Radiographic Bone Texture Analysis Using Deep Learning Models for Early Rheumatoid Arthritis Diagnosis

Yun-Ju Huang
  CGMH: Chang Gung Memorial Hospital   https://orcid.org/0000-0002-6226-6635

Miao Shun
  PAII labs, Bethesda, USA

Kang Zheng
  PAII labs, Bethesda, USA

Le Lu
  PAII labs, Bethesda, USA

Yuhang Lu
  PAII labs, Bethesda, USA

Chihung Lin
  Chang Gung Memorial Hospital

Chang-Fu Kuo (✉ zandis@gmail.com)
  Division of Rheumatology, Allergy and Immunology, Chang Gung Memorial Hospital, Taoyuan, Taiwan 2. School of Medicine, Chang Gung University, Taoyuan, Taiwan   https://orcid.org/0000-0002-9770-5730

Abstract

Background

Rheumatoid arthritis (RA) is characterized by altered bone microarchitecture (radiographically referred to as 'texture') of periarticular regions. We hypothesize that deep learning models can quantify periarticular texture changes to aid in the classification of early RA.

Methods

The second, third, and fourth distal metacarpal areas from hand radiographs of 892 early RA and 1236 non-RA patients were segmented for the Deep Texture Encoding Network (Deep-TEN; texture-based) and residual network-50 (ResNet-50; texture and structure-based) models to predict the probability of RA. The performances were measured using the area under the curve of the receiver operating characteristics curve (AUROC). Multivariate logistic regression was used to estimate the odds ratio (OR) with 95% confidence intervals (CIs) for RA.

Results

The AUROC for RA was 0.69 for the Deep-TEN and 0.73 for the ResNet-50 model. The positive predictive values of a high texture score to classify RA using the Deep-TEN and ResNet-50 models were 0.64 and 0.67, respectively. High mean texture scores were associated with age- and sex-adjusted ORs (95% CI) for RA of 3.42 (2.59–4.50) and 4.30 (3.26–5.69) using the Deep-TEN and ResNet-50 models, respectively. The moderate and high RA risk groups determined by the Deep-TEN model were associated with adjusted ORs (95% CIs) of 2.48 (1.78–3.47) and 4.39 (3.11–6.20) for RA, respectively, and those using the ResNet-50 model were 2.17 (1.55–3.04) and 6.91 (4.83–9.90), respectively.

Conclusion

Fully automated quantitative assessment for periarticular texture by deep learning models can help the classification of early RA.

Background

Rheumatoid arthritis (RA) is an autoimmune disease characterized by symmetric polyarthritis at peripheral small joints, especially the proximal interphalangeal, metacarpal phalangeal, and radiocarpal joints. The progression of RA, including bone structural and textural changes, can be assessed via conventional radiographs, computed tomography (CT), magnetic resonance imaging (MRI), or densitometry.(1–3) Conventional radiography is an inexpensive and reproducible technique that can assist in RA screening, diagnosis, evaluation, and monitoring by indicating joint space narrowing, erosions, and periarticular bone microarchitecture (radiographically present as texture) changes such as osteoporosis.(4) However, it is challenging to assess periarticular texture changes using plain radiographs quantitatively; this poses a problem because the extent of periarticular osteoporosis could be an early indication of RA.

Fractal analysis is one of the techniques used to determine bone texture characteristics from radiographs. A fractal dimension is a measure of the space-filling capacity of a pattern and can be used to indicate irregular pattern complexity with self-similarity at different scales.(5) A particular type of fractal analysis is fractal signature analysis (FSA), which is a computerized textural analysis method used to measure vertical and horizontal trabeculae based on the fractal dimensions of the bone structure over a range of trabecular widths.(6–8) FSA has been previously applied

for bone architecture measurements and disease progression using radiographs in cases of knee osteoarthritis,(9–14) osteoporosis treatment response after administration of risedronate,(15) hand osteoarthritis,(16) and hip osteoarthritis.(17) Furthermore, differences in fractal signatures in RA radiographs among three types of bone conditions, namely normal, osteopenic, and eroded bone, have also been assessed.(18) While these previous studies indicate that bone disease classification and disease progression assessment can be performed by examining textures radiographically using techniques such as FSA, the present techniques are based on fixed descriptors for texture features. They are not capable of 'learning' latent features in radiology films that may indicate disease classification or progression.

Deep learning methods such as those based on multilayer convolutional neural networks (CNNs) are robust alternatives for various image analysis tasks, including image classification and segmentation.(19, 20) CNNs are capable of automatically learning and extracting hidden structural and textural bone features from radiographs to classify them and quantify their features, which are often not apparent to the human eye, such as those of periarticular osteoporosis and trabecular abnormalities. Therefore, we hypothesized that deep learning algorithms would be capable of identifying textural feature changes in periarticular regions of the phalanges, which could indicate signs of early RA. In addition, these textural features may be used to diagnose early RA using conventional radiographic images of the hand by dividing patients into different risk groups.

