phalanx bone and metacarpal bone region of the MCP joint (§ 2.1). (ii) An un-supervised ResNet-like based deep registration network is proposed to quantify the rigid transformation parameters of the proximal phalanx bone and metacarpal bone region (§ 2.2). (iii) The JSN progression can be obtained by calculating the displacement difference of the y-axis between the two bones. (B) Four rigid transformation parameters are shown that are used in this work; dz: scaling, $d\theta$: rotation, dx: displacement on the x-axis, dy: displacement on the y-axis.

1.3. Our contributions

In this work, a deep learning-based methodology is proposed for JSN quantification. The following are our original contribution:

1. Implementation of an image segmentation network based on U-net++ to accurately segment joint images.
2. Proposal of a ResNet-like deep registration network to measure bone displacement.
3. Achievement of sub-pixel accuracy in monitoring joint space during the early stage of RA.
4. Significant improvement in robustness for scaling, rotation, and noise compared to related works. This reduces mismatching caused by inconsistent angles between the upper and lower bones of the joint, variable spatial resolutions of radiography images, and inconsistent projection angles.
5. Provision of a misalignment visualization that enables radiologists and rheumatologists to assess the reliability of quantification. This feature has important implications for the future clinical application of our method.

The rest of this paper is organized as follows. § 2 describes the implementation of our methodology; including joint segmentation network and JSN progression quantification network and introduces the clinical datasets used in this work. § 3 presents and discusses the segmentation and registration results. § 4 concludes this work and discusses possible future research directions for computer-aided monitoring of RA.

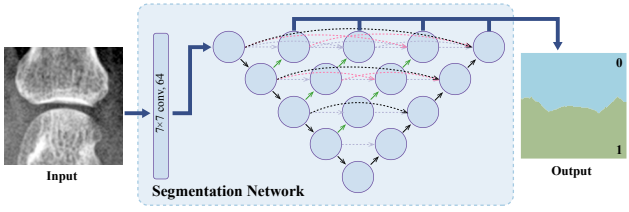

Figure 3: Diagram of our segmentation network. The segmentation network contains one convolutional layer (kernel size: 7×7, channels: 64) and a 5-layer Unet++ network.

2. Methodology and materials

In this work, a deep learning based JSN quantification method is proposed. The proposed method aims to improve the sensitivity, accuracy and robustness of JSN progression monitoring in the early stages of RA. As shown in Fig. 2, the proposed work contains two networks: a U-net++ based joint segmentation network and a ResNet-like deep registration network for JSN progression quantification.

2.1. Joint segmentation

A network based on U-net++ with an added convolution layer is proposed for joint segmentation, as illustrated in Fig. 3. Consider a MCP joint as an example, the proximal phalanx bone and metacarpal bone are segmented separately using the U-net++ network. This enables the individual measurement of displacement for the upper part (proximal phalanx bone) and lower part (metacarpal bone) of the

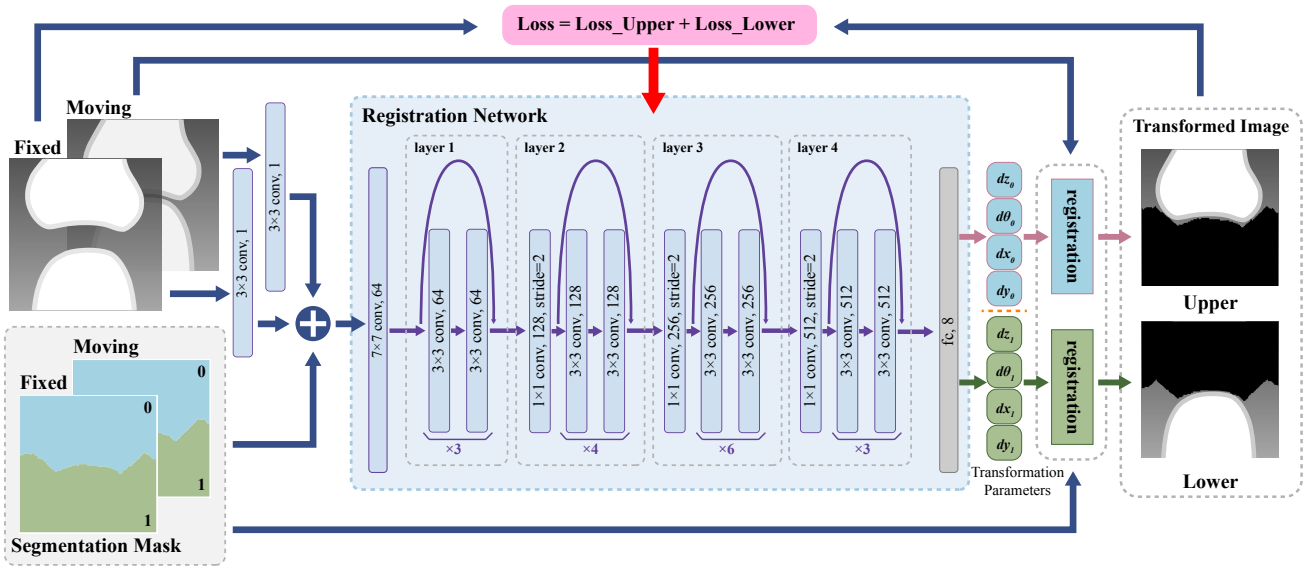

Figure 4: Structure diagram of image registration network. In this case, after a convolution layer of 3×3 convolution kernels and a 1-channel convolution base and combined with its corresponding segmentation mask, the input image set is input into the registration network. The registration network contains a layer of convolution and 4 layers of residual convolution modules with 64, 128, 256 and 512 channels, respectively. The final transformation parameters are obtained after a fully connection layer with 8 output values. These transformation parameters are used to deform the moving image to generate a transformed image. The difference between the generated transformed image and the fixed image is defined as the loss, which is used to optimize the registration network.

joint. The output of the segmentation network is defined as S, where 0 represents the metacarpal bone region (upper region) and 1 represents the proximal phalanx bone region (lower region).

2.2. JSN quantification by image registration

In this subsection, an unsupervised intra-subject rigid registration network is proposed for quantifying JSN progression. The pipeline can be explained as follows: (i) Quantify the transformation parameters between the baseline and follow-up radiographic images using the registration network, as shown in Fig. 4. (ii) Calculate JSN progression based on the vertical displacement difference between the upper and lower regions of the joint.

2.2.1. Transformation Parameterization

A 3D rigid transformation can be parametrized by three in-plane and one out-of-plane transformation parameters (Kaiser et al., 2014), as shown in Fig. 2 (A). The in-plane transformation parameters include two displacement parameters dx, dy and one rotation parameter $d\theta$. The out-of-plane transformation parameter is the scaling parameter dz. In our registration network, the upper and lower regions are registered simultaneously. Two sets of parameters are introduced, namely: $\{P_0 \mid dz_0, d\theta_0, dx_0, dy_0\}$ for the upper region, and $\{P_1 \mid dz_1, d\theta_1, dx_1, dy_1\}$ for the lower region. The vertical displacement parameter dy of P_0 and P_1 is used to calculate the JSN progression.

