Erzincan Üniversitesi
Fen Bilimleri Enstitüsü Dergisi
2023, 16(3), 940-969
ISSN: 1307-9085, e-ISSN: 2149-4584
Derleme Makalesi

Erzincan University
Journal of Science and Technology
2023, 16(3), 940-969
DOI: 10.18185/erzifbed.1211547
Review

Machine Learning/Deep Learning in Rheumatological Screening: A Systematic Review

Zehra Aysun Altikardes[1], Emre Canayaz[2*], Alparslan Unsal[3]

1 Vocational School of Technical Sciences, Marmara University, Istanbul, Turkey
2 Department of Electrical and Electronics Engineering, Institute of Pure and Applied Sciences, Marmara University, Istanbul, Turkey
2 Department of Radiology, School of Medicine, Adnan Menderes University, Aydin, Turkey

Received:29/11/2023, **Revised:**08/07/2023 , **Accepted:**10/07/2023 , **Published:** 31/12/2023

Abstract
Machine learning and deep learning techniques have been used in many fields, especially automatic image processing techniques, in recent years. In light of these developments, it has become inevitable to develop applications in the medical field. This study focuses on the past few years of research using machine learning and deep learning methods in the context of image processing in the field of rheumatology. This review provides researchers with the latest information on the use of deep learning and machine learning and inspires them to generate new ideas in their research by analyzing image processing systems performed by these artificial intelligence methods. In the proposed systematic review, 28 articles covering the application of deep learning and machine learning methods in the domain of rheumatology with the aim of digital image processing in the last 18 years were evaluated. Experiments emphasize that machine learning and deep learning methods provide significant segmentation accuracy and better case classification accuracy for various rheumatologic diseases like rheumatoid arthritis, osteoarthritis, and ankylosing spondylitis. Lastly submitted review presents possible different research ideas for related researchers to concentrate on for their future studies.

Keywords: Artificial Intelligence, Machine Learning, Deep Learning, Rheumatology

Romatolojik Görüntülemede Makine Öğrenimi/Derin Öğrenme: Sistematik Bir İnceleme

Öz

Son yıllarda otomatik görüntü işleme teknikleri başta olmak üzere birçok alanda makine öğrenmesi ve derin öğrenme teknikleri kullanılmaktadır. Bu gelişmeler ışığında medikal alanda uygulamaların geliştirilmesi kaçınılmaz hale gelmiştir. Bu çalışma, romatoloji alanında görüntü işleme bağlamında makine öğrenimi ve derin öğrenme yöntemlerini kullanan son birkaç yıldaki araştırmalara odaklanmaktadır. Bu inceleme, araştırmacılara derin öğrenme ve makine öğreniminin kullanımı hakkında en son bilgileri sağlamayı ve yapay zekâ yöntemleri tarafından gerçekleştirilen görüntü işleme sistemlerini analiz ederek araştırmalarında yeni fikirler üretmeleri için onlara ilham vermeyi hedeflemektedir. Önerilen sistematik incelemede, son 18 yılda dijital görüntü işleme amacıyla derin öğrenme ve makine öğrenmesi yöntemlerinin romatoloji alanında uygulanmasını kapsayan 28 makale değerlendirilmiştir. Deneyler, makine öğrenimi ve derin öğrenme yöntemlerinin, romatoid artrit, osteoartrit ve ankilozan spondilit gibi çeşitli romatolojik hastalıklar için önemli segmentasyon doğruluğu ve nispeten daha iyi vaka sınıflandırma doğruluğu sağladığını vurgulamaktadır. Son olarak gönderilen inceleme, ilgili araştırmacıların gelecekteki çalışmaları için odaklanmaları için olası farklı araştırma fikirleri sunmakta ve araştırmacıların kullandığı veri tabanları hakkında bilgi sağlamaktadır.
Anahtar Kelimeler: Yapay Zekâ, Makine Öğrenmesi, Derin Öğrenme,Romatoloji

*Corresponding Author: emre.canayaz@marmara.edu.tr

Zehra Aysun Altikardes, https://orcid.org/0000-0003-3875-1793
Emre Canayaz, https://orcid.org/0000-0002-3695-3642
Alparslan Unsal, https://orcid.org/0000-0002-0407-159X

1. Introduction

Medical image processing is a constantly changing discipline. Over the last ten years, the morphology of medical images, image processing in full-color depths, image data lessening, image recognition, and knowledge-based medical image analysis systems have all gotten much attention `(Zharkova, 2007)). Therefore, artificial intelligence (AI) enables researchers to cope with the massive amounts of data collected. AI is a general phrase that uses a computer to imitate intelligent behavior with little human interaction (Hamet & Tremblay, 2017). The early applications of AI in the medical area, according to historical texts, happened primarily in the 1960s and 1970s (Becker, 2019). In 1959, Arthur Samuel was the first to utilize machine learning to teach a computer to play checkers using human guidance (Samuel, 2000).

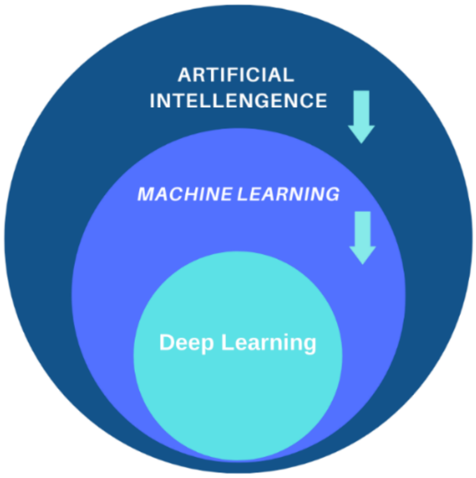


Figure 1: Relation between Artificial Intelligence, Machine Learning, and Deep Learning

Machine learning (ML) is a subsection of AI specializing in statistics and computer science. ML works on learning relationships from datasets collected by computer algorithms and uses example data or previous experience to improve a performance criterion (Khan, 2002). There are many different definitions of deep learning (DL) (Bengio & LeCun, 2007; Zhang, Yang, Lin, Ji, & Gupta, 2018). Still, we can summarize that DL is based on discovering several layers of representation, expecting that higher-level characteristics can reflect the data's abstract semantics. The more abstract terms learned from a deep network, the more resilient they should be against intra-class heterogeneity. The usage of convolutional architectures is a significant component of DLs effectiveness in image categorization. DL is based on discovering several layers of representation in the hopes that higher-level characteristics can capture the data's abstract semantics. A deep network's abstract expressions are intended to deliver better results (Algan & Ulusoy, 2021; Bengio & LeCun, 2007; Bressem et al., 2020; Kräter et al., 2021; LeCun & Bengio, 1995). In a word, deep learning is a machine learning approach that employs a deep neural network (Figure 1), which is an artificial neural network (ANN) containing two or more hidden layers with distinct characteristics (Figure 2) (Kim & Tagkopoulos, 2019).

