Clinical Rheumatology
https://doi.org/10.1007/s10067-019-04487-4

ORIGINAL ARTICLE

Detection of rheumatoid arthritis from hand radiographs using a convolutional neural network

Kemal Üreten[1,2] · Hasan Erbay[3] · Hadi Hakan Maraş[4]

Received: 17 December 2018 / Revised: 28 January 2019 / Accepted: 17 February 2019

Abstract

Introduction Plain hand radiographs are the first-line and most commonly used imaging methods for diagnosis or differential diagnosis of rheumatoid arthritis (RA) and for monitoring disease activity. In this study, we used plain hand radiographs and tried to develop an automated diagnostic method using the convolutional neural networks to help physicians while diagnosing rheumatoid arthritis.

Methods A convolutional neural network (CNN) is a deep learning method based on a multilayer neural network structure. The network was trained on a dataset containing 135 radiographs of the right hands, of which 61 were normal and 74 RA, and tested it on 45 radiographs, of which 20 were normal and 25 RA.

Results The accuracy of the network was 73.33% and the error rate 0.0167. The sensitivity of the network was 0.6818; the specificity was 0.7826 and the precision 0.7500.

Conclusion Using only pixel information on hand radiographs, a multi-layer CNN architecture with online data augmentation was designed. The performance metrics such as accuracy, error rate, sensitivity, specificity, and precision state shows that the network is promising in diagnosing rheumatoid arthritis.

Keywords Convolutional neural network · Deep learning · Plain hand radiographs · Rheumatoid arthritis

Introduction

Rheumatoid arthritis (RA) is an inflammatory arthritis that affects as much as 1% of the population and may cause deformities in the joints and permanent disability. It may affect all synovial joints, especially the metacarpophalangeal (MCP), proximal interphalangeal (PIP), and wrist joints. In rheumatoid arthritis, soft tissue swelling, and periarticular osteoporosis are seen in the early stages, and joint space narrowing, erosions, and deformities are seen later [1–4]. A modified Sharp score can be used to determine the damage caused by rheumatoid arthritis in the hand and wrist [5].

RA diagnosis is established by a joint evaluation of the findings of examination, laboratory results, and imaging of the patient. One of the seven criteria for RA classification, according to the 1987 American College of Rheumatology (ACR), is the presence of juxta-articular erosion in hand radiographs. Synovitis, proven by imaging, is one of the 2010 ACR/EULAR criteria for RA classification, indicating the importance of imaging in an RA diagnosis [6, 7]. Imaging methods used in RA diagnosis include plain radiographs, ultrasonography, MRI, and CT. Plain radiography is the most widely used and is the first-line imaging method for diagnosis and differential diagnosis of RA and for monitoring the activity of the disease. Plain radiographs are relatively inexpensive and easily accessible imaging method [7]. Modifying antirheumatic drugs (DMARDs) and biological drugs are used in the treatment of RA [8, 9].

Recent improvements in computer technology have generated a rich pool of data with high resolution. In addition, image processing has improved greatly due to improvements

✉ Kemal Üreten
kemalureten@yahoo.com

1 Department of Rheumatology, Faculty of Medicine, Kırıkkale University, 71450 Kırıkkale, Turkey

2 Department of Computer Engineering (MSc), Çankaya University, Ankara, Turkey

3 Department of Computer Engineering, Faculty of Engineering, Kırıkkale University, Kırıkkale, Turkey

4 Department of Computer Engineering, Faculty of Engineering, Çankaya University, Ankara, Turkey

Published online: 08 March 2019

in central processing units (CPU) and graphics processing units (GPU) technology. If a dataset is large enough, heterogeneous images can be classified successfully. Computerized methods, especially those based on image analysis and machine learning, can facilitate the identification of pathological findings, improve accuracy, and support the physician's workflow [10, 11].

A convolutional neural network (CNN) is a deep learning method based on a multilayer neural network structure. A multilayer neural network has at least three layers: an input layer, a hidden layer, and an output layer. The data are presented to the network via the input layer, the desired output is obtained from the output layer, and the associated rules are learned in the hidden layer [10, 11]. In deep learning models, inputs are always in vector form; however, in medical images, structural information contained in adjacent pixels or voxels is important, and vectorization destroys this structural and configurational knowledge in images. CNN was designed to make better use of spatial and configuration information. CNN uses three mechanisms, consisting of a local receiver, weight sharing, and subsampling. A CNN has a series of convolution layers, pooling layers, and connected layers. The role of a convolution layer is to detect local features at different positions in the input feature maps. The pooling layer follows a convolution layer to downsample the feature maps of the previous convolution layer. After each convolution layer, a rectified linear unit (ReLU) is applied to give a nonlinear activation function [10, 11].