2.2.2. Registration network

The rigid transformation parameters of the upper and lower regions are obtained by simultaneously registering both regions. The detailed operation is described as follows.

Given a fixed joint image F and a moving joint image G, they can be divided into upper and lower regions. The segmentation mask is denoted as S, where 0 represents the upper bone region and 1 represents the lower bone region.

The transformation matrix of the upper region, denoted as t_0, and the lower region, denoted as t_1, are generated based on the parameter sets P_0 or P_1 obtained through the proposed network, which is defined in Eq. 1.

$$t = \begin{pmatrix} dzcosd\theta & -dzsind\theta & dxdzcosd\theta - dydzsind\theta \\ dzsind\theta & dzcosd\theta & dxdzsind\theta + dydzcosd\theta \\ 0 & 0 & 1 \end{pmatrix} \quad (1)$$

The transformation matrix t_0 and t_1 is subsequently applied for the transformation function T of the moving image ðlŘž to transform the spatial position of each pixel and generate the transformed image G_0, G_1, which can be defined as shown in Eq. 2:

$$G_0 = T\left(G, t_0\right) \quad G_1 = T\left(G, t_1\right) \quad (2)$$

Thus, the upper region F_0 and lower region F_1 of the fixed image, the upper region G_0' and lower region G_1' of the transformed image can be expressed as follows:

$$\begin{aligned} F_0 &= F * \neg S \quad F_1 = F * S \\ G_0' &= G_0 * \neg S \quad G_1' = G_1 * S \end{aligned} \quad (3)$$

The transformed image G' can then be obtained by combining its upper region G'_0 and lower region G'_1, as shown in Eq. 4.

$$G' = G'_0 + G'_1 \tag{4}$$

In this registration network, the mean squared error (MSE) of the Euclidean distance is used as the loss, as defined in Eq. 5. m and n denote the width and height of $F(x, y)$, respectively.

$$L(F, G) = \sqrt{\frac{1}{m \times n} \sum_{y=1}^{n} \sum_{x=1}^{m} (F(x, y) - G(x, y))^2} \tag{5}$$

Here, $L(F, G)$ represents the Euclidean distance loss between image F and image G. The loss in our registration network includes both the upper and lower parts. For example, given a fixed image F and a transformed image G', the transformed loss $L(F, G')$ can be defined as follows:

$$L(F, G') = \alpha \times L(F_0, G'_0) + \beta \times L(F_1, G'_1) \tag{6}$$

Here, α and β represent the weights used to balance the loss of the upper and lower parts of the joint. They are set to $\alpha = \beta = 0.5$. Then, the original loss $L(F, G)$ of the moving image G can be similarly calculated.

The transformation parameters P_0 and P_1 obtained from registering the upper and lower joint regions are used to generate the final results. The vertical displacements dy_0 and dy_1 are used to calculate JSN progression. Thus, the joint space difference between the fixed image F and the moving image G can be described as follows:

$$\mathrm{JSN}_{fg} = dy_0 - dy_1 \tag{7}$$

2.2.3. Network architecture

The architecture of the proposed registration network is illustrated in Fig. 4. This network is based on a residual convolution module (He et al., 2016). It is implemented to obtain transformation parameters between the fixed image and the moving image. The network takes the moving image, fixed image and their segmentation masks as input, which are combined as four channels. To reduce noise interference in the radiographic images of both fixed and moving images, a single-layer convolution is applied before entering the registration network. The network's layer information and the number of parameters are illustrated in Table. 2. After feature extraction by the registration network, eight output parameters are obtained through a fully connected layer. These parameters include two sets of registration parameters for the upper and lower bone regions of the joint, which are the outputs of the network. These two sets of transformation parameters are then used to transform the moving image and its segmentation mask according to Eq. 2 and Eq. 3, respectively, resulting in the transformed image. During the training stage of the network, the loss function is the mean squared error (MSE) of the distance between the fixed image and the transformed image, as defined in Eq. 5.

Table 2
Layer types, output shapes, and the number of parameters of the proposed ResNet-like network.

Layers	Output shape	Number of parameters
Conv2d+BN	(1, 224, 224)	10+2
Conv2d+BN	(1, 224, 224)	10+2
Conv2d+BN	(64, 112, 112)	12,544+128
ReLU	(64, 112, 112)	0
MaxPool2d	(64, 56, 56)	0
Conv2d+BN	(128, 56, 56)	73,728+256
ResidualBlock×3	(128, 56, 56)	747,008
Conv2d+BN	(256, 28, 28)	294,912+512
ResidualBlock×4	(256, 28, 28)	4,165,632
Conv2d+BN	(512, 14, 14)	1,179,648+1024
ResidualBlock×6	(512, 14, 14)	26,095,616
Conv2d+BN	(512, 7, 7)	2,359,296+1024
ResidualBlock×3	(512, 7, 7)	12,064,768
Fully Connected	(1, 8)	4,104
Tanh	(1, 2)	0
Tanh	(1, 2)	0
Linear	(1, 4)	0
Total	-	**47,000,224**

ResidualBlock: A Residual Block consists of the following layers: Conv2d, BatchNorm2d, ReLU, Conv2d, BatchNorm2d, ReLU **BN**: BatchNorm2d

2.3. Implementation

The joint segmentation and registration networks were trained and tested separately. The networks were implemented using Python language and PyTorch package (Paszke et al., 2019) on a workstation with a single GPU (NVIDIA GeForce GTX 2080 Ti). The implementation details of the networks are described as follows.

2.3.1. Segmentation network

The loss function of the segmentation network consisted of sigmoid and binary cross entropy loss (BCELoss). It was optimized using the RMSProp optimizer (Hinton et al., 2012), with ρ set to 0.9 and epsilon set to 0.0001. The initial learning rate was 0.00001. Training was carried out over 150 epochs with a batch size of 30. The network was trained three times. The results were the averaged to reduce the impact of random initialization.

2.3.2. Registration network

The registration network was trained using the Adam optimizer (Kingma and Ba, 2014) , with an initial learning rate of 0.001. ReduceLROnPlateau was used for adaptive learning rate reduction. The training process consisted of 500 epochs, with a batch size of 80. Furthermore, to minimize the impact of random initialization during training, the proposed registration network was trained three times. The result was obtained by averaging the three runs.