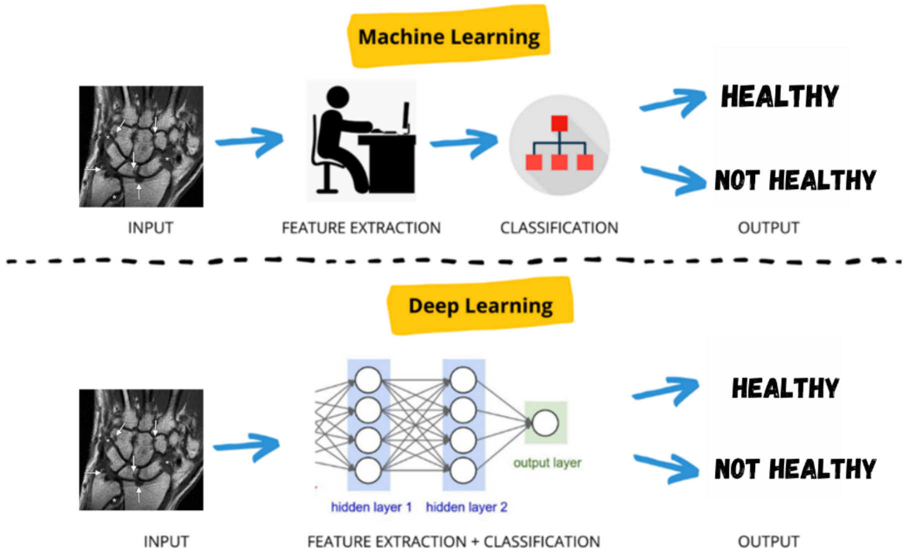


Figure 2: General workflow of ML and DL Networks. As stated in the figure, in ML applications, feature the user must define extractions, then the system performs classification. On the other hand, in DL applications, all feature extraction, and classification steps are automated by the system itself.

Due to the features of DL and ML to detect and process different data and establish connections between them, it finds use in various fields like text manipulation, speech recognition, and image classification (Deng et al., 2013; Devlin, Chang, Lee, & Toutanova, 2018; Kowsari et al., 2017; J. Liu, Chang, Wu, & Yang, 2017; Majumder, Poria, Gelbukh, & Cambria, 2017; Nanni, Costa, Aguiar, Silla Jr, & Brahnam, 2018). In recent years, ML and DL models have become more prevalent in applied medicine in biomedical research and healthcare, such as image labeling, image annotation, segmentation, data harmonization, cancer diagnosis, and gene expression data analysis. Because of these features, DL and ML systems are becoming more popular in rheumatology imaging studies, too.

The main objective of this review is to provide researchers with a document in rheumatology that offers them the latest information on the use of DL and ML and inspires them to generate new ideas in their research by analyzing the image processing systems performed by these AI methods. This study's readers will keep track of recent trends in AI usage in rheumatology image processing and find a table with the data and system features used in these systems. Moreover, some questions such as" Do we need AI for image processing rheumatology?", "Is it beneficial to use DL and ML in image processing" and "What kind of DL/ ML technologies do we need?" were tried to be answered.

2. Material and Methods

The presented literature review was conducted according to the PRISMA methodology (Page et al., 2021). In identifying sources for this systematic literature review, varying large citation databases such as the National Library of Medicine, the Cochrane Library, Scopus, Web of Science, IEEE Explore, ELSEVIER, Springer, and Science Direct were used. First of all, the terms Artificial Intelligence, Machine Learning, Deep Learning, Rheumatology, and Image Processing were combined with AND commands with various combinations. Afterward, combined phrases were searched for establishing a list of research articles to form a primary source. The sources used in the presented review have the most recent publication dates, going back no further than 2004 and including studies published in 2021.

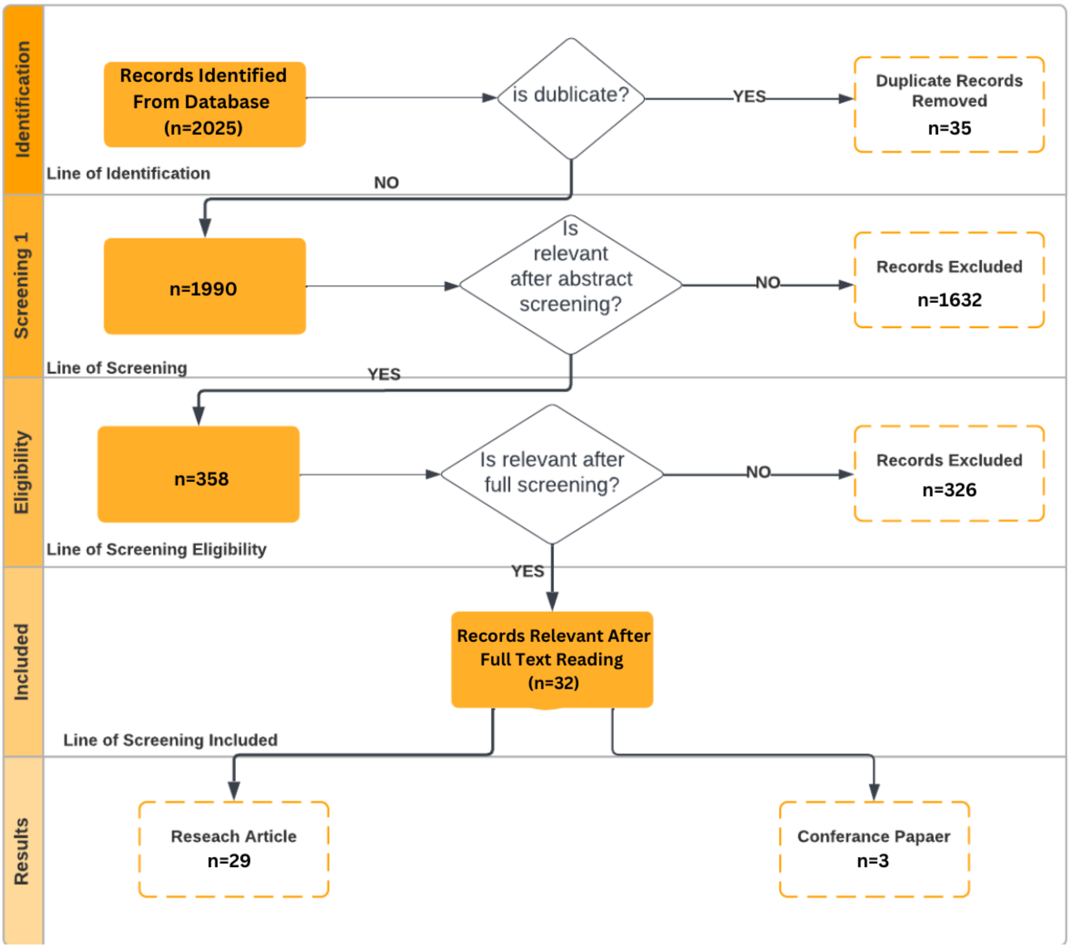


Figure 3: *PRISMA methodology followed in the presented review*

Figure 3 presents a flowchart of how the PRISMA methodology was used for the presented study. The analysis and filtering steps of the 2025 publications accessed here are shown in detail. The eligibility criteria of the review were as follows:

- Research should be written in English and must not be published before 2004
- Machine learning algorithms and/or deep learning methodology should be applied in the rheumatology area
- Research should use medical images to detect anatomical structures and/or target rheumatologic disease.
- Research should contribute to the primary purpose of the presented literature review and has an impact on the related area of study.
- Research should have been peer-reviewed by the related authorities.