Materials and methods

Image acquisition

The images in the raw dataset were taken from examined patients between Jan 1, 2012, and Oct 1, 2018, in the rheumatology outpatient clinic of Medical Faculty of University of Kırıkkale. Data entails 180 radiograph images of both hands with different sizes. Eighty-one of the patients are normal, while 99 suffer from RA. The characteristics of the dataset is presented in Table 1.

Table 1 Characteristics of the dataset

Cases	180
Normal	81
RA	99
Image resolution	Various
Image format	Radiographs

Data pre-processing and splitting

By cropping the right hands [12] from radiographs whilst keeping the same aspect ratio from the raw dataset, the images in the dataset were gathered. Resizing the images to 160 × 240 and normalizing pixel values to ensure reduction of the range between zero and one were performed.

The dataset was split randomly into a training set and a test set by using a stratified selection strategy. The splitting is shown in Table 2.

Data augmentation

Image data augmentation is a necessary step to hinder CNN from learning irrelevant patterns, overfitting, and memorizing exact details of training images. Moreover, it boosts overall performance. Online data augmentation was conducted by random horizontal and vertical translation of images up to three pixels and by rotating the images with an angle up to 20°.

The architecture of the CNN model proposed

The CNN model architecture that we proposed includes six groups of convolution layers: consist of a convolution layer, a batch normalization layer, a ReLU and five max-pooling layers followed by one fully connected layer, and a Softmax layer. The details of the proposed CNN architecture are clearly presented in Table 3.

Results

A MATLAB® programming environment was used for the implementation of the CNN model architecture. A dataset of 135 right-hand radiographs were trained through the network: 61 were normal, 74 had RA. It was then tested on 45 radiographs that consist of 20 normal and 25 RA. Figure 1 presents a sample from the dataset.

Plots of accuracy and loss over several iterations for the proposed model are illustrated in Figs. 2 and 3. It is clear that during the training process, the training accuracy (blue line) and validation accuracy (dashed black line) approach each other. The loss subplot also shows that both the training loss and validation loss are low and decrease throughout the process.

Table 2 Training and testing numbers

Cases	Training	Test
Normal	61	20
Rheumatoid arthritis	74	25
Total	135	45

Table 3 Layered architecture of the network

Layer	Name	Layer type	Layer description
1	İnput	Image input	240 × 160 × 1 images with "zero center" normalization
2	conv_1	Convolution	32 11 × 11 × 1 convolutions with stride [1 1] and padding [1 1 1 1]
3	BN_1	Batch normalization	Batch normalization with 32 channels
4	relu_1	ReLU	ReLU
5	maxpool_1	MaxPooling	2 × 2 max pooling with stride [2 2] and padding [0 0 0 0]
6	conv_2	Convolution	64 9 × 9 × 32 convolutions with stride [1 1] and padding [1 1 1 1]
7	BN_2	Batch normalization	Batch normalization with 64 channels
8	relu_2	ReLU	ReLU
9	maxpool_2	MaxPooling	2 × 2 max pooling with stride [2 2] and padding [0 0 0 0]
10	conv_3	Convolution	128 7 × 7 × 64 convolutions with stride [1 1] and padding [1 1 1 1]
11	BN_2	Batch normalization	Batch normalization with 128 channels
12	relu_2	ReLU	ReLU
13	maxpool_3	MaxPooling	2 × 2 max pooling with stride [2 2] and padding [0 0 0 0]
14	conv_4	Convolution	256 5 × 5 × 128 convolutions with stride [1 1] and padding [1 1 1 1]
15	BN_4	Batch normalization	Batch normalization with 256 channels
16	relu_4	ReLU	ReLU
17	maxpool_4	MaxPooling	2 × 2 max pooling with stride [2 2] and padding [0 0 0 0]
18	conv_5	Convolution	512 3 × 3 × 256 convolutions with stride [1 1] and padding [1 1 1 1]
19	BN_5	Batch normalization	Batch normalization with 512 channels
20	relu_5	ReLU	ReLU
21	maxpool_5	MaxPooling	2 × 2 max pooling with stride [2 2] and padding [0 0 0 0]
22	conv_6	Convolution	1024 3 × 3 × 512 convolutions with stride [1 1] and padding [1 1 1 1]
23	BN_6	Batch normalization	Batch normalization with 1024 channels
24	relu_6	ReLU	ReLU
26	Fc	Fully connected	Two fully connected layers
27	Softmax	Softmax	Softmax
28	classOutput	Classification output	crossentropyex with classes "normal" and "rheumatoid"

Thirty-two convolutional kernels, measuring 11 × 11 × 1, taken from the first convolutional layer on the input images, measuring 240 × 160 × 1, are displayed in Fig. 4.