Methods

Patient characteristics and study design

In this study, we developed a deep learning-based image processing model to automatically detect and segment distal metacarpal bones as regions of interest (ROIs) in plain radiography images of both hands; the extracted radiographic features were used to classify the images for early RA. Our proposed model was trained, tested, and validated using data recorded at Chang Gung Memorial Hospital, Taiwan. In particular, digital anterior-posterior radiographs of bilateral hands from early RA and non-RA patients aged 18 years or older were retrospectively collected to form the primary dataset. The radiographs of RA were collected within one year of the initial diagnosis of RA, which was based on the 2010 European League Against Rheumatism / American College of Rheumatology (EULAR-ACR) classification criteria for RA.(21) The RA diagnosis was confirmed by two rheumatologists after a thorough chart review. This study was approved by the Institutional Review Board of Chang Gung Memorial Hospital, Taiwan. The requirement for signed informed consent was waived because the data used in this study were derived from partial hand radiographs obtained from de-identified digitized patient data to prevent any confidentiality concerns.

Datasets for CNN training and testing

Our CNN model for early RA classification was trained using a random set of 3,740 radiographs obtained from 892 RA and 1236 non-RA patients, which represent 80% of the primary dataset; this random set was further partitioned into training (80%) and validation (20%) datasets. It is noteworthy that multiple hand radiographs from the same patient were considered as independent radiographs in the training set. The final trained model was then tested using the remaining 20% of the primary dataset—consisting of 905 radiographs from 228 RA and 272 non-RA patients—as the test dataset. The digital radiographs included in our primary dataset were obtained at 50 kVp using the same radiography system (Fujifilm Healthcare); these radiographs were greyscaled and had resolutions ranging from 1192 × 1536 to 3015 × 2505 pixels.

Segmentation of ROIs

For the image pre-processing, model training, and validation tasks in our study, we used the high-performance computing systems available at the Center for Artificial Intelligence in Medicine, Chang Gung Memorial Hospital, Taiwan. Deep learning algorithms were used to segment the distal third of metacarpal bones and analyze the corresponding radiographic textural features from the radiographs. A curve-graph convolutional network (GCN) was trained for fully automated segmentation of the second, third, and fourth metacarpal bone images. The specific AI methodology for the GCN-based automated anatomical tissue segmentation approach used in this study has been described in a previous work (arXiv:2007.03052v2 [cs.CV], Accepted: MICCAI 2020). In summary, a novel GCN-based contour transformer network (CTN), which is a one-shot anatomy segmentor with a naturally built-in human-in-the-loop mechanism, was used to segment the ROIs in the radiographs by learning a contour evolution behaviour process. The CTN was trained to fit a contour to the required object boundary by learning from one labelled image exemplar; this network takes the image exemplar and an unlabelled image as inputs, and then detects contours with similar features as those in the image exemplar in the unlabelled image. Three losses were considered to ensure that the CTN was 'one-shot' trainable. This segmentation model was then connected to a classification model to realize a fully automatic process for RA classification.

The set of segmented images was augmented via random rotation (-180° to + 180°) and intensity jittering (brightness: -0.2 to + 0.2; contrast: -0.2 to 0.2). Subsequently, the obtained images were resized to 192 × 192 pixels before texture feature extraction. The deep texture encoding network (Deep-TEN) was the base architecture used to generate textural feature vectors in our study.(22) Finally, the texture feature vectors were used for RA classification of radiographs.

Algorithm and training of proposed RA classification models

We developed a deep learning algorithm based on the Deep-TEN model to extract bone textural features from hand radiographs. The proposed algorithm is based on a multilayer CNN with parameters that are structured as a hierarchy of layers. In general, a CNN image classification model scans an image to extract and aggregate structural and textural features from it. With a large amount of data, such a model can learn the essential features necessary to fit and identify ROIs for a problem, which, in our case, is the classification of radiography images for RA.

Deep learning models can extract texture representations using a pre-trained generic CNN model (such as the ResNet-18 or ResNet-50 models) considering both texture and structure or specific models considering texture alone.(23) The Deep-TEN model used as the base architecture in our proposed algorithm is a texture-specific model that includes a novel encoding layer on top of the convolutional layers of the generic ResNet-18 model.(24) Therefore, the Deep-TEN model is a specialized model that can detect and extract image texture features with superior performance, and is thus especially useful for material and texture recognition.(22) Because the features extracted by the Deep-TEN model are learnable, the proposed model is dynamic and does not rely on any fixed feature set. Our proposed model architecture is shown in Fig. 1. The vectors generated by the proposed model represent the orderless textural features; however, the structural features are excluded from these extracted representations. Separate models were trained for the second, third, and fourth distal metacarpal bones, and a final ensemble model was developed using the five trained models by averaging their outputs for the three metacarpal bones in an input image. Furthermore, we trained a ResNet-50 model to classify the radiographs for RA using the extracted ROIs for comparison with our proposed model; in this case, aside from the textural features, the structural changes in the images were also considered for RA classification. In previous works, the ResNet model has been shown to be useful for RA diagnosis, either using clinical information (25) or using diffuse optical tomography images.(26) Both Deep-TEN and ResNet-50 models take the ROIs as input and provide a continuous RA risk probability value between zero and one as an output. Patients were divided into groups of low, moderate, and high RA risk based on this output value. The dataset of the original training

radiographs was split into a subject-stratified 5-fold cross (FC) validation set. A final ensemble model was created from the corresponding five trained models by averaging their outputs for the three metacarpal bones in the input image.