The reasons for exclusions were as follows:

- Full text could not be obtained
- Research was presented only as abstract
- Main focus of the research was not eligible for the presented review
- Research that addressed only theoretical concepts
- Research that did not provide detailed information about the applied method

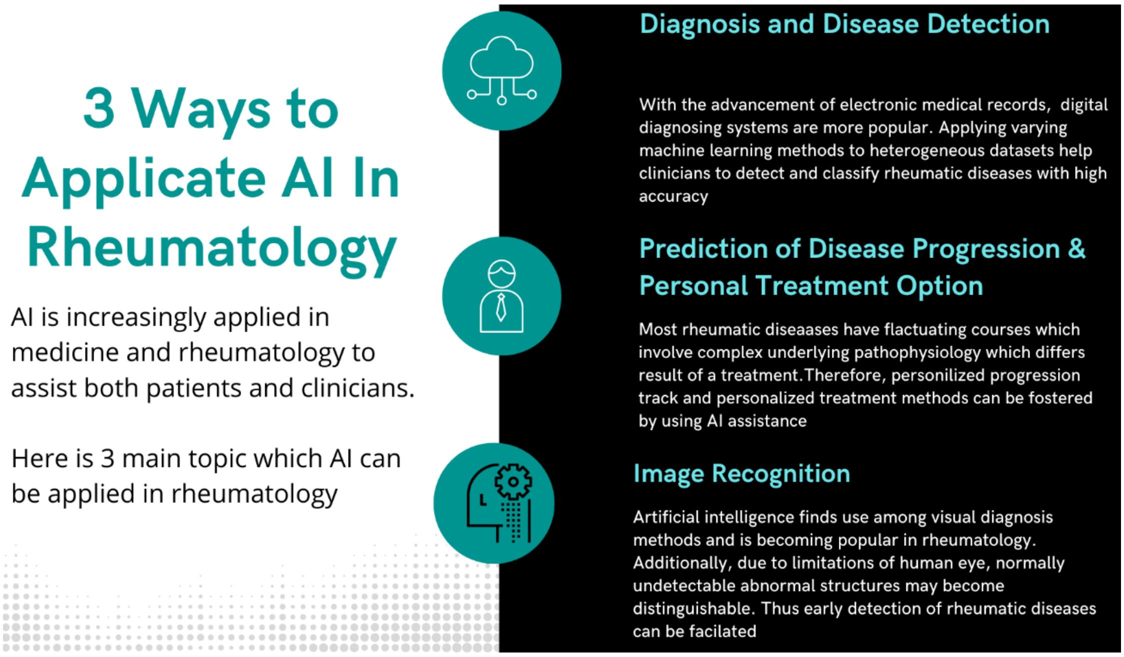
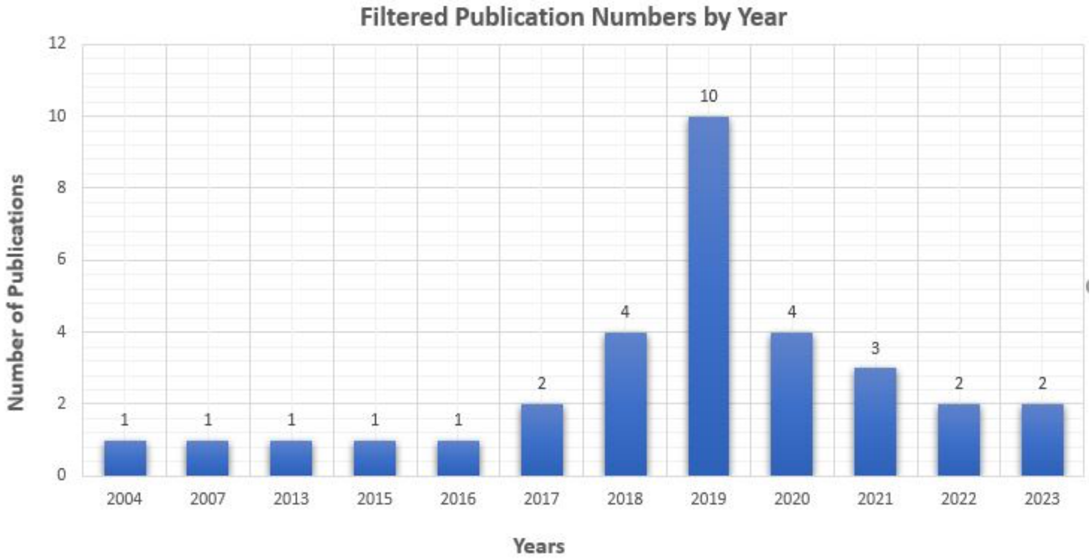


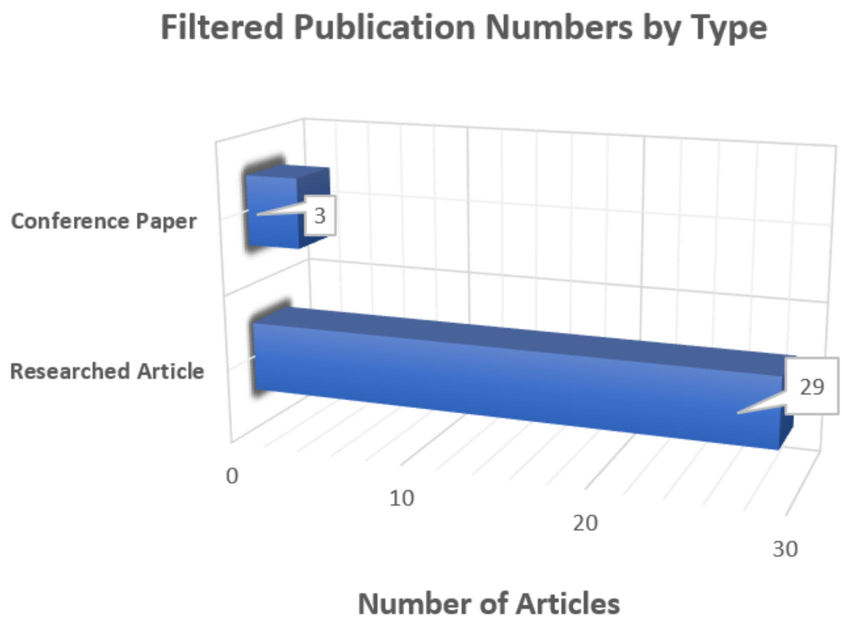
Figure 4: Three different ways to applicate AI technologies in rheumatology. 1. Diagnosing and Disease Detection 2. Prediction of Disease Progression and Personal Treatment Options 3. Image Recognition. The first two main headlines were eliminated to perceive novel AI practices in Image Processing Area.