The results of applying the above feature to the image in Fig. 1 are shown in Fig. 5.

Fig. 1 Original image

To assess the efficiency of the developed CNN model, a number of performance metrics, including sensitivity, specificity, precision, F1 score, false negative rate, false discovery rate, false positive rate, negative predictive value, and classification accuracy were used. These metrics were in turn assessed through the confusion matrix shown in Fig. 6. Table 4 shows the details of the performance statistics.

Thirty-three patients out of 45 were correctly distinguished using the proposed CNN model. However, the model did not succeed in identifying five normal and seven RA patients. Figure 6 presents the confusion matrix for the test dataset. The network achieved 73.33% accuracy.

Discussion

Recent advances in computer technology have facilitated the acquisition of high-resolution images and processing of images. These developments have revolutionized the fields of medical imaging, artificial intelligence, and machine learning, resulting in a number of successful studies in the field of medical imaging. Deep learning is quickly proving to be a state-of-the-art technique, and CNN is an application of this deep learning approach. Medical image analysis groups around the world are rapidly entering this field and applying CNN, and successful results have begun to emerge [10], with more than 300 publications involving CNN in 2015 and 2016 [10, 13].

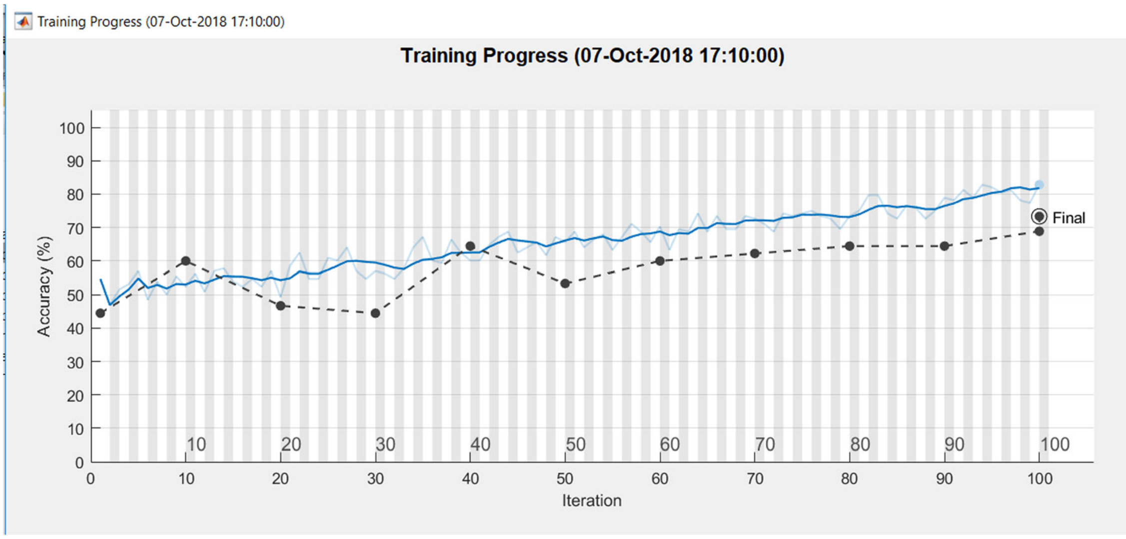


Fig. 2 Model accuracy

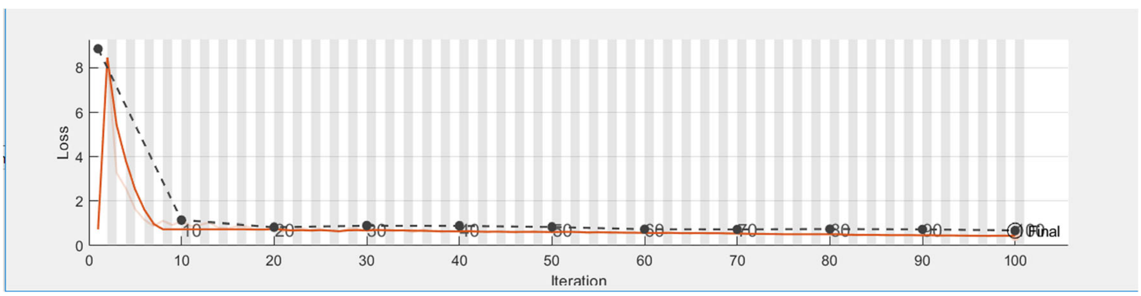
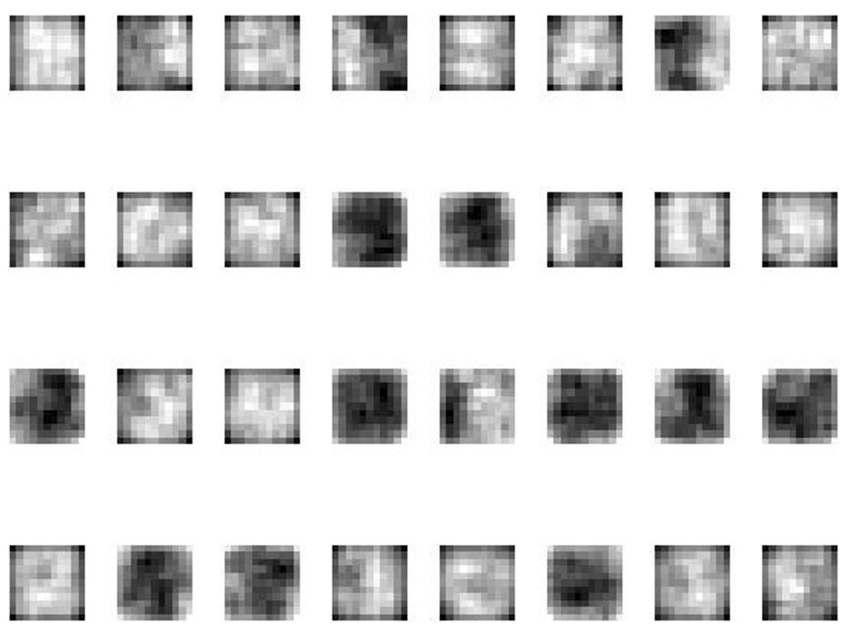


Fig. 3 Model loss

Plain hand radiographs have been used in several studies by applying a multilayer neural network. In a study conducted by Duryea et al., successful results were obtained from the use of a neural network to determine the localization of the joints commonly involved in osteoarthritis and RA in plain hand radiography [14]. Banerjee et al. succeeded in detecting osteophytes