Evaluation of the proposed model

The performance of our proposed model for RA classification of hand radiographs was evaluated using the test dataset. The receiver operator characteristic (ROC) curve was used to visualize the performance of the classification model for RA prediction, and the area under the ROC curve (AUROC) was used to indicate model performance, where a value of '1' indicates perfect prediction and a value of 0.5 or less indicates that the model has no class separation ability. Separate ROC curves were obtained for the Deep-TEN and ResNet-50 models, and the corresponding AUROCs and 95% confidence intervals (CIs) were also estimated.(27)

Statistical analysis

Summary statistics for patients with and without RA were compiled and compared. The performances of the Deep-TEN and ResNet-50 models were compared with the obtained RA classification results; in addition, metrics such as sensitivity, specificity, and positive predictive value were calculated. Differences were considered to be significant if there was a two-tailed P value of less than 0.05. Multivariate logistic regression was used to assess the association between the RA risk groups and RA diagnosis, and the odds ratios (ORs) and 95% CIs for RA were calculated with adjustments for age and sex. The image processing, deep learning model building, and training were based on Python programming language with the deep learning framework of Pytorch. All statistical analyses were conducted using the SAS program, version 9.4 (SAS Institute Inc., Cary, NC, USA).

Results

Patient characteristics

In this study, we acquired de-identified digitized medical data of 1119 RA patients, which were then split into the training/validation (n = 891) and test (n = 228) datasets such that both sets had patient data with similar age and sex distributions. The patient characteristics of all patients (RA and non-RA) are listed in Table 1. Furthermore, the median disease duration (interquartile range (IQR)) was 35 (14, 294) and 21 (14, 49) days in the training/validation and testing sets, respectively.

Performance comparison between the Deep-TEN and ResNet-50 models

The Deep-TEN model achieved an AUROC of 0.69 (95% CI: 0.64–0.74) for RA classification based on textural features obtained from patient radiographs; this performance was similar to that of the ResNet-50 model, which had an AUROC of 0.73 (95% CI: 0.69–0.77). Figure 2 shows the ROC curves of the Deep-TEN and ResNet-50 models for RA classification; from these curves, it can be observed that the Deep-TEN model, which uses only textural features for classification, is capable of classifying patient radiographs for early RA with a performance similar to ResNet-50 model, which considers both textural and structural features. Using the Youden's index, the cut-offs for the texture score for RA classification were obtained as 0.43 and 0.45 for the Deep-TEN and ResNet-50 models, respectively. The sensitivity, specificity, and positive predictive value of a high texture score to classify early RA were 0.67, 0.62, and 0.64 for the Deep-TEN model and 0.68, 0.67, and 0.67 for the ResNet-50 model.

Texture risk group and RA prediction

High mean texture scores with age- and sex-adjusted ORs (95% CI) of 3.42 (2.59–4.50) and 4.30 (3.26–5.69) were obtained using the Deep-TEN and ResNet-50 models, respectively, for RA prediction (see Table 2). Based on the results listed in Table 2, it can be deduced that the sex of patients did not have any significant effect on the models' RA classification performance. Further, we partitioned the predicted texture score into tertiles in order to differentiate the patients into three risk groups for RA. Table 3 lists the mean texture scores for RA risk in the three different risk categories. Using the Deep-TEN model, the moderate and high RA risk groups had age- and sex-adjusted ORs (95% CIs) of 2.48 (1.78–3.47) and 4.39 (3.11–6.20), respectively, compared with the low RA risk group. Similarly, using the ResNet-50 model, the age- and sex-adjusted ORs (95% CI) for RA were 2.17 (1.55–3.04) and 6.91 (4.83–9.90) in the moderate and high RA risk groups, respectively, compared with the low RA risk group.

Table 1. Characteristics of the patients in our study.

Characteristic	Training/Validation Set (N = 2128)			Test Set (N = 500)		
	Non-RA (N = 1237)	RA (N = 891)	P-value	Non-RA (N = 272)	RA (N = 228)	P-value
Age, mean ± S.D., years	57.1 ± 14.8	58.2 ± 12.7	<0.0001	57.0 ± 12.7	57.1 ± 12.2	0.880
Sex, n (%)						
Male	291 (23.5%)	172 (19.3%)	0.022	60 (22.1%)	40 (17.5%)	0.209
Female	946 (76.5%)	719 (80.7%)		212 (77.9%)	188 (82.5%)	
Disease duration, days	-	35 (14, 294)	-	-	21 (14, 49)	-

Table 2. Texture scores and RA risk prediction.