3. Results and Discussion

In this section, filtered publications were examined. Also, in Table 1, related studies are summarized detailly and listed according to their publication date. General Rheumatologic AI studies could be grouped as in Figure 4. When the eligibility stage of the PRISMA methodology was completed, 32 studies had been identified for review. Thirty- two of the filtered studies can be summarized as research articles, and three as conference papers. Results were summarized in Figure 5. Additionally, the data source of the filtered studies was presented in Table 1.



a.



b.

Figure 5: Distribution of results obtained after filtering studies by year (a) and type (b)

When the datasets used in the studies were examined, it was observed that most of them preferred to use the clinical datasets. However, it has been determined that some studies prefer datasets that can be accessed by other researchers (Table 2). Among the online datasets studies, the Osteoarthritis Initiative dataset (OAI) was the most preferred one.

In addition, another study (Carano et al., 2004) devised a technique for quantifying bone abnormalities in MRI scans of RA patients. According to a system based on machine learning algorithms, k-NN showed the most potential for multicenter clinical trials. It was also argued that supervised classifiers are more sensitive to patient variations and that training data might lead to errors. Another research group (Tripoliti, Fotiadis, & Argyropoulou, 2007) tried to develop an automated system that segments and quantifies inflammatory tissue of the hand in

RA using enhanced T1- weighted MRI images. A fuzzy-based system was used in their study, and both manual and automated segmentation was applied to related images. Their system managed to achieve 97.71% sensitivity while classifying MRI images of the patients. Prasoon's study (Prasoon et al., 2013) tend to develop a system that calculates cartilage due to OA from knee MRI scans using both CNN and ML algorithms. According to their claim 3D, multi-scale features performed better with 99.93% accuracy and 81.92% sensitivity.

Another research team (Segen, Kulbacki, & Wereszczyński, 2015) used ultrasound images for assessing synovitis with the assistance of image point feature description and k-NN and SVM. The proposed unsupervised learning system managed to locate joints and bones by registering structural descriptions of the joint region. The clustering approach employs an inter-model distance metric, which is defined as the minimum of the objective function and quantifies a structural description disagreement. Another study group (Antony, McGuinness, O'Connor, & Moran, 2016) asserted that they created a deep CNN structure that uses MRI data to quantify the severity of knee OA automatically. In comparison, CNN features inferred from the fine-tuned BVLC reference CaffeNet yield better classification accuracy. Moreover, it claimed that SVM based automated method detected and extracted the knee MRI images. Researchers (Ashinsky et al., 2017) developed an ML-based system for detecting OA development in human cartilage systems using MRI images. T2- weighted MRI images were used, and 75% accuracy was achieved. Additional to that research, in 2017, another researcher (Xue, Zhang, Deng, Chen, & Jiang, 2017) trained a CNN-based DL system to detect OA from hip X-Ray images with 95.00% sensitivity and claimed that DL could be used in the practice of intelligent medical image diagnosis. After 2018, DL-type systems' numbers increased drastically (Table 1). Tiulpin (Tiulpin, Thevenot, Rahtu, Lehenkari, & Saarakkala, 2018) proposed a DL model based on CNN and Siamese neural network architecture to spot knee OA from plain radiographs automatically. To achieve this goal, they presented attention maps of the ROI with 66.71% multiclass accuracy. After a long period, another researcher considered the fuzzy clustering method to calculate bone marrow edema (BME) from wrist MRI images with early arthritis and claimed that false-positive events occurred due to not suppressing fat and fat tissue identified before the examination (Aizenberg et al., 2018). Another automated CNN-based system was developed to perceive cartilage lesions in the knee joint was developed by Lui (F. Liu et al., 2018). The system analyzed MRI images of individuals with different cartilage degeneration levels. According to their claim, when compared to a clinical radiologist, the cartilage lesion identification method exhibited greater sensitivity but poorer specificity at the optimum Youden index. Another technology was shown to help radiologists by offering a unique quantitative strategy for automatically diagnosing bone degeneration on hand radiographs with an accuracy of 80.5 percent (Murakami, Hatano, Tan, Kim, & Aoki, 2018). Researchers claimed that they used a pre-trained image classification network due to the hardness of the collection of medical images. Another study group (Norman, Pedoia, Noworolski, Link, & Majumdar, 2019) advanced a DL system using U-net and DenseNet methods to detect and classify OA according to KL grading. They asserted that U-net managed to label knee joint regions with 98.3% accuracy automatically. Moreover, DenseNet architecture classified no OA, mild, moderate, and severe cases with 83.7%, 70.2%, 68.9%, and 86.0%, respectively. Hirvasniemi (Hirvasniemi et al., 2019), Elastic Net Regularization was performed on the hip, pelvic X-Ray images to predict the incidence of rHOA or total hip replacement (THR) over ten years. Orenge's study (Orange et al., 2018) used doppler US images on VGG-16 and Inception-v3 DL architectures to determine Omeract-Eular Synovitis Score (OESS). Researchers compared the output of their system with an experienced radiologist and claimed that the proposed system had the highest accuracy due to the comparison.

Even though the lack of image numbers, Schaefer (Schaefer, Krawczyk, Doshi, & Merla, 2013) managed to develop an ML ensemble-based classification model for the automatic detection of scleroderma capillary patterns with an accuracy of 83.3%. In their research, nail fold capillaroscopy images were used as input. Using US images belonging to the MEDUSA database as input to the CNN-based DL system, Hemalatha (Hemalatha, Vijaybaskar, & Thamizhvani, 2019) managed to classify different grades of RA and differentiate the other structural issues such as skin and bone. Additional to US-based systems, another system was developed using metacarpal head US images to provide a method for automatically selecting bone structures and regions (Fiorentino, Moccia, Cipolletta, Filippucci, & Frontoni, 2019). The related system was developed using DL networks such as VGG16 architecture and Inception-v3 architecture. According to their claim, VGG16 architecture had better performance than Inception-v3 and proposed a system that could aid in the diagnosis process and training of young residents. Brahim's study group (Brahim et al., 2019) created a ML-based computer-aided diagnostic system to detect early knee OA with X-ray images provided from the OA initiative database. In the proposed system, a Fourier Filter was applied to images. After that, multivariate linear regression was applied to reduce variability between subjects with OA and those classified as healthy. Their system's OA detection accuracy was 82.98%. A ML and DL hybrid system (Hirano et al., 2019) was presented to determine finger joint damage caused by RA. The proposed system ignored intercarpal joint damage due to better performance in distinguishing between other joints. Model recognized proximal interphalangeal joints (PIP), interphalangeal joints (IP), and metacarpophalangeal joints (MCP) with 95.3% accuracy.