Fig. 4 Features used in convolution layer conv 1



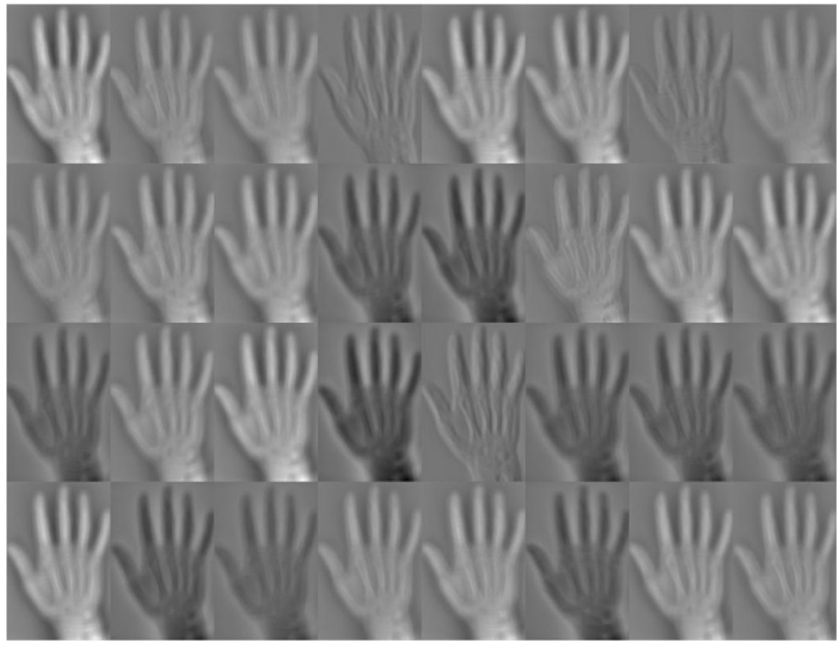


Fig. 5 Features shown in Fig. 4 applied to the image in Fig. 1

using hand radiography with a cellular neural network method using various filters. The results were shown to be accurate in over 90% of the images studied [15]. In children, two studies for the detection of bone age were performed on hand radiographs using a CNN method, and very successful results were obtained [16, 17]. Murakami et al. applied a CNN method to diagnose RA based on the detection of bone erosion in the hand joints. The region of interest was automatically set to the contour line of the segmented phalanges. They found a true positive rate of 80.5% and a false positive rate of 0.84% [18]. In the same study, they

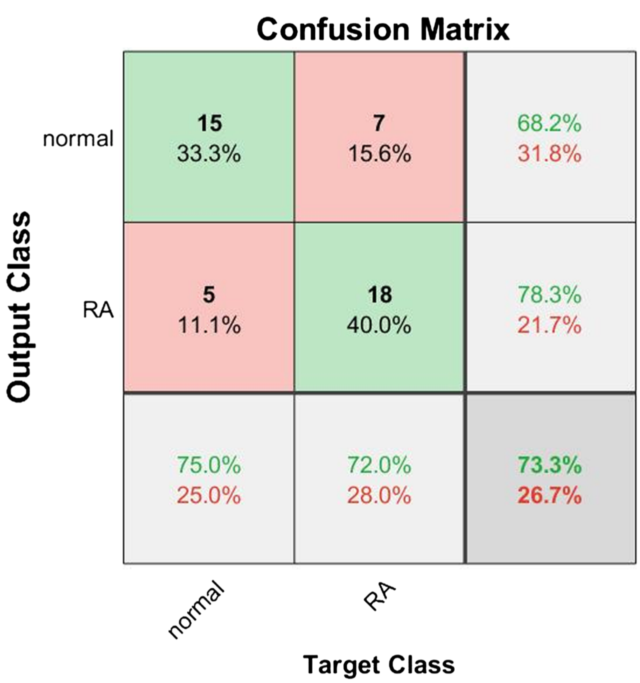


Fig. 6 Confusion matrix

worked on to detect only erosions seen in RA. Recall that in addition to erosions and soft tissue swelling, periarticular osteoporosis is seen in early stages of RA, and in later stages, symmetric narrowing in the joint space and deformities is seen. Herein, we tried to determine not only erosions but also other changes seen in RA. We proposed a deep learning CNN model with online data augmentation. The accuracy of the model was 73.33%. The sensitivity of the network was 68.18% whereas specificity was 78.26. The precision was 0.75. From these performance statistics, the model can be used to assist specialist during the diagnosis process. Even for non-specialist, the model might help at the initial examination.

A patient with pain, swelling, or stiffness in the hand joints may be examined by a general practitioner, orthopedic specialist, internal medicine specialist, physical therapist, or

Table 4 Performance statistics

Measure	Value
Sensitivity	0.6818
Specificity	0.7826
Precision	0.7500
Negative predictive value	0.7200
False positive rate	0.2174
False discovery rate	0.2500
False negative rate	0.3182
Accuracy	0.7333
F1 score	0.7143
Matthews correlation coefficient	0.4672
Elapsed time	3 min and 13 s
Hardware resource	Single GPU

rheumatologist. A diagnosis of RA is made by jointly evaluating the findings of the examination of the patient, the laboratory results, and imaging. Laboratory findings such as sedimentation and CRP elevation, anti-CCP, and rheumatoid factor positivity may not be present in all patients at the same time. Although conventional radiography (CR) is an easily accessible imaging method, the physician's experience in assessing RA findings using CR is important, and this is a time-consuming process. There may also be no experienced physicians in a center where CR is available. With early treatment, the permanent deformities seen in RA can be prevented. For physicians who do not have sufficient experience in evaluating RA-related changes using CR, an automated method of diagnosing RA from CR may be useful, and our findings can support this. In this case, the patient can be referred to the rheumatologist early, if necessary, and can therefore obtain early treatment. Experienced doctors such as radiologists and rheumatologists can also use this method when making their final decisions.

Although the radiographs of our patients showed both early and late radiological changes due to RA, we did not calculate the Modified Sharp Score for the radiographs, and further work is needed in this regard. We are also continuing to improve our dataset. In ongoing work, we are working on the diagnosis of osteoarthritis (OA) from hand radiographs, and initial results have been promising. In future work, we aim to be able to distinguish between RA and OA from hand radiographs.