	Predicted results	RA (N = 453), n (%)		Control (N = 452), n (%)		Mean texture Scores (95% CI)	Crude OR (95% CI)		Adjusted OR (95% CI)	
Deep-TEN Model										
All patients	RA	304	(67.3%)	170	(37.5%)	0.57 (0.56–0.57)	3.42	(2.60–4.49)	3.42	(2.59–4.50)
	Non-RA	148	(32.7%)	283	(62.5%)	0.33 (0.33–0.34)	1		1	
Male patients	RA	46	(57.5%)	28	(31.8%)	0.57 (0.56–0.57)	2.90	(1.54–5.45)	3.06	(1.61–5.80)
	Non-RA	34	(42.5%)	60	(68.2%)	0.32 (0.31–0.33)	1		1	
Female patients	RA	258	(69.4%)	142	(38.9%)	0.56 (0.56–0.57)	3.55	(2.62–4.82)	3.60	(2.64–4.89)
	Non-RA	114	(30.7%)	223	(61.1%)	0.33 (0.33–0.34)	1		1	
ResNet-50 Model										
All patients	RA	306	(67.7%)	148	(32.7%)	0.64 (0.63–0.66)	4.32	(3.27–5.70)	4.30	(3.26–5.69)
	Non-RA	146	(32.3%)	305	(67.3%)	0.27 (0.26–0.28)	1		1	
Male patients	RA	52	(65.0%)	25	(28.4%)	0.65 (0.62–0.68)	4.68	(2.44–8.99)	4.93	(2.54–9.58)
	Non-RA	28	(35.0%)	63	(71.6%)	0.24 (0.21–0.26)	1		1	
Female patients	RA	254	(68.3%)	123	(33.7%)	0.64 (0.63–0.66)	4.23	(3.11–5.76)	4.24	(3.11–5.77)
	Non-RA	118	(31.7%)	242	(66.3%)	0.28 (0.27–0.29)	1		1	

Adjusted OR is adjusted for age

Table 3. RA prediction scores for different risk groups obtained using the proposed Deep-TEN and ResNet-50 models.

Texture scores/Risk groups	RA/Total people (%)	Mean texture scores (95% CI)	Crude OR		Adjusted OR	
			OR	95% CI	OR	95% CI
Deep-TEN model						
Low	93/301 (30.9%)	0.31 (0.30−0.31)	1		1	
Moderate	159/302 (52.6%)	0.45 (0.44−0.45)	2.49	(1.78−3.47)	2.48	(1.78−3.47)
High	200/302 (66.2%)	0.62 (0.61−0.63)	4.39	(3.12−6.17)	4.39	(3.11−6.20)
ResNet-50 model						
Low	87/301 (28.9%)	0.24 0.21−0.24)	1		1	
Moderate	142/302 (47.0%)	0.46 (0.45−0.47)	2.18	(1.56−3.06)	2.17	(1.55−3.04)
High	223/302 (73.8%)	0.71 (0.71−0.74)	6.94	(4.86−9.93)	6.91	(4.83−9.90)
Adjusted OR is adjusted for sex and age						

Discussion

In this study, we demonstrated that radiographic textural features of distal metacarpal bones could indicate early signs of RA. Because of the complexity of the high-dimensional textural features in radiographs, simple mathematical operations such as FSA cannot be used to describe them. In contrast, deep learning methods can provide an overall insight into the complex textural bone properties and yield risk scores based on them, thereby enabling the classification early RA and stratifying patients into different risk groups. Thus, deep learning methods can be used for automatic reporting of RA risk based on plain radiographs; this risk information could then be incorporated into standard clinical risk analysis for early RA prediction.

We compared two deep learning models, namely the Deep-TEN and ResNet-50 models, for RA classification. Based on our results, the performance of both models is similar in terms of binary classification into RA and non-RA radiographs. However, the primary difference between both models is that the Deep-TEN model only takes into account the textural information from radiographs for RA prediction, while the ResNet-50 model considers both their textural and structural features. For example, bone erosions resulting in a change in bone contour are not considered by the Deep-TEN model because it is a structural feature change. Therefore, the ResNet-50 model performs slightly better at identifying patients at high risk of RA. In contrast, the Deep-TEN model is better at separating the patients into three risk groups for RA based on changes in the texture, thereby forming a more homogenous risk continuum. Hence, the selection of a deep learning model for RA prediction in clinical settings would depend on clinical needs.

The 1987 ACR classification criteria for RA(28) define erosion or unequivocal bony decalcification (periarticular osteoporosis) in hand and wrist posteroanterior radiographs as one of the radiographic features relevant to RA diagnosis. Periarticular osteoporosis, which is a bone textural feature, is an osseous morphologic indication that is observed before the occurrence of bone erosions and joint space narrowing.(29, 30) Early periarticular osteoporosis, which is characterized by the loss of trabecular size and reduction in the number of metaphyseal regions, is difficult to detect and quantify via traditional hand radiography; therefore, X-ray radiogrammetry,(31) CT,(32) and MRI(33) have been applied to detect periarticular osteoporosis in previous studies. However, the application of these approaches in clinical settings is hampered by their high costs. In the 2010 EULAR-ACR classification criteria for RA, (21) information on RA diagnoses based on clinical features such as joint involvement or symptom duration as well as using laboratory tests for anti-citrullinated peptide antibodies, rheumatoid factor, C-reactive protein, and erythrocyte sedimentation rate were included. Radiographic bone texture changes were not emphasized as in the previous 1987 ACR criteria(28) because early indications of bone erosion and periarticular osteoporosis were difficult to assess objectively from plain radiographs, and this could have led to delayed RA diagnosis. Traditionally, conventional radiography was considered to be less sensitive to early indications of RA. Nevertheless, in recent times, with the assistance of machine learning techniques, as we have observed in our study, conventional radiography could perhaps be useful for early RA classification.