2.4. Dataset

To evaluate the performance of our network, we prepared a clinical dataset in compliance with the guidelines

Table 3
Patient information in the clinical dataset

	Mean ± SD	Range
Age at enrollment (year)	56.11 ± 13.79	20.68 ∼ 88.00
Amount of radiography	4.30 ± 2.54	3 ∼ 17
Follow-up period (year)	4.04 ± 3.44	0.88 ∼ 12.10

Table 4
Configuration parameters for radiographic imaging

	SARC	HMCRD	SCGH
Model	DR-155HS2-5	Radnext 32	KXO-50G
Manufacturer	Hitachi	Hitachi	Toshiba
Aluminum filter (mm)	1.5	0.5	NO
Tube voltage (kV)	42	50	45
Tube current (mA)	100	100	250
Exposure time (mSec)	20	25	14
Source to image (cm)	100	100	100
Resolution (mm/pixel)	0.175	0.15	0.15
Image size (pixel)	2010×1490	2010×1490	2010×1490
Bit depth (bit)	12	10	10

SARC: Sagawa Akira Rheumatology Clinic.
HMCRD: Hokkaido Medical Center for Rheumatic Diseases.
SCGH: Sapporo City General Hospital.

of the Declaration of Helsinki and obtained approval from the Ethics Committee of the Faculty of Health Sciences, Hokkaido University (approval number: 19 - 46). The dataset used in this study consisted of 675 hand posteroanterior projection (PA) radiographs from 80 patients with rheumatoid arthritis (RA). Among these patients, 88.5% were female. Detailed patient information is summarized in Table 3. The images were obtained from three different institutions: Sagawa Akira Rheumatology Clinic (Sapporo, Japan), Hokkaido Medical Center for Rheumatic Diseases (Sapporo, Japan), and Sapporo City General Hospital (Sapporo, Japan). Each institution has its own X-ray systems, and the dataset is managed using the digital imaging and communications in medicine (DICOM) standard. For detailed information about the imaging parameters, please refer to Table 4.

To extract finger joint images from the hand images, we utilized the finger joint detection method described in (Ou et al., 2022). The finger joints were scored by a radiologist with extensive training. Only early-stage RA cases with a SvdH hand score of 0 were included in the dataset. This selection criterion was applied because narrowed joint space can impact segmentation accuracy, and bone erosion can cause damage to the bone margin in advanced RA cases. Additionally, we retained only the images from patients who underwent hand radiography at least three times to enable the calculation of standard deviation. The distribution of data for different joints is presented in Table 5.

For the segmentation task, we divided the dataset into an 80% training set, consisting of 4,854 finger joint images with corresponding expert annotations, and a 20% testing set, consisting of 1,213 finger joint images. For the registration task, we needed paired images of fixed and moving joints. To construct the dataset, we considered each joint with multiple images. We created pairs by pairing up images smaller than the middle index with images larger than the middle index within each joint image set. The resulting dataset included 1,597 finger joint image pairs for training and 3,604 finger joint image pairs for testing.

Table 5
The amount of finger joint images in our clinical dataset

	IP	PIP	MCP	OverAll
Thumb	561	N/A	569	
Index	N/A	636	672	
Middle	N/A	647	599	
Ring	N/A	514	626	
Small	N/A	560	683	
Overall	**561**	**2357**	**3149**	**6067**

IP: Interphalangeal joint.
PIP: Proximal interphalangeal joint.
MCP: Metacarpophalangeal joint.

In the clinical assessment, we included 15 patients with rheumatoid arthritis from Hokkaido Medical Center for Rheumatic Diseases. Each patient had baseline and 52-week follow-up radiographic images of both hands. The dataset evaluated the five metacarpophalangeal (MCP) joints and four proximal interphalangeal (PIP) joints of each hand using the Genant-modified Sharp score (GSS) (Genant et al., 1998).

3. Experiments and discussion

3.1. Segmentation experiments

3.1.1. Segmentation evaluation

The segmentation performance of U-net++ in this work was quantitatively evaluated using manual annotations as the ground truth. The evaluation employed the following six metrics (Zhang, 1996):

- **Mean Intersection over Union (mIoU)**: Measures the ratio of the intersection to the union of the two sets of ground truth and predicted results.
- **Sensitivity (SEN)**: Represents the percentage of ground truth regions that are correctly segmented.
- **Specificity (SPC)**: Indicates the percentage of non-ground truth regions that are correctly segmented.
- **Dice Similarity Coefficient (DSC)**: Quantifies the similarity between the prediction and the ground truth.
- **Accuracy (ACC)**: Measures the percentage of correctly predicted pixels to the total number of pixels.

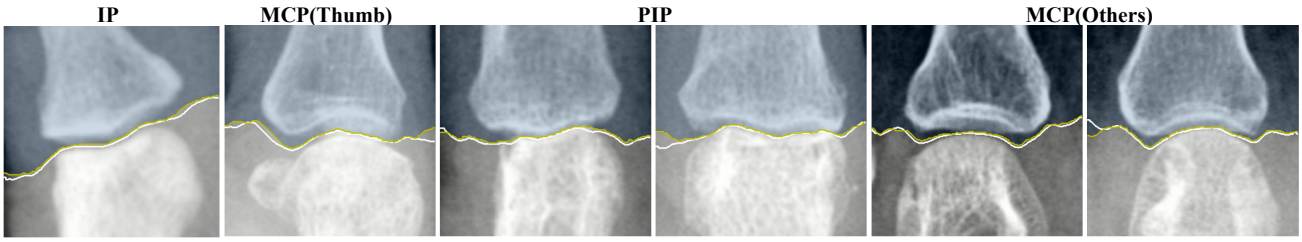

Figure 5: Experiments of the proposed segmentation network. White lines represent the manual label of segmentation, and yellow lines represent the predicted segmentation by using the network.

Table 6
The performance in different evaluation metrics

	mIoU	SEN	SPC	DSC	ACC
IP	0.95900	0.97782	0.98292	0.97896	0.98028
PIP	0.96227	0.97115	0.99191	0.98056	0.98166
MCP	0.95396	0.96422	0.99094	0.97615	0.97821
Overall	**0.95779**	**0.96835**	**0.99052**	**0.97819**	**0.97980**

3.1.2. Segmentation results

Different evaluation metrics were employed to assess the IP, PIP, and MCP joints individually. The results, shown in Fig. 5 and Table 6 indicate that our automatic segmentation method yields accurate and sensitive results compared to the ground truth.

3.2. Registration experiments

3.2.1. Registration evaluation

In radiographic images, noise can greatly affect visual measurements (Ou et al., 2023). Phantom experiments have shown that manually annotated data can have sub-pixel mean errors, leading to sub-pixel deviations in algorithm evaluation. Hence, metrics that do not rely on manually annotated data are used for evaluating algorithm performance. The experiments conducted on our registration network are summarized in Table 7 using four metrics: the standard deviation σ as defined in (Ou et al., 2019), the standard deviation σ' as defined in (Langs et al., 2008), the mismatching ratio, and the transformed loss.