Compliance with ethical standards

Disclosures None.

Ethics approval Ethical approval of this study was obtained from the Kırıkkale University Non-Interventional Research Ethics Committee on Dec 2, 2018, no. 2018.09.02.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Lee DM, Weinblatt ME (2001) Rheumatoid arthritis. Lancet 358(9285):903–911
2. Kourilovitch M, Galarza-Maldonado C, Ortiz-Prado E (2014) Diagnosis and classification of rheumatoid arthritis. J Autoimmun 48:26–30
3. Gabriel SE (2001) The epidemiology of rheumatoid arthritis. Rheum Dis Clin N Am 27(2):269–281
4. Renner W, Weinstein A (1988) Early changes of rheumatoid arthritis in the hand and wrist. Radiol Clin N Am 26(6):1185–1193
5. Sharp JT, Young DY, Bluhm GB, Brook A, Brower AC, Corbett M, Decker JL, Genant HK, Gofton JP, Goodman N (1985) How many joints in the hands and wrists should be included in a score of radiologic abnormalities used to assess rheumatoid arthritis? Arthritis Rheum 28(12):1326–1335
6. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO III, Birnbaum NS, Burmester GR, Bykerk VP, Cohen MD (2010) 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 62(9):2569–2581
7. Vyas S, Bhalla AS, Ranjan P, Kumar S, Kumar U, Gupta AK (2016) Rheumatoid arthritis revisited–advanced imaging review. Pol J Radiol 81:629–635
8. Kim Y, Oh H-C, Park JW, Kim I-S, Kim J-Y, Kim K-C, Chae D-S, Jo W-L, Song J-H (2017) Diagnosis and treatment of inflammatory joint disease. Hip Pelvis 29(4):211–222
9. Smolen JS, Landewe R, Breedveld FC, Buch M, Burmester G, Dougados M, Emery P, Gaujoux-Viala C, Gossec L, Nam J, Ramiro S, Winthrop K, de Wit M, Aletaha D, Betteridge N, Bijlsma JW, Boers M, Buttgereit F, Combe B, Cutolo M, Damjanov N, Hazes JM, Kouloumas M, Kvien TK, Mariette X, Pavelka K, van Riel PL, Rubbert-Roth A, Scholte-Voshaar M, Scott DL, Sokka-Isler T, Wong JB, van der Heijde D (2014) EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update. Ann Rheum Dis 73(3):492–509. https://doi.org/10.1136/annrheumdis-2013-204573
10. Greenspan H, Van Ginneken B, Summers RM (2016) Guest editorial deep learning in medical imaging: overview and future promise of an exciting new technique. IEEE Trans Med Imaging 35(5):1153–1159
11. Shen D, Wu G, Suk H-I (2017) Deep learning in medical image analysis. Annu Rev Biomed Eng 19:221–248
12. Koh JH, Jung SM, Lee JJ, Kang KY, Kwok SK, Park SH, Ju JH (2015) Radiographic structural damage is worse in the dominant than the non-dominant hand in individuals with early rheumatoid arthritis. PLoS One 10(8):e0135409. https://doi.org/10.1371/journal.pone.0135409
13. van Ginneken B (2017) Fifty years of computer analysis in chest imaging: rule-based, machine learning, deep learning. Radiol Phys Technol 10(1):23–32
14. Duryea J, Zaim S, Wolfe F (2002) Neural network based automated algorithm to identify joint locations on hand/wrist radiographs for arthritis assessment. Med Phys 29(3):403–411
15. Banerjee S, Bhunia S, Schaefer G (2011) Osteophyte detection for hand osteoarthritis identification in X-ray images using CNNs. In: Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE. IEEE, pp 6196–6199
16. Lee H, Tajmir S, Lee J, Zissen M, Yeshiwas BA, Alkasab TK, Choy G, Do S (2017) Fully automated deep learning system for bone age assessment. J Digit Imaging 30(4):427–441
17. Spampinato C, Palazzo S, Giordano D, Aldinucci M, Leonardi R (2017) Deep learning for automated skeletal bone age assessment in X-ray images. Med Image Anal 36:41–51
18. Murakami S, Hatano K, Tan J, Kim H, Aoki T (2018) Automatic identification of bone erosions in rheumatoid arthritis from hand radiographs based on deep convolutional neural network. Multimed Tools Appl 77(9):10921–10937