In many clinical situations, the automatic evaluation of radiographs using deep learning will be of great medical value, because such a system could potentially support RA diagnosis as a screening tool for RA in both general clinics and specialized hospitals. Furthermore, our proposed CNN model can estimate the bone texture score and predict RA from radiographs within one second per image, which is considerably faster than analyses by human clinicians. Thus, our proposed model could save time and be used as a diagnostic tool in countries where the number of available rheumatologists or radiologists is low. Furthermore, it can be used by family physicians to refer their patients to RA specialists based on the diagnostic predictions by the model. Moreover, because this is a computerized model, intraobserver and interobserver variabilities can be avoided if it is applied in clinical trials related to RA research.

Compared with our current work, previous attempts to use CNNs for the interpretation of hand radiography images of RA patients did not consider the distinctive textural or structural changes that occur in the joints of RA patients. In particular, Kemal et al. used 180 hand radiographs to train their CNN model for RA diagnosis and achieved an accuracy, sensitivity, and specificity of 73.33%, 0.68, and 0.78, respectively.(34) Toru et al. proposed a model that achieved accuracies of 49.3–65.4% for joint space narrowing and 70.6–74.1% for bone erosion detection on 30 hand radiographs; their model was trained using 186 radiographs.(35) Because these two studies used downsampled images of the entire radiographs, subtle textural changes were not considered. The CNN model has been applied not only to radiography images but also to other image modalities. For example, Jakob et al. used CNN to assess synovitis activity from ultrasound images and achieved an accuracy of 86.4%, sensitivity 0.864, and specificity of 0.864.(36) Lun et al. developed a CNN-based segmentation method for the wrist using T2-weighted fat-suppressed MRI images for early RA detection.(37)

Despite the advantages of our proposed CNN-based approach for the detection of early RA indications, our study has the following limitations. First, we only analyzed the texture of the second, third, and fourth distal metacarpal bones for RA classification of radiographs. Thus, further investigation is required to confirm whether the inclusion of radiographic images of other parts of the hand as input to the proposed CNN model would increase its RA risk classification performance. Second, the training data used in the current study are from patients with early RA (for most patients, RA was diagnosed less than a year prior to the study). Thus, later temporal changes in the bone texture

or structure due to RA as the disease progresses were not considered in our work. Third, the complexity of the proposed deep learning model with millions of parameters prevents a straightforward interpretation of the results by human doctors and clinicians.

Conclusions

In this study, we proposed a deep learning model that can detect bone texture changes related to RA from hand radiographs, which, when coupled with automatic joint detection and segmentation, can help the classification of early RA.

Abbreviations

RA
rheumatoid arthritis; Deep-TEN:Deep Texture Encoding Network; ResNet-50:residual network-50; AUROC:area under the curve of the receiver operating characteristics curve; OR:odds ratio; CI:confidence intervals; CT:computed tomography; MRI:magnetic resonance imaging; FSA:fractal signature analysis; CNNs:convolutional neural networks; ROIs:regions of interest; EULAR-ACR:European League Against Rheumatism / American College of Rheumatology; GCBN:graph convolutional network; CTN:contour transformer network; FC:fold cross; ROC:receiver operator characteristic; IQR:interquartile range

Declarations

Authors' contributions

All authors interpreted the data and critically reviewed and revised the article for important intellectual content. All authors read and approved the final manuscript

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

Study in this manuscript were conducted in accordance with the Declaration of Helsinki and were approved by independent ethics committees.

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests

Author details

[1]Division of Rheumatology, Allergy and Immunology, Chang Gung Memorial Hospital, Taoyuan, Taiwan [2]School of Medicine, Chang Gung University, Taoyuan, Taiwan [3]PAII labs, Bethesda, Maryland, USA [4]Center for Artificial Intelligence in Medicine, Chang Gung Memorial Hospital, Taoyuan, Taiwan

References

1. Østergaard M, Boesen M. Imaging in rheumatoid arthritis: the role of magnetic resonance imaging and computed tomography. Radiol Med. 2019;124(11):1128–41.

2. Peters M, van Tubergen A, Scharmga A, Driessen A, van Rietbergen B, Loeffen D, et al. Assessment of Cortical Interruptions in the Finger Joints of Patients With Rheumatoid Arthritis Using HR-pQCT, Radiography, and MRI. Journal of bone mineral research: the official journal of the American Society for Bone Mineral Research. 2018;33(9):1676–85.

3. Gauri LA, Fatima Q, Diggi S, Khan A, Liyakat A, Ajay BR. Study of Bone Mineral Density (BMD) in Patients with Rheumatoid Arthritis and its Co-relation with Severity of the Disease. J Assoc Phys India. 2017;65(4):26–30.

4. Drosos AA, Pelechas E, Voulgari PV. Conventional radiography of the hands and wrists in rheumatoid arthritis. What a rheumatologist should know and how to interpret the radiological findings. Rheumatol Int. 2019;39(8):1331–41.