Gray-level co-occurrence matrices (GLCM) and local binary patterns to generate input features of CT sacroiliac joint images to k-NN and random forest algorithms and Inception-v3 DL architecture. According to their research, age is an important factor in disguising AS-based tissue differences with age-based erosions. They claimed they managed to differentiate the difference with an AUC score of 0.96 with eight-fold cross-validated input data. Moreover, Faleiros (Faleiros et al., 2020) built a system to classify active inflammatory sacroiliitis from MRI scans utilizing machine learning techniques like SVM and MLP and a feature selection approach to minimize the input's dimensionality. The best performance was observed with the Wrapper feature selection method with a 10-fold cross-validated training set (sensitivity = 100%, specificity = 95.6%, and accuracy = 84.7%). The model proposed by Brui (Brui et al., 2020) investigated cartilage segmentation in wrist MRI images using the U-net DL approach.

Author/ Year	Type	Objective	Input	Artificial Intelligence Method	Results
(Carano et al., 2004)	Research Article	Develop a technique for measuring bone changes in RA patients' hands that needs little user engagement.	13 RA patients of at least six months duration and less than five years before	Three different ML classifiers were used: 1. Multivariate Gaussian s(MVG), 2. k-nearest neighbor (k-NN), 3. K-means (KM).	**1.** Temporal categorization rates were 90.1% for KM, 89.5% for MVG, and 86.7% for k-NN. **2.** k-NN has a high potential for application in multicenter clinical studies. **3.** Because a considerable variation from the norms of the training data might result in errors, k-NN and MVG are more sensitive to changes among patients and exams as supervised classifiers.
(Tripoliti et al., 2007)	Research Article	Using contrast-enhanced T1-weighted MRI, an agile technique for segmenting and measuring inflammatory tissue of the hand in rheumatoid arthritis patients was developed.	300 MRI images from 25 rheumatoid arthritis (RA) patients, and after one year follow up 204 MRI images from 17 of them	Fuzzy C-means Algorithm	**1.** The sensitivity and positive predictive rates, respectively, are 97.71% and 83.35 %t. **2.** Inflammation was measured before and after treatment, and a manual segmentation comparison was made. **3.** The variations in absolute percentage between automatic and manual segmentation of 17 patients range from 0.12 to 37.98%.
(Prasoon et al., 2013)	Research Article	Calculating cartilage due to osteoarthritis causes disability	114 knee MRI Scans	Tripilar CNN and kNN	**1.** Although the proposed Model uses 2D features at a single scale, it performs better than a state-of-the-art method using 3D multi-scale features **2.** For Triplanar CNN, DSC 0,8249 (± 4.26), Accuracy 99,93%, (± 1,86%), Sensitivity 81.92% (± 7.62%) **3.**For State-of-the-art , DSC 0,8135 (± 4.87), Accuracy 99,92%,(± 2.31%) , Sensitivity 80.52% (± 8,95%)

(Segen et al., 2015)	Research Article	The development of an automated method for assessing synovitis from ultrasound pictures is underway to decrease medical expenses and enhance patient care.	Synthetic data from Project MEDUSA	Image point feature descriptor (like SURF and SIFT) for future vector for classifier and k-NN, nearest description cluster, and SVM	**1.** Proposed system managed to locate joints and bones by registering structural descriptions of the joint region. **2.** A preliminary result is provided, which comprises a description of a registration approach that iteratively improves registration quality and an example of its use using synthetic data care.
(Antony et al., 2016)	Research Article	Develop a technique to automatically estimate the severity of knee osteoarthritis (OA) from radiographs using deep CNN.	Osteoarthritis Initiative dataset (0.E.1) Kellgren & Lawrence (KL) graded 4446 radiographs)	VGG 16-layer net, VGG-M-128, CaffeNet pre-trained DL Models	**1.** The primary contributions of this article, according to the researchers, are the use of CNNs and regression loss to estimate knee OA severity. **2.** CNN characteristics inferred from the BVLC reference that has been fine-tuned CaffeNet outperforms the state-of-the-art classification accuracy. **3.** A SVM-based technique recognizes and separates the knee joints from knee OA radiographs automatically.
(Ashinsky et al., 2017)	Research Article	To see how well a machine learning system can detect in vivo MRI of human articular cartilage for OA development.	The osteoarthritis initiative (OAI) control and incidence cohorts (0.E.1.,0.C.2) T2-weighted knee images were used to identify 68 individuals.	ML algorithm weighted neighbor distance using compound hierarchy of algorithms representing morphology (WND-CHRM),	**1.** ML algorithm applied to T2 maps can give valuable prognostic information for the progression of OA. **2.** Sensitivity, specificity, and accuracy were 74%, 76%, and 75%, respectively. **3.** The 10^{th}-order polynomial registration model required many hours of processing time per subject.

(Xue et al., 2017)	Research Article	Train a CNN Model which automatically diagnoses OA from hip X-Ray images	420 hip X-ray images	CNN	**1.** The proposed model attained a 95.0 percent sensitivity and a 90.7 percent specificity balance. **2.** In comparison to chief doctors, the model has a 92.8 percent accuracy. **3.** Compared to doctors, the model's performance and system sensitivity are lower. **4.** Researchers claimed that DL could be used. in the practice of intelligent medical picture diagnosis
(Tiulpin et al., 2018)	Research Article	Simple radiographs can be used to diagnose knee OA while also offering transparency into the decision-making process of clinicians.	3,000 subjects (5,960 knees plain radiography)	CNN, Siamese deep neural networks	**1.** Displaying attention maps that show which regions of interest affected the network's choice. **2.** Researchers have provided a standardized dataset for knee X-ray OA diagnosis algorithms. **3.** Radiographic OA diagnosis with an average multi-class accuracy of 66.71%
(Aizenberg et al., 2018)	Research Article	The main goal was to explore if bone marrow edema (BME) could be automatically quantified on MRI of the wrist in persons with early arthritis.	573 early arthritis patients' MRI images (mean age: 54.7 years); 354 females (mean age, 53.0 years); 219 males (mean age, 57.5 years)	Fuzzy Clustering to calculate the BME score	**1.** The Pearson correlation between quantitative and visual BME values was r=50.83, P <0.001, across 485 individuals. **2.** BME quantification on MRI of the wrist might be an excellent alternative to visual scoring. **3.** By simply adding these areas of interest to the atlas, this framework may be readily expanded to include more regions of the wrist and other joints. **4.** Even though false-positive events owing to fat suppression are uncommon, they must be recognized before this study may be used to avoid false positives.