- **Standard deviation** σ: Evaluates the accuracy of JSN progression quantification.
- **Standard deviation** σ': Represents the percentage of mismatching cases and measures the robustness of the algorithm.
- **Mismatching ratio**: Represents the percentage of mismatching cases and measures the robustness of the algorithm.
- **transformed loss**: Quantifies the difference between the transformed image and the fixed image, providing an assessment of the registration network's performance.

The standard deviation σ is computed to assess the reliability of the registration network without the ground truth (Ou et al., 2023). The standard deviation represents the variation among multiple measurements. For example, when comparing images F and G, the JSN_{FG-I} between image F and image G can be indirectly calculated by introducing intermediate image I, as follows:

$$JSN_{FG-I} = JSN_{FI} + JSN_{IG} \tag{8}$$

Considering a set of images, the $\overline{JSN_{fg}}$ can be obtained by taking the average of multiple measurements.

$$\overline{JSN_{FG}} = \frac{1}{n}\sum_{I=1}^{n} JSN_{FG-I} \tag{9}$$

Therefore, the standard deviation σ_{fg} of JSN_{fg} can be defined as follows:

$$\sigma_{FG} = \sqrt{\frac{1}{n}\left(\sum_{I=1}^{n} JSN_{FG-I} - \overline{JSN_{FG}}\right)^2} \tag{10}$$

We utilize the standard deviation introduced in (Langs et al., 2008) to demonstrate the reliability of our method. In the case of two images F and G, we create image G_j by adding a random translation (-3, +3 pixels) to image G along the x-axis and y-axis. The standard deviation σ'_{FG} of images F and G can be defined by using the JSN_{FG-j} between image F and image G_i.

$$\overline{JSN_{FG}} = \frac{1}{10}\sum_{j=1}^{10} JSN_{FG-j} \tag{11}$$

$$\sigma'_{FG} = \sqrt{\frac{1}{10}\left(\sum_{j=1}^{10} JSN_{FG-j} - \overline{JSN_{FG}}\right)^2} \tag{12}$$

When calculating the standard deviation σ'_{FG}, outliers will be removed as mismatches.

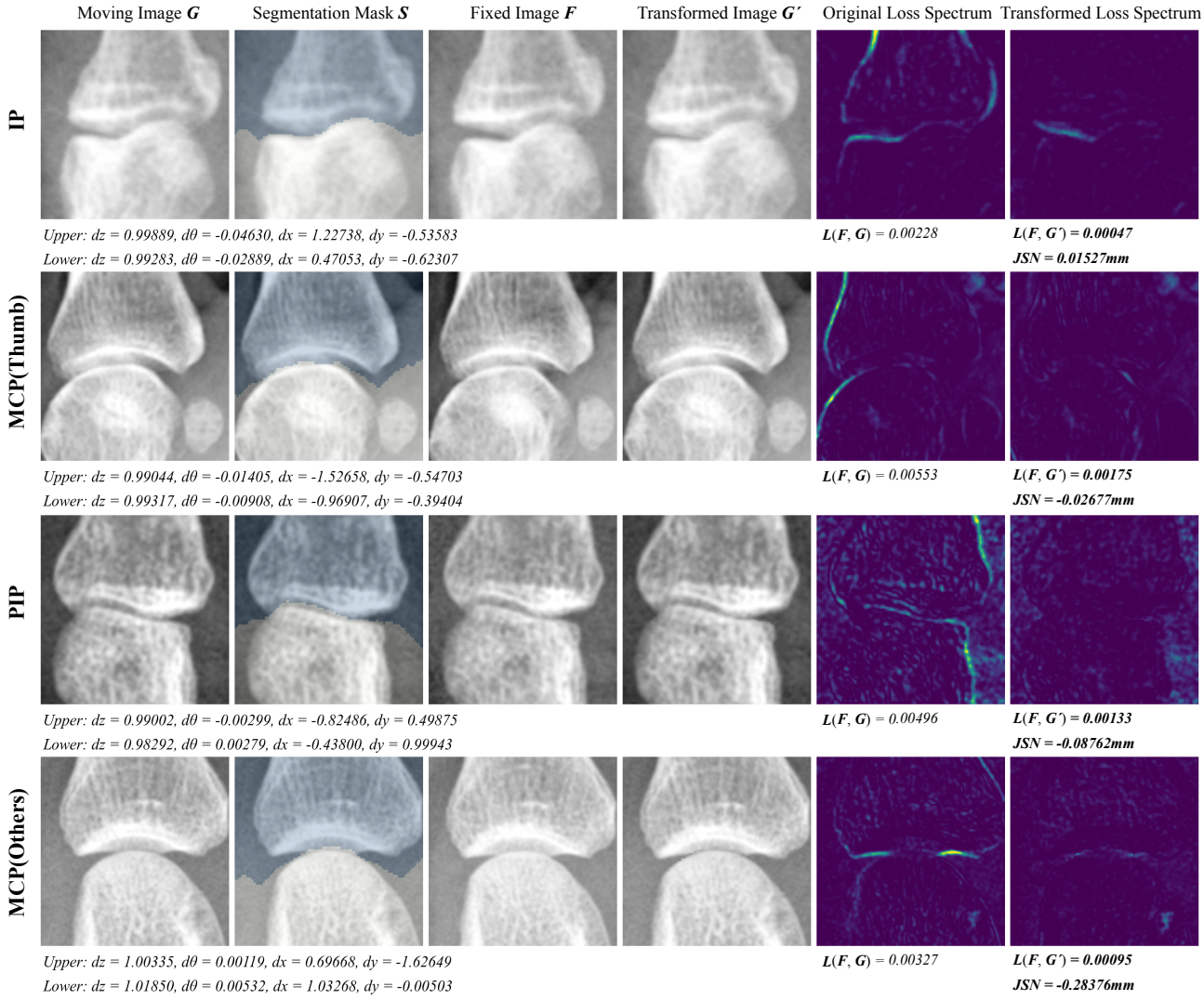

Figure 6: Experiments of the proposed registration network. Considering that the MCP joint of the thumb is different from the other fingers, we have divided the finger joints into four categories; IP, MCP (thumb), PIP and MCP (others). The first three columns on the left are inputs, the moving image G, the segmentation mask S and the fixed image F. The fourth column is the transformed image G'. The two columns on the right are the original Euclidean distance loss spectrum E_{FG} and the transformed loss spectrum $E_{FG'}$. The quantified transformation parameters P_0, P_1, the original loss $L(F, G)$ and the transformed loss $L(F, G')$ are listed below the images.

3.2.2. Registration results

The experimental results of our registration network for various finger joints are shown in Fig. 6. The finger joints are divided into four categories: IP, MCP (thumb), PIP, and MCP (others). In the original loss spectrums and the transformed loss spectrums, the highlighted regions represent relative displacement between fixed image and moving/transformed image. As shown in Fig. 6, the proposed registration network can effectively reduce the highlight regions, especially around the bone margin. This is significant as the bone margin information is crucial for determining JSW or JSN progression in clinical data, rather than the bony texture information, which can vary due to changes in bone thickness or imaging angles. Therefore, the loss around bone margin region is more important. These irregular variations on bony texture are also the primary cause of loss in the rigid registration network, as shown in Fig. 7. This demonstrates that the proposed registration network can accurately quantify the transformation parameters between the fixed image and moving image.