5. de Melo RHC, Conci A. How Succolarity could be used as another fractal measure in image analysis. TELECOMMUN SYST. 2013;52(3):1643–55.

6. Larsen A, Dale K, Eek M. Radiographic evaluation of rheumatoid arthritis and related conditions by standard reference films. Acta Radiol Diagn (Stockh). 1977;18(4):481–91.

7. Sharp JT, Young DY, Bluhm GB, Brook A, Brower AC, Corbett M, et al. How many joints in the hands and wrists should be included in a score of radiologic abnormalities used to assess rheumatoid arthritis? Arthritis rheumatism. 1985;28(12):1326–35.

8. Sharp JT, Lidsky MD, Collins LC, Moreland J. Methods of scoring the progression of radiologic changes in rheumatoid arthritis. Correlation of radiologic, clinical and laboratory abnormalities. Arthritis rheumatism. 1971;14(6):706–20.

9. Lynch JA, Hawkes DJ, Buckland-Wright JC. Analysis of texture in macroradiographs of osteoarthritic knees using the fractal signature. Physics in medicine biology. 1991;36(6):709–22.

10. Buckland-Wright JC, Lynch JA, Macfarlane DG. Fractal signature analysis measures cancellous bone organisation in macroradiographs of patients with knee osteoarthritis. Ann Rheum Dis. 1996;55(10):749–55.

11. Messent EA, Ward RJ, Tonkin CJ, Buckland-Wright C. Cancellous bone differences between knees with early, definite and advanced joint space loss; a comparative quantitative macroradiographic study. Osteoarthritis cartilage. 2005;13(1):39–47.

12. Kraus VB, Collins JE, Charles HC, Pieper CF, Whitley L, Losina E, et al. Predictive Validity of Radiographic Trabecular Bone Texture in Knee Osteoarthritis: The Osteoarthritis Research Society International/Foundation for the National Institutes of Health Osteoarthritis Biomarkers Consortium. Arthritis & rheumatology (Hoboken, NJ). 2018;70(1):80–7.

13. Hirvasniemi J, Niinimäki J, Thevenot J, Saarakkala S. Bone Density and Texture from Minimally Post-Processed Knee Radiographs in Subjects with Knee Osteoarthritis. Ann Biomed Eng. 2019;47(5):1181–90.

14. Kraus VB, Feng S, Wang S, White S, Ainslie M, Brett A, et al. Trabecular morphometry by fractal signature analysis is a novel marker of osteoarthritis progression. Arthritis rheumatism. 2009;60(12):3711–22.

15. Buckland-Wright JC, Messent EA, Bingham CO 3rd, Ward RJ, Tonkin C. A 2 year longitudinal radiographic study examining the effect of a bisphosphonate (risedronate) upon subchondral bone loss in osteoarthritic knee patients. Rheumatology. 2007;46(2):257–64.

16. Wolski M, Podsiadlo P, Stachowiak GW. Directional fractal signature methods for trabecular bone texture in hand radiographs: data from the Osteoarthritis Initiative. Medical physics. 2014;41(8):081914.

17. Papaloucas CD, Ward RJ, Tonkin CJ, Buckland-Wright C. Cancellous bone changes in hip osteoarthritis: a short-term longitudinal study using fractal signature analysis. Osteoarthritis cartilage. 2005;13(11):998–1003.

18. Buckland-Wright JC, Lynch JA, Bird C. Microfocal techniques in quantitative radiography: measurement of cancellous bone organization. Br J Rheumatol. 1996;35(Suppl 3):18–22.

19. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436–44.

20. Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. Insights into imaging. 2018;9(4):611–29.

21. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO 3. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Ann Rheum Dis. 2010;69(9):1580–8. rd, et al.

22. Zhang H. Deep TEN: Texture Encoding Network. IEEE Conference on CVPR. 2017.

23. Liu L, Chen J, Fieguth P, Zhao G, Chellappa R, Pietikäinen M. From BoW to CNN: Two Decades of Texture Representation for Texture Classification. Int J Comput Vis. 2018.

24. Zhang H, Xue J, Dana K, editors. Deep TEN: Texture Encoding Network. 2017 IEEE Conference on Computer Vision and Recognition P. (CVPR); 2017 21–26 July 2017.

25. Fukae J, Isobe M, Hattori T, Fujieda Y, Kono M, Abe N, et al. Convolutional neural network for classification of two-dimensional array images generated from clinical information may support diagnosis of rheumatoid arthritis. Scientific reports. 2020;10(1):5648.

26. Dehghani H, Feng Y, Lighter D, Zhang L, Wang Y, editors Deep Neural Networks Improve Diagnostic Accuracy of Rheumatoid Arthritis using Diffuse Optical Tomography. Clinical and Preclinical Optical Diagnostics II; 2019 2019/06/23; Munich: Optical Society of America.

27. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988;44(3):837–45.

28. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis rheumatism. 1988;31(3):315–24.

29. Fouque-Aubert A, Chapurlat R, Miossec P, Delmas PD. A comparative review of the different techniques to assess hand bone damage in rheumatoid arthritis. Joint bone spine. 2010;77(3):212–7.