(F. Liu et al., 2018)	Research Article	The main purpose was to detect cartilage lesions in the side of a knee joint using a fully automated DL base system	175 patients MRI image set with T2-weighted, fat suppression	CNN	**1.** Obtained high diagnostic performance for detecting cartilage degeneration (for evaluation 1, Sensitivity:84.1% , Specificity: 80.5% , ROC:0.917 ; for evaluation 2 Sensitivity:85.2% , Specificity: 87.9% , ROC:0.914) **2.** When compared to a clinical radiologist, the cartilage lesion identification method exhibited greater sensitivity but poorer specificity at the optimum Youden index.
(Murakami et al., 2018)	Research Article	To assist radiologists by presenting a new quantitative approach for identifying bone degeneration on hand radiographs automatically.	129 cases do DL learning and hand radiographs of 30 patients with RA for testing performance	CNN	**1.** Automatically performs a segmentation process to extract the region of interest of phalanges regions. **2.** Proposed system classifies the bone erosion of the related areas with 80.5% accuracy and 0.84% false-positive rate. **3.** Researchers used a pre-trained image classification network due to the hardness of collecting medical images
(Norman et al., 2019)	Research Article	Develop a fully automated algorithm for detecting OA and classifying according to Kellgren Lawrence (KL) grading.	OAI dataset (age = 61.2 ± 9.2 years, BMI = 32.8 ± 15.9 kg/m2, 42/58 male/female split)	U-Net for joint Detection and DenseNets for grading	**1.** In 1000 randomly selected instances, the U-net for knee joint labeling was successful 98.3% of the time. **2.** No OA, mild, moderate, and severe OA testing sensitivity rates were 83.7, 70.2, 68.9, and 86.0%, respectively, according to DenseNets. The specificity rates were 86.1, 83.8, 97.1, and 99.1%, respectively. **3.** The single-blind research with internal radiologists demonstrates the KL classification's inter-observer reliability, which has been reported to vary from 0.51 to 0.8. As a result, the claimed classification accuracies might be up to 30% higher.
(Hirvasniemi et al., 2019)	Research Article	The purpose of this study was to see how well radiography-based bone texture characteristics in the proximal femur and acetabulum predicted the incidence of radiographic	987 hips pelvic radiographs	Elastic net (ML) was used to predict rHOA	**1.** Bone texture analysis, according to the researchers, Bone texture analysis gives new information for predicting the incidence of rHOA or total hip replacement (THR) over the latest data for predicting the incidence of rHOA or total hip replacement (THR) over ten years. **2.** At the 10-year follow-up, 435 (44%) of the 987 hips that did not have rHOA at the start of the study had rHOA. 471 (71%) of the 667 hips with JSN grade 0 had JSN grade >1 at the 10-year follow-up. 526

		hip osteoarthritis (rHOA) over ten years.			(86%) of the 613 hips with OST grade 0 had OST grade > 1 at the 10-year follow-up. **3.** The AUCs for models that included age, gender, and body mass index (BMI) to predict incident rHOA, JSN, and OST were 0.59, 0.54, and 0.51, respectively, for models that included age, gender, and BMI to predict incident rHOA, JSN, and OST.
(Orange et al., 2018)	Research Article	Using the OMERACT-EULAR Synovitis Scoring (OESS) system, see whether neural network architecture can be used to evaluate RA disease activity on Doppler US pictures.	1342 Doppler Ultrasound Image	2 CNNs (VGG-16 Architecture and Inception-v3 Architecture)	**1.** The researchers stated that this was the first CNN method that used the OMERACT-EULAR Synovitis Scoring System to classify arthritic disease activity. **2.** When compared to an expert rheumatologist, the neural network evaluating healthy/diseased scores had the highest accuracy, with a sensitivity of 0.864 and 0.875 and specificity of 0.864 and 0.864, respectively.
(Schaefer et al., 2013)	Conference Paper	Designing an automated approach for analyzing and categorizing nail fold capillaroscopy (NC) images.	16 Subjects, 60 NC Images	Machine Learning Ensemble Classification	**1.** From texture features from microscope images of Scleroderma Capillary Pattern was identified using the Ensemble Classification Method **2.** System classified images with 83.3% accuracy.
(Hemalatha et al., 2019)	Research Article	It was proposed that automatic identification of the various stages of arthritis and detection of other structural areas such as skin and bone be implemented.	MEDUSA database. 276 ultrasound (US) images with manually annotated	CNN	**1.** Because the severity of arthritis differs from person to person, segmentation of distinct anatomical areas such as skin, bone, and the joint is required for effective disease progression discrimination and differentiation. **2.** Average Accuracy: 88.52%, False Positive Rate: 1.41%

(Fiorentino et al., 2019)	Conference Paper	This study aims to provide a method for automatically selecting informative US rheumatology pictures.	214 balanced metacarpal head US images	VGG16 and Inception V3 CNNs	**1.** The best results were obtained using VGG16. (ROC:0.90) **2.** The findings suggest the feasibility of using this technique in actual clinical practice to aid in the diagnosis process and the training of young residents. **3.** The proposed approach might be a valuable tool for selecting important frames that computer-assisted algorithms can analyze to help with diagnosis. It could be used in a variety of anatomical districts and imaging modalities.
(Brahim et al., 2019)	Research Article	Develop a computer-aided diagnostic (CAD) system that combines knee X-ray imaging and machine learning algorithms to detect early knee OA.	1024 X-ray knee images from the OAI database	Ml classifiers: Naive Bayes and random forest	**1.** Images were subjected to a Fourier filter before being subjected to multivariate linear regression (MLR) to decrease variability between OA and healthy participants. **2.** An accuracy of 82.98%, a sensitivity of 87.15%, and a specificity of up to 80.65% were reached in the identification of OA. **3.** The suggested approach, according to the researchers, can diagnose OA early.
(Hirano et al., 2019)	Research Article	To assess radiographic finger joint deterioration in RA, use a deep-learning model.	Among 216 radiographs of 108 patients with RA	Joint detection by using ML and determining the score of destruction with CNN	**1.** The patients who took part in this trial were prone to joint damage. **2.** With a sensitivity of 95.3%, the model correctly identified PIP, IP, and MCP joints. **3.** The model tended to overlook intercarpal joints. Because the model was trained using images of the PIP, IP, and MCP joints simultaneously, it's probable that it won't be able to tell the difference between them.