Our experiments show that the standard deviation σ, the mismatching ratio, and the mean original loss of IP joint images are much higher than other joints. There is substantial evidence that thumb movements are more independent compared to other fingers (Ingram et al., 2008). As a result, the IP joint exhibits distinct characteristics when the hand posture is altered. Inconsistent hand posture is primarily shown radiographically as rotation or scaling in PIP and MCP joints. However, the rotation of the thumb MCP joint can lead to rolling in the thumb, thereby altering

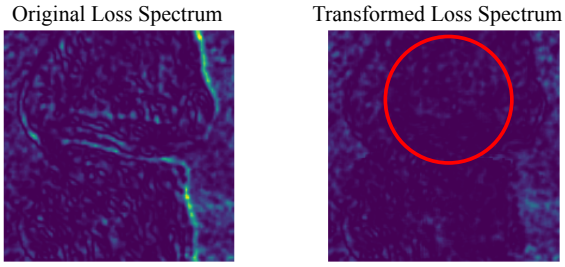

Figure 7: The bony texture varies as the bone thickness/diameter varies (in response to muscle activity or weight) or due to any changes in the imaging angle. Therefore, it is difficult to reduce the loss on the bone surface region to 0, as shown in the highlighted region. These irregular variations of bony texture are also a major part of the transformed loss.

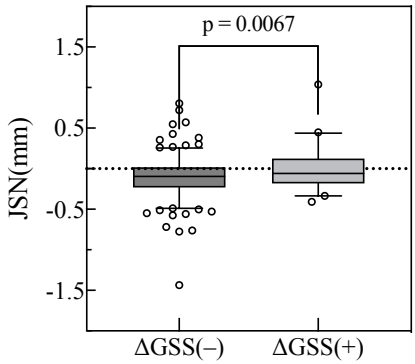

Figure 8: Comparison of the JSN progression derived from the proposed method in terms of radiographic JSN progression. JSN(mm), the difference in joint space width at baseline and the 52nd week. ΔGSS, the difference in GSS at baseline and the 52nd week.

the projection angle of the IP joint. This can significantly affect the IP joint characteristics, leading to decreased quantification accuracy and potential mismatches. Improving the robustness of algorithms and reducing mismatches in the IP joints remains a challenge.

Figure 9 demonstrates the distribution and the relationship between original loss and transformed loss in various joint images. Due to the differing characteristics of various joints, the distribution varies. As mentioned earlier, the mean and variance of the original loss and the transformed loss for IP joint images are higher compared to others. In all three kinds of finger joint images, our proposed registration network can effectively control the loss. In 94.6% of the registration cases, the transformed loss is less than half compared to the original loss.

It is important to note that the transformed loss is difficult to decrease infinitely in rigid registration due to variations in bone features, including bony texture, margin information from finger bending, and bone erosion, across multiple radiographic images. These variations pose significant obstacles to achieving successful rigid registration. Moreover, experiments involving these factors typically yield high original loss, leading to a higher occurrence of mismatch cases in the high original loss region. In contrast, our proposed method, as shown in Fig.9, demonstrates high robustness and a low mismatching ratio in this region. Table 7 further highlights the effectiveness of our method in controlling the mismatching ratio of various finger joint images, particularly PIP and MCP joint images.

3.2.3. Clinical assessment

We conducted a statistical analysis to compare the results obtained by our proposed method with the results of the GSS (Genant et al., 1998) completed by a rheumatologist. This analysis was performed on a clinical dataset consisting of 135 joints from 15 patients with RA. The dataset included images of five MCP joints and four PIP joints. We analyzed the differences and correlations separately.

For each patient, we obtained the GSS values for the baseline and the follow-up at the 52nd week. The difference between these values was calculated and referred to as ΔGSS. We categorized the joints based on the sign of ΔGSS: ΔGSS(+) for joints with a positive ΔGSS according to the GSS results (n = 41, 15.19%), and ΔGSS(-) for the remaining joints (n = 229, 84.81%). Additionally, we measured the joint space narrowing (JSN) at the baseline and follow-up using our proposed method.

Firstly, we found that the JSN of the finger joints with GSS progression [ΔGSS(+)] was significantly greater than the JSN of the joints without GSS progression [ΔGSS(-)] (Unpaired T test, p = 0.0067 (one-tail), 95% confidence interval: 0.02143 to 0.1831). The mean JSN for ΔGSS(-) and ΔGSS(+) was -0.1079 mm and -0.005626 mm, respectively, as shown in Fig. 8. The presence of individual outlier points in the figure can be attributed to incorrect joint detection results and cases of excessive narrowing or absence of joint space.

Secondly, we observed a significant correlation between the JSN of the finger joints measured by our proposed method and the progression of ΔGSS (Pearson Correlation, p = 0.0003, r = 0.2167).

3.2.4. Comparison with related works

During clinical radiographic imaging, inconsistent hand posture of the patient or different imaging equipment can result in rotation or scaling of the bones. As the RA progresses, bone erosion gradually destroys the margin information of bones. These changes in bone margin information due to rotation, scaling, or bone erosion pose a challenge for rigid image registration based JSN progression quantification in RA.

Based on experiments of PIPOC-based JSN progression quantification in (Ou et al., 2023), PIPOC is susceptible to noise, rotation, and scaling, making inconsistent hand posture a major reason for mismatches. This inconsistency can be broadly divided into two cases. (i) The inconsistent angle between the upper and lower bones of the joint, as show in the first row of Fig. 10. This inconsistent joint angle is shown radiographically as a rotation. (ii) The bending of

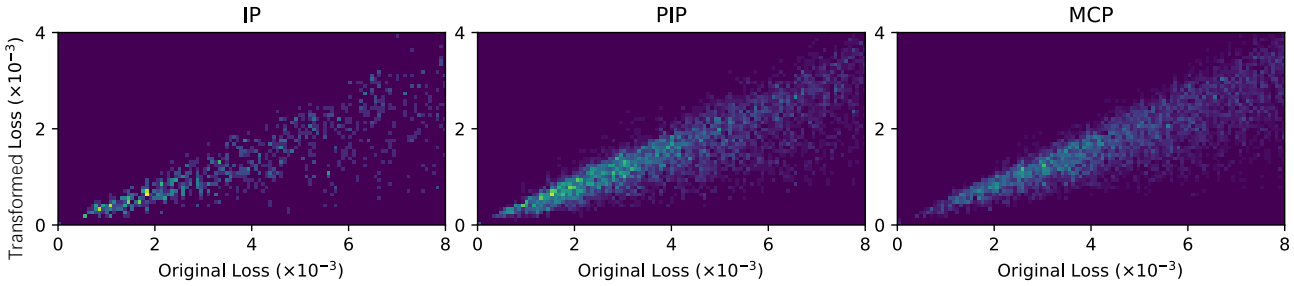

Figure 9: Heat maps of the original loss $L(F, G)$ and the transformed loss $L(F, G')$. This set of heat maps demonstrate the distribution and relationship between the original loss and the transformed loss.