30. Hauser B, Riches PL, Wilson JF, Horne AE, Ralston SH. Prevalence and clinical prediction of osteoporosis in a contemporary cohort of patients with rheumatoid arthritis. Rheumatology. 2014;53(10):1759–66.

31. Pfeil A, Haugeberg G, Renz DM, Reinhardt L, Jung C, Franz M, et al. Digital X-ray radiogrammetry and its sensitivity and specificity for the identification of rheumatoid arthritis-related cortical hand bone loss. J Bone Miner Metab. 2017;35(2):192–8.

32. Feehan LM, Li LL, McKay HA. Micro-structural bone changes in early rheumatoid arthritis persist over 1-year despite use of disease modifying anti-rheumatic drug therapy. BMC Musculoskelet Disord. 2017;18(1):521.

33. Schett G. Bone marrow edema. Ann N Y Acad Sci. 2009;1154:35–40.

34. Üreten K, Erbay H, Maraş HH. Detection of rheumatoid arthritis from hand radiographs using a convolutional neural network. Clinical rheumatology. 2019.

35. Hirano T, Nishide M, Nonaka N, Seita J, Ebina K, Sakurada K, et al. Development and validation of a deep-learning model for scoring of radiographic finger joint destruction in rheumatoid arthritis. Rheumatology advances in practice. 2019;3(2):rkz047.

36. Andersen JKH, Pedersen JS, Laursen MS, Holtz K, Grauslund J, Savarimuthu TR, et al. Neural networks for automatic scoring of arthritis disease activity on ultrasound images. RMD open. 2019;5(1):e000891.

37. Wong LM, Shi L, Xiao F, Griffith JF. Fully automated segmentation of wrist bones on T2-weighted fat-suppressed MR images in early rheumatoid arthritis. Quantitative imaging in medicine surgery. 2019;9(4):579–89.
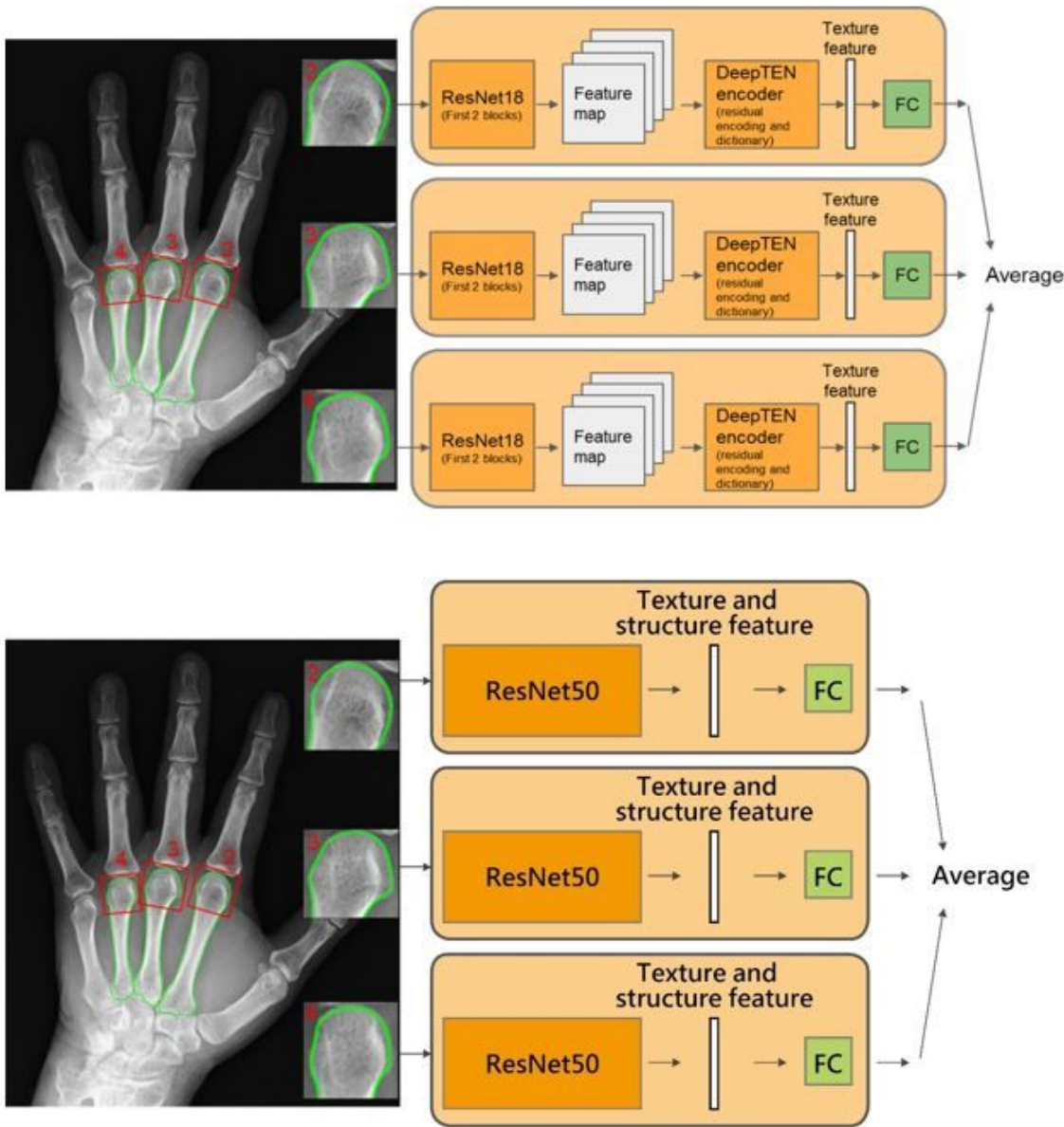
Figures