(Shenkman et al., 2019)	Research Article	Develop a system that diagnoses and grades Sacroiliitis from pelvis CT scans automatically to enable early diagnosis (SIJ-grade)	242 anonymized axial CT scans from the Sheba Medical Center	Random Forest for detecting region of interest (ROI) U-Net Classifier, custom slice CNN classifier	**1.** Proposed algorithm had two-phase: An offline training phase and an online classification base **2.** 91.9% and 86% classification accuracy with 95% and 82% sensitivity and 0.97 and 0.57 AUC respectively. **3.** Researchers claimed that the proposed system could grade and diagnose sacroiliitis.
(Pedoia et al., 2019)	Research Article	The ability of deep-learning algorithms to detect and stage the severity of meniscus and patellofemoral cartilage lesions in OA and anterior cruciate ligament (ACL) patients was investigated.	1478 MRI studies with various stages of OA	For automatic segmentation: 2D U-Net For Severity staging: 3D-CNN	**1.** Meniscus lesion detection had an 89.81 percent sensitivity and an 81.98 percent specificity, whereas cartilage detection had an 80.0 percent sensitivity and an 80.27 percent specificity. **2.** The suggested model was tested on a dataset approximately ten times bigger and more diversified than the previous one, which included OA and ACL patients before and after reconstruction. **3.** Researchers claim that although it is too early to make any statements regarding the shift in workflow or comment on how these procedures will directly assist patients, exploring the direction they envisage for these approaches is essential.
(Castro-Zunti, Park, Choi, Jin, & Ko, 2020)	Research Article	By analyzing computed tomography (CT) data, propose a statistical ML and DL-based classifier to predict erosion, an early Ankylosing Spondylitis sign	681 grayscale JPG images, each featuring a single sacroiliac joint	Gray-level co-occurrence matrices (GLCM) and local binary patterns to generate input features to machine learning algorithms (k-NN, random forest) and InceptionV3 CNN InceptionV3 backboned DL Model	**1.** Random forest classifiers outperform k-NN classifiers after 8-fold cross-validation, with average accuracy, recall, and area under receiver operator characteristic curve (ROC AUC) for erosion vs. young control patients of 96.0 %, 92.9 %, and 0.97, respectively, and 82.4 %, 80.6 %, and 0.91 for erosion vs. young control patients. 2. A DL classifier trained without limiting validation loss surpassed all (mixed young and elderly) control patients in terms of erosion, with cross-validation accuracy, recall, and ROC AUC of 99.0 %, 97.5 %, and 0.97, respectively.

(Faleiros et al., 2020)	Research Article	To classify active inflammatory sacroiliitis in magnetic resonance images, researchers tested the applicability of traditional ML models and feature selection methods.	56 sacroiliac joint MRI images	Support Vector Machine (SVM), the Multilayer Perceptron (MLP), and the Instance- Based Algorithm	**1.** ReliefF and Wrapper methods used to select important features of the image properties **2.** Segmentation and Selecting ROI processes were performed in Adobe Photoshop **3.** Based on 10-fold cross-validation using the training dataset, the MLP classifier employs six features selected using the Wrapper feature selection approach, and has the best performance, with sensitivity = 100 %, specificity = 95.6 %, and accuracy = 84.7 %.
(Brui et al., 2020)	Research Article	Examine the performance of a customized convolutional neural network (CNN) designed explicitly for segmenting wrist cartilage from 2D MR images.	11 subjects, 20 Multi-slide MRI Scan	CNN	**1.** In the wrist cartilage segmentation challenge, CNN architecture considerably beat the traditional image-based U-Net (DSC = 0.86 and 0.64, respectively). **2.** The proposed network exhibited resilience in the presence of numerous anatomical features and joint deformities that resemble or contrast with cartilage.
(Bressem et al., 2020)	Research Article	Using centrally scored images from two observational cohort studies, create and validate a DL artificial neural network to diagnose definite radiographic sacroiliitis.	Two different image datasets: Patients with Axial Spondyloarthritis: Multicountry Registry of Clinical Characteristics (PROOF) (n= 1553) and German Spondyloarthritis Inception Cohort (GESPIC) (n=458).	ResNet-50 CNN architecture	**1.** Deep artificial neural networks accurately identify definite radiographic sacroiliitis required for axSpA diagnosis and classification. **2.** The neural network performed remarkably well in diagnosing definite radiographic sacroiliitis, with AUCs of 0.97 and 0.94 for the validation and test datasets, respectively. **3.** For the cutoff considering both measures equally, the validation set's sensitivity and specificity were 88 % and 95 %, respectively, but the test set's findings were 92 % and 81 %.

(Maziarz, Krason, & Wojna, 2021)	Conference Paper	Develop a DL algorithm that learns to locate joints on X-ray pictures while diagnosing two types of joint damage: constriction and erosion.	RA2 DREAM Challenge Data 367 patients with four images per patient (both hands and feet)	CNN	**1.** The proposed DL model was able to pinpoint the location of joints and the severity of RA. **2**. The DL method can identify two types of joint damage: narrowing and erosion. **3.** To bridge the gap between classification and regression problems, researchers recommended using local label smoothing. **4.** Researchers employed annotations of the centroid of all joint sites to improve the effectiveness of training signals.
(Chaturvedi, 2021)	Research Article	To demonstrate a two-stage technique (DeepRA) that integrates object recognition, convolution neural networks, and attention to accurately forecast overall and joint level narrowing and erosion from patient radiographs.	RA2 DREAM Challenge Data	RetinaNet and CNN	**1.** Two-stage model which combines object detection method and CNN to predict narrowing and erosion of the joints according to SvH score system **2.** The researchers utilized an attention mechanism to assist the model in focusing on essential areas of X-Ray pictures of the hands and feet. **3.** To represent the model's predictive power, additional visualization was added on top of the medical images.
(Han et al., 2021)	Research Article	Creating a new segmentation network depending on DL to detect BME on hip joint MRI images	141 cases (101 for training and 40 for validation)	CNN and Resnet-50	**1**. 31 cases were correctly classified out of 40 test cases. **2**. The accuracy rate of the proposed system is 85.7% **3**. Automatic computer-based system analyzing MRI images are helpful and has the potential for grading AS and early detection of AS
Ribas er al., 2022 (Ribas, Riad, Jennane, & Bruno, 2022)	Research Article	Using the principles of complex network theory, textural characteristics associated with OA were extracted from radiographic knee X-ray images to train DL models for early detection of OA	688 Knee radiogrphas from Osteoarthritis Initiative dataset	DL Models: AlexNet, VGG, GoogleNet, InceptionV3, ResNet, DenseNet and EfficientNet	**1.** To choose the set of thresholds automatically, proposed model uses a feature vector which was formed Euclidean distance of pixel nodes. **2.** Proposed model's performance was compared with other OI DL based researches and it performed with 81.69% accuracy **3.** Researchers asserted that the results of the suggested technique are likely representative of early subchondral bone modifications that take place in the core compartment of the tibia prior to unambiguous radiological OA detection on X-Ray images.