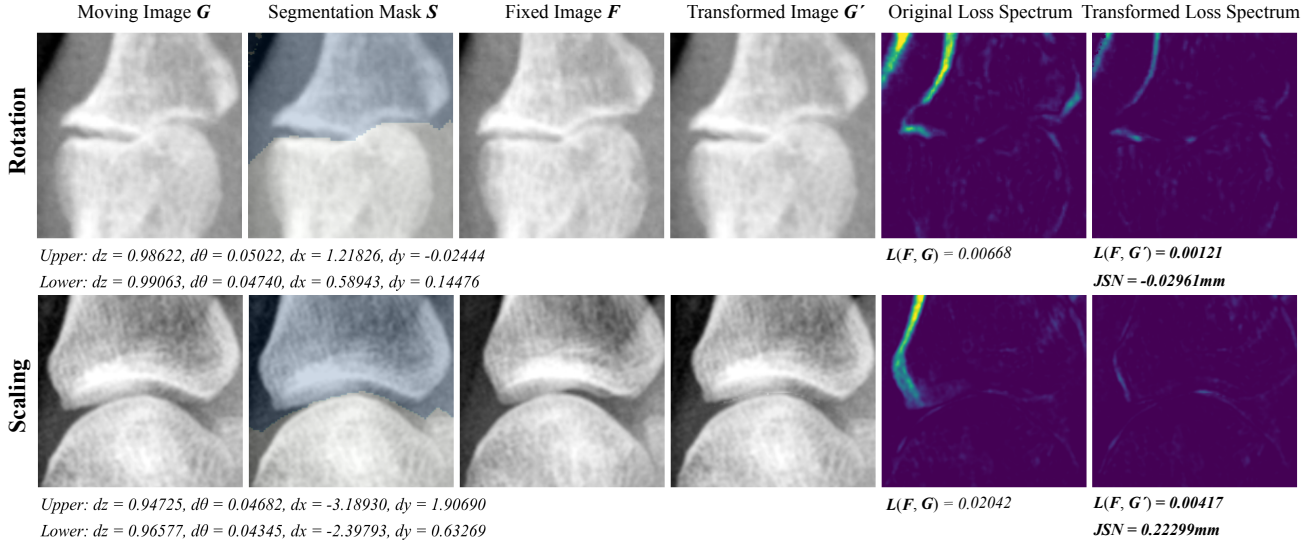

Figure 10: Experiments of the proposed registration network in specific cases. Those figures show the robustness of the proposed registration network to rotation and scaling.

the fingers, as show in the second row of Fig. 10. In this case, there will be obvious scale differences between the upper and lower bones. In the paper (Ou et al., 2023), it is reported that due to the characteristics and limitation of PIPOC, the inconsistency of the angles or scales of upper and lower bones can easily cause mismatches. As shown in the transformed loss spectrums of Fig. 10, the registration network proposed in this work can improve the robustness for rotation and scaling, and it can accurately quantify the angle and scale difference. This improvement can effectively reduce the mismatching ratios when compared to PIPOC, as shown in Table 8.

Table 8 summarizes the comparison with other joint space quantification work in RA. We can observe that the image registration-based JSN progression quantification framework can achieve lower standard deviation σ' when compared to the margin detection-based JSW quantification framework. This shows that the image registration-based

Table 7
Mean standard deviation in millimeters, the mismatching ratios and the transformed Euclidean distance loss for our image registration network.

	Standard Deviation σ			Standard Deviation σ'			Mismatching Ratio (%)			Transformed Loss		
	IP	PIP	MCP	IP	PIP	MCP	IP	PIP	MCP	IP	PIP	MCP
Thumb	0.0864	N/A	0.0688	0.0415	N/A	0.0394	3.31	N/A	0.53	0.0049	N/A	0.0031
Index	N/A	0.0496	0.0689	N/A	0.0354	0.0373	N/A	1.41	1.07	N/A	0.0031	0.0029
Middle	N/A	0.0491	0.0571	N/A	0.0400	0.0331	N/A	0.65	0.01	N/A	0.0021	0.0028
Ring	N/A	0.0423	0.0610	N/A	0.0358	0.0313	N/A	0.14	0.44	N/A	0.0020	0.0031
Small	N/A	0.0410	0.0745	N/A	0.0294	0.0312	N/A	0.54	0.34	N/A	0.0023	0.0035
Overall	**0.0864**	**0.0455**	**0.0661**	**0.0415**	**0.0351**	**0.0345**	**3.31**	**0.68**	**0.48**	**0.0049**	**0.0024**	**0.0031**

Table 8
Comparison with related works in algorithm accuracy. The mean standard deviation in millimeter and the mismatching ratios for respective joints.

	Method	Dataset (Images)	Resolution (mm/pixel)	Output Metric	Standard Deviation σ				Standard Deviation σ'				Mismatching Ratio (%)			
					IP	PIP	MCP	OverAll	IP	PIP	MCP	OverAll	IP	PIP	MCP	OverAll
Langs et al. TMI'08	ASM	160MCP	0.0846	JSW	-	-	-	-	-	-	0.0800	0.0800	-	-	-	-
Ou et al. JBHI'23	PIPOC	549	0.175	JSN	0.093	0.053	0.050	0.056	-	0.0095	0.0061	0.0076	7.2	3.5	2.8	3.5
This work	**CNN**	**705**	**0.175/0.15**	**JSN**	**0.086**	**0.046**	**0.066**	**0.066**	**0.0415**	**0.0351**	**0.0345**	**0.0370**	**3.31**	**0.68**	**0.48**	**1.49**

ASM: Active Shape Models. **PIPOC**: Partial image phase only correlation. **CNN**: Convolutional Neural Network.

Table 9
Comparison with related works in algorithm performance. The computational complex, the number of parameters, time required.