Figure 1

Flow of machine learning The first model (above) depicts our proposed CNN model, which produces a texture descriptor; it is comprised of the first two blocks of the ResNet-18 model, which are then followed by a Deep TEN encoder. The second model (below) depicts a CNN model with the ResNet-50 network as the backbone architecture, which produces a dense feature map, followed by a GCN that detects the boundary of the metacarpal bones. Subsequently, a texture score was assigned to each subchondral region detected by the GCN. The data set of the original training radiographs was split into a subject-stratified 5-fold cross (FC) validation set. A final ensemble model was created from the corresponding five trained models by averaging their outputs for the three metacarpal bones in the input image.
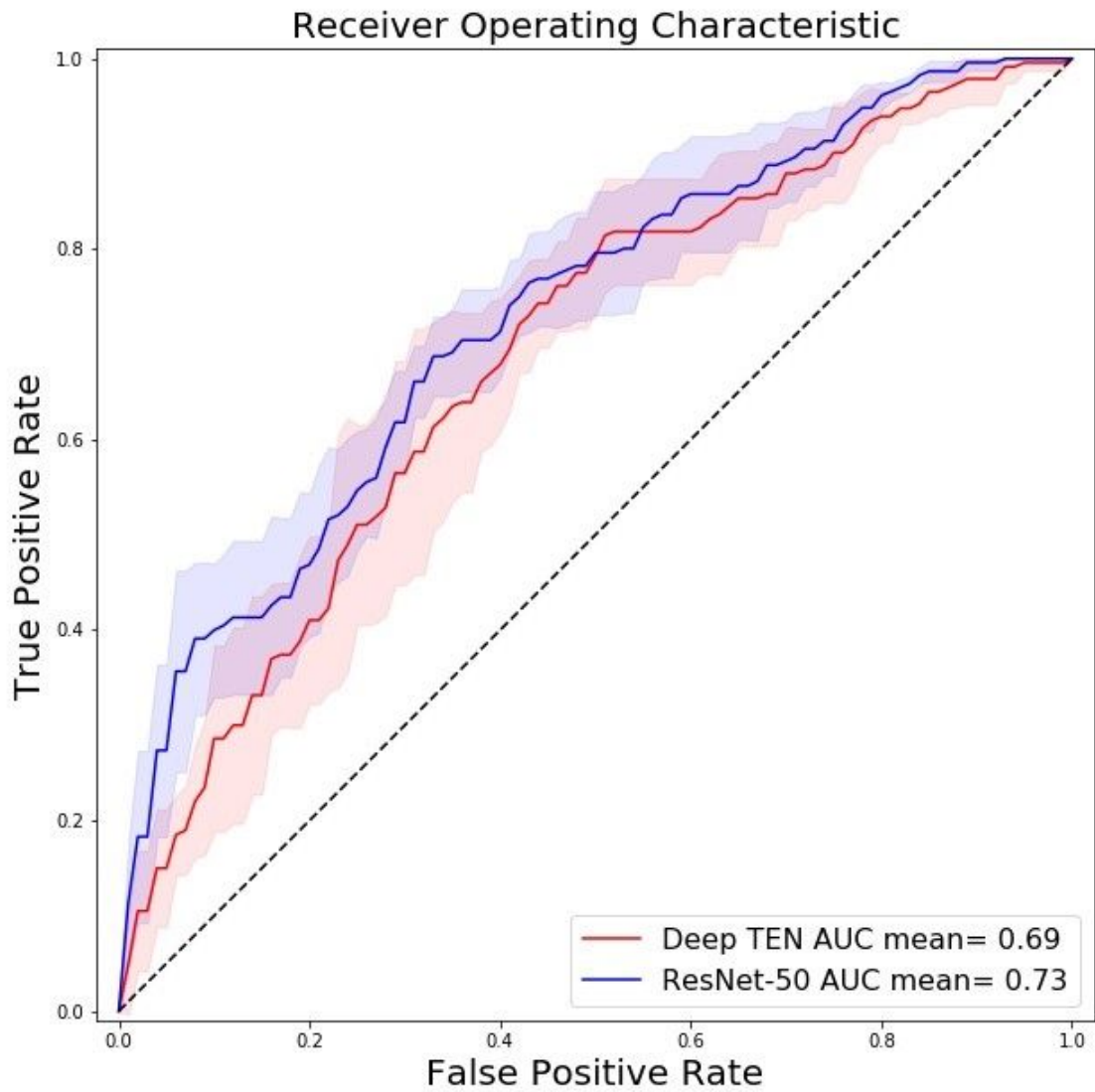
Figure 2

ROC curves of the Deep-TEN and ResNet-50 models for RA classification