Wu et al., 2022 (Wu et al., 2022)	Research Article	To determine if an automatic classification of RA metacarpophalangeal joint conditions in ultrasound images is feasible using DL in order to provide a more unbiased, computerized, and rapid method for detecting RA in a clinical setting.	1337 2D RA ultrasound images	DenseNet Based DL	**1.** The effectiveness of the proposed model in classifying synovial proliferation and distinguishing between healthy and pathological cases with high accuracy and AUC values was demonstrated by ROC analysis. **2.** The use of heat maps generated by the model's class activation mapping method identifies the most significant areas for classification, thereby shedding light on synovial joint conditions. **3.** The results, according to the researchers, imply the model's robustness and potential clinical utility in accurately identifying and classifying various synovial joint conditions scenarios.
Subhash et al. 2023 (Subhash & Kureshi, 2023)	Research Article	Researchers have proposed a CNN system designed for autonomously learning the characteristics of hand radiographs and estimating their classification based on a large data set in order to detect RA stages.	For the investigation, 290 photographs of hands of patients of varying proportions were collected. (130 Normal, 160 with RA)	ConvNet based DL	**1.** Researchers emphasized that the CNN classifier obtains greater precision than other classifiers (SVM, ANN) on both training and test sets, emphasizing its superior performance. **2.** CNN requires considerable storage space, time, and computational capacity for measurement, as it compares all stored training images to the test image. **3.** The article prioritized efficient performance measurement over teaching efficiency, recognizing that although CNN training is time-consuming, it facilitates rapid classification of new test data, which aligns with practical criteria for model application.
Aarthi et al. (KS, Selvakumar, Sathyamangalam, & Nadu, 2023)	Research Article	Utilizing CNN in DL, the primary objective was to develop a system capable of identifying rheumatoid arthritis.	7174 DR knee images which contains four different classes of RA	CNN	**1.** RA has significant detrimental effects on one's quality of life, resulting in suffering, disability, and premature mortality. **2.** The proposed method provides advantages over human assessors by employing image augmentation to better the quality of the training dataset, resulting in enhanced model performance during training. **3.** Regardless of possible drawbacks in explainability, DL was acknowledged as a high-quality strategy in dataset evaluation and AI, and the proposed network architecture demonstrated exceptional performance in classification and prediction, which was further enhanced by various processing operations.

Table 1: Data sources used in selected studies

	Dataset	Number of Studies	Studies Reference
Online	Osteoarthritis Initiative dataset (OAI) (https://nda.nih.gov/oai)	6	(Antony et al., 2016; Ashinsky et al., 2017; Brahim et al., 2019; Norman et al., 2019; Ribas et al., 2022; Tiulpin et al., 2018)
	RA2 DREAM Challenge Data Automated Scoring of Radiographic Damage in Rheumatoid Arthritis through Synapse ID (https://www.synapse.org/#!Synapse:syn20545111/files/)	2	(Chaturvedi, 2021; Maziarz et al., 2021)
	Multicenter Osteoarthritis Study (MOST) dataset (https://most.ucsf.edu/)	1	(Tiulpin et al., 2018)
	Cohort Hip and Cohort Knee (CHECK) (Wesseling et al., 2009)	1	(Hirvasniemi et al., 2019)
	Patients with Axial Spondyloarthritis: Multicountry Registry of Clinical Characteristics (PROOF) (Poddubnyy et al., 2021)	1	(Bressem et al., 2020)
	German Spondyloarthritis Inception Cohort (GESPIC) (https://clinicaltrials.gov/ct2/show/NCT01277419)	1	(Bressem et al., 2020)
Clinical		20	(Aizenberg et al., 2018; Brui et al., 2020; Carano et al., 2004; Castro-Zunti et al., 2020; Faleiros et al., 2020; Fiorentino et al., 2019; Han et al., 2021; Hemalatha et al., 2019; Hirano et al., 2019; KS et al., 2023; F. Liu et al., 2018; Murakami et al., 2018; Orange et al., 2018; Pedoia et al., 2019; Prasoon et al., 2013; Schaefer et al., 2013; Segen et al., 2015; Shenkman et al., 2019; Subhash & Kureshi, 2023; Tripoliti et al., 2007; Wu et al., 2022; Xue et al., 2017)

Despite the low input image number, they claimed that the CNN structure outperformed and demonstrated well match with the manual segmentation process (Sørensen–Dice similarity coefficient (DSC) =0.81). The study also investigated the manual wrist cartilage segmentation's inter-and intra-observer variability (DSC = 0.78-0.88 and 0.9, respectively). Researchers (Bressem et al., 2020) used the ResNet-50 DL architecture model on two different image data sets: Patients with Axial Spondyloarthritis: Multi-country Registry of Clinical Characteristics (PROOF) and German Spondyloarthritis Inception Cohort (GESPIC). The study's major purpose was to develop and test a DL-ANN for diagnosing definitive radiographic sacroiliitis. For the cutoff considering both measures equally, the validation set's sensitivity and specificity were 88 percent and 95 percent, respectively, but the test set's findings were 92% and 81%. Another algorithm for locating joints on X-ray pictures diagnoses two types of RA joint damage: constriction and erosion (Maziarz et al., 2021), and researchers used RA2 DREAM Challenge data as input. Also, the proposed multi-task CNN-based DL model had an alternate system for label smoothing, which integrates classification and regression information into a single loss. Similarly, another algorithm (Chaturvedi, 2021) uses RA2 DREAM Challenge data to predict joint level narrowing and erosion of RA patients' X-ray images. Researchers combined object detection and CNN architecture and developed a two-stage model.

In 2022, Using data from the Osteoarthritis Initiative, Ribas et al. (Ribas et al., 2022) was to develop a sophisticated network-based method for detecting knee osteoarthritis. Their proposed system employed a complex network architecture to analyze knee joint data and identified osteoarthritis-related patterns. The system integrated X-ray images and clinical variables to enhance detection and prognosis accuracy. According to researchers claim, results indicated that the proposed knee OA detection method was competitive and potentially promising with 81.69% accuracy. Likewise, in 2022 Wu et al. (Wu et al., 2022) for RA patients, a classification model based on DL neural networks and the OMERACT-EULAR synovitis scoring system was developed. The model yielded satisfactory results, with an AUC of 0.886 and an accuracy rate of 82.1% for classifying synovial proliferation and an AUC of 0.901 and an accuracy rate of 80.4% for classifying healthy versus malignant cases, respectively. The findings demonstrate the viability of a CNN architecture for evaluating joint proliferation levels in ultrasound images and provide a potential automated method for triaging patients and assessing RA conditions.

Current research indicates that