	Method	Dataset (images)[†]	Input Size	Output[‡]	Parameters	Time Required (s)
Langs et al. TMI'08	ASM	X-ray MCP joint(160)	-	JSW	-	-
Miao et al. TMI'16	CNN	TKA(100)&VIPS(7)&XEF(94)	18×52×52	TP	642,019,500	0.1100(GTX 980)
Mahapatra et al. MLMI'18	GAN	NIH ChestXray14(1087)	2×n×n	Mask&DF	-	0.5000(Tesla K40)
Miao et al. AAAI'18	CNN	CBCT(116sets)&Clinical(28sets)	2×128×128	TP	-	1.7666(Titan Xp)
Ou et al. JBHI'23	PIPOC	X-ray hand(549)	-	JSN	-	0.0121(CPU)
This work	**CNN**	**X-ray hand(705)**	**4×224×224**	**TP&JSN**	**47,000,224**	**0.0280(GTX 2080Ti)**

[†] **TKA**: Total Knee Arthroplasty Kinematics. **VIPS**: Virtual Implant Planning System. **XEF**: X-ray Echo Fusion.
[‡] **DF**: Deformation field. **TP**: Transformation parameters

framework has lower uncertainty and greater potential for accuracy and sensitivity. Neural network is a non-linear function (Gawlikowski et al., 2021), which can lead to higher uncertainty when compared to traditional image processing algorithms. The major uncertainty in this work is aleatoric uncertainty, that exist due to noise, and is irreducible by improving the quality or quantity of data (Indrayan and Malhotra, 2017; Mehrtash et al., 2020). As the standard deviations σ shown in the Table 8, the impact of aleatoric uncertainty can be controlled, and can attain the standard deviation σ close to the PIPOC-based JSN progression quantification work in (Ou et al., 2023). This demonstrates that our proposed network can achieve lower mismatching ratio while ensuring similar accuracy to previous studies.

Furthermore, considering the specific focus of our registration method on image registration in this domain and the abundance of related studies addressing various tasks, we conducted a comparative analysis with other models (including both rigid and non-rigid registration models) in terms of model parameters and time consuming. Table. 9 reveals that our method exhibits fewer model parameters than other models, particularly with larger input data. Moreover, while our time required is higher compared to CPU-based registration method PIPOC, it is slightly lower than other deep learning-based registration methods using a GPU.

4. Conclusion and Future work

In this work, we propose a deep learning method for joint space narrowing (JSN) progression quantification in rheumatoid arthritis (RA). The proposed method includes an image segmentation network based on U-net++, and a ResNet-like deep intra-subject rigid registration network for displacement quantification. Our experiments demonstrated that image registration based JSN progression quantification framework exhibits greater advantages and potential in terms of accuracy and sensitivity compared to two other mainstream frameworks for joint space quantification, namely margin detection based joint space width (JSW) quantification and machine learning classification-based scoring.

Image registration is a hot research topic in the field of medical image analysis, particularly in recent years, where non-rigid image registration algorithms have found extensive application scenarios. However, considering the emphasis on the overall bone displacement during JSN progression quantification in RA, this work utilized a rigid registration network to calculate the relative bone displacement.

Compared to non-rigid registration networks, this network has clear advantages in terms of computation time, computational complexity, and parameter requirements. Compared to existing image registration based JSN progression quantification works, this work significantly improves the robustness for scaling, rotation, and noise while maintaining almost comparable accuracy and sensitivity. This work can handle complex clinical situations and reduce mismatches due to inconsistent angle and spatial resolution of radiography images. Additionally, our approach provides a misalignment visualization as a reliability indicator that can be used by radiologists and rheumatologists to assess the quantification reliability, thus, making it a promising tool for future clinical applications.

The JSN progression measured in this work was analyzed in comparison to the GSS measured by rheumatologists. The results indicate that the JSN of the finger joints with GSS progression [ΔGSS(+)] was significantly greater than the JSN of the joints without GSS progression [ΔGSS(-)] (Unpaired T test, $p = 0.0067$ (one-tail), 95% confidence

interval: 0.02143 to 0.1831). And there is a positive correlation between measured JSN progression and ΔGSS (Pearson Correlation, p =0.0003, r=0.2167).

Our experimental results demonstrate that our proposed rigid convolutional neural registration network can be used for quantifying JSN progression in RA, offering advantages of high accuracy, high sensitivity, and high robustness.

Recently, non-rigid registration network based on deformation fields have received significant attention and development. Our algorithmic process is a kind of regional image registration. An interesting direction for future research could be the incorporation of segmentation information to immobilize the target region of the deformation field. This approach draws on the advantages of the deformation field, enabling the quantification of JSN progression in complex joint regions, such as wrist joints. This could lead to more comprehensive monitoring in the early stages of RA and provide novel ideas for registration-based joint space measurements.

5. Acknowledgments

This work was supported in part by the Japan Society for the Promotion of Science (JSPS) Grants-in-Aid for Scientific Research (KAKENHI) under Grant 21K07611, and in part by JST SPRING under Grant JPMJSP2119. The authors would like to sincerely thank Akira Sagawa, MD, PhD, Sagawa Akira Rheumatology Clinic (Sapporo, Japan), Masaya Mukai, MD, PhD, Sapporo City General Hospital (Sapporo, Japan) and Kazuhide Tanimura, MD, Hokkaido Medical Center for Rheumatic Diseases (Sapporo, Japan) for image data preparation.

CRediT authorship contribution statement

Haolin Wang: Methodology, Software, Validation, Investigation, Writing − original draft, Experiment, Visualization. **Yafei Ou:** Conceptualization, Methodology, Validation, Investigation, Writing − original draft, Writing − review & editing, Experiment, Visualization, Supervision, Project administration. **Wanxuan Fang:** Data Curation. **Prasoon Ambalathankandy:** Investigation, Writing − review & editing. **Naoto Goto:** Formal analysis. **Gen Ota:** Formal analysis. **Taichi Okino:** Data curation. **Jun Fukae:** Data curation. **Kenneth Sutherland:** Writing − review & editing. **Masayuki Ikebe:** Resources, Funding acquisition, Supervision. **Tamotsu Kamishima:** Resources, Data Curation, Funding acquisition, Supervision.

References

Aletaha, D., Smolen, J.S., 2018. Diagnosis and management of rheumatoid arthritis: a review. Jama 320, 1360–1372.

Chen, X., Diaz-Pinto, A., Ravikumar, N., Frangi, A.F., 2021. Deep learning in medical image registration. Progress in Biomedical Engineering 3, 012003.

Chen, X., Wang, X., Zhang, K., Fung, K.M., Thai, T.C., Moore, K., Mannel, R.S., Liu, H., Zheng, B., Qiu, Y., 2022. Recent advances and clinical applications of deep learning in medical image analysis. Medical Image Analysis , 102444.

Ehrhardt, J., Werner, R., Schmidt-Richberg, A., Handels, H., 2010. Statistical modeling of 4d respiratory lung motion using diffeomorphic image registration. IEEE transactions on medical imaging 30, 251–265.

Fu, Y., Lei, Y., Wang, T., Curran, W.J., Liu, T., Yang, X., 2020. Deep learning in medical image registration: a review. Physics in Medicine & Biology 65, 20TR01.

Gawlikowski, J., Tassi, C.R.N., Ali, M., Lee, J., Humt, M., Feng, J., Kruspe, A., Triebel, R., Jung, P., Roscher, R., et al., 2021. A survey of uncertainty in deep neural networks. arXiv preprint arXiv:2107.03342 .

Genant, H.K., Jiang, Y., Peterfy, C., Lu, Y., Rédei, J., Countryman, P.J., 1998. Assessment of rheumatoid arthritis using a modified scoring method on digitized and original radiographs. Arthritis & Rheumatism: Official Journal of the American College of Rheumatology 41, 1583–1590.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778.

Van der Heijde, D., 2000. How to read radiographs according to the sharp/van der heijde method. The Journal of rheumatology 27, 261–263.

Hinton, G., Srivastava, N., Swersky, K., 2012. Neural networks for machine learning lecture 6a overview of mini-batch gradient descent. Cited on 14, 2.

Hirano, T., Nishide, M., Nonaka, N., Seita, J., Ebina, K., Sakurada, K., Kumanogoh, A., 2019. Development and validation of a deep-learning model for scoring of radiographic finger joint destruction in rheumatoid arthritis. Rheumatology advances in practice 3, rkz047.

Huo, Y., Vincken, K.L., van der Heijde, D., De Hair, M.J., Lafeber, F.P., Viergever, M.A., 2015. Automatic quantification of radiographic finger joint space width of patients with early rheumatoid arthritis. IEEE Transactions on Biomedical Engineering 63, 2177–2186.

Indrayan, A., Malhotra, R.K., 2017. Medical biostatistics. CRC Press.

Ingram, J.N., Körding, K.P., Howard, I.S., Wolpert, D.M., 2008. The statistics of natural hand movements. Experimental brain research 188, 223–236.

Kaiser, M., John, M., Heimann, T., Brost, A., Neumuth, T., Rose, G., 2014. 2d/3d registration of tee probe from two non-orthogonal c-arm directions, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 283–290.

Kato, K., Yasojima, N., Tamura, K., Ichikawa, S., Sutherland, K., Kato, M., Fukae, J., Tanimura, K., Tanaka, Y., Okino, T., et al., 2019. Detection of fine radiographic progression in finger joint space narrowing beyond human eyes: phantom experiment and clinical study with rheumatoid arthritis patients. Scientific reports 9, 8526.

Kingma, D.P., Ba, J., 2014. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 .

Langs, G., Peloschek, P., Bischof, H., Kainberger, F., 2008. Automatic quantification of joint space narrowing and erosions in rheumatoid arthritis. IEEE transactions on medical imaging 28, 151–164.

Mahapatra, D., Ge, Z., Sedai, S., Chakravorty, R., 2018. Joint registration and segmentation of xray images using generative adversarial networks, in: Machine Learning in Medical Imaging: 9th International Workshop, MLMI 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Proceedings 9, Springer. pp. 73–80.

Mehrtash, A., Wells, W.M., Tempany, C.M., Abolmaesumi, P., Kapur, T., 2020. Confidence calibration and predictive uncertainty estimation for deep medical image segmentation. IEEE transactions on medical imaging 39, 3868–3878.

Miao, S., Piat, S., Fischer, P., Tuysuzoglu, A., Mewes, P., Mansi, T., Liao, R., 2018. Dilated fcn for multi-agent 2d/3d medical image registration, in: Proceedings of the AAAI Conference on Artificial Intelligence.

Miao, S., Wang, Z.J., Liao, R., 2016. A cnn regression approach for real-time 2d/3d registration. IEEE transactions on medical imaging 35, 1352–1363.

Minh, N.N., Nguyen, N., Ngoc, C.N., Duy, T.T., Huy, T.N., Do, B.N., Viet, T.T., 2022. Application of imagej software for the quantification of hand joint space narrowing in patients with rheumatoid arthritis. Current Rheumatology Reviews 18, 136–143.

Nakatsu, K., Morita, K., Yagi, N., Kobashi, S., 2020. Finger joint detection method in hand x-ray radiograph images using statistical shape model and support vector machine, in: 2020 International Symposium on Community-centric Systems (CcS), IEEE. pp. 1–5.

Neylon, J., Min, Y., Low, D.A., Santhanam, A., 2017. A neural network approach for fast, automated quantification of dir performance. Medical physics 44, 4126–4138.

Okino, T., Ou, Y., Ikebe, M., Tamura, K., Sutherland, K., Fukae, J., Tanimura, K., Kamishima, T., 2023. Fully automatic software for detecting radiographic joint space narrowing progression in rheumatoid arthritis: phantom study and comparison with visual assessment. Japanese Journal of Radiology 41, 510–520.

Ou, Y., Ambalathankandy, P., Furuya, R., Kawada, S., Kamishima, T., Ikebe, M., 2022. Joint space narrowing progression quantification with joint angle correction in rheumatoid arthritis, in: 2022 14th Biomedical Engineering International Conference (BMEiCON), IEEE. pp. 1–5.

Ou, Y., Ambalathankandy, P., Furuya, R., Kawada, S., Zeng, T., An, Y., Kamishima, T., Tamura, K., Ikebe, M., 2023. A sub-pixel accurate quantification of joint space narrowing progression in rheumatoid arthritis. IEEE Journal of Biomedical and Health Informatics 27, 53–64.

Ou, Y., Ambalathankandy, P., Shimada, T., Kamishima, T., Ikebe, M., 2019. Automatic radiographic quantification of joint space narrowing progression in rheumatoid arthritis using poc, in: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), IEEE. pp. 1183–1187.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., et al., 2019. Pytorch: An imperative style, high-performance deep learning library. Advances in neural information processing systems 32.

Peloschek, P., Langs, G., Weber, M., Sailer, J., Reisegger, M., Imhof, H., Bischof, H., Kainberger, F., 2007. An automatic model-based system for joint space measurements on hand radiographs: initial experience. Radiology 245, 855–862.

Pfeil, A., Renz, D.M., Hansch, A., Kainberger, F., Lehmann, G., Malich, A., Wolf, G., Böttcher, J., 2013. The usefulness of computer-aided joint space analysis in the assessment of rheumatoid arthritis. Joint Bone Spine 80, 380–385.

Platten, M., Kisten, Y., Kälvesten, J., Arnaud, L., Forslind, K., van Vollenhoven, R., 2017. Fully automated joint space width measurement and digital x-ray radiogrammetry in early ra. RMD open 3, e000369.

Rydell, E., Forslind, K., Nilsson, J.Å., Karlsson, M., Åkesson, K.E., Jacobsson, L.T., Turesson, C., 2021. Predictors of radiographic erosion and joint space narrowing progression in patients with early rheumatoid arthritis: a cohort study. Arthritis research & therapy 23, 1–13.

Stoel, B., 2020. Use of artificial intelligence in imaging in rheumatology–current status and future perspectives. RMD open 6, e001063.

Zhang, Y.J., 1996. A survey on evaluation methods for image segmentation. Pattern recognition 29, 1335–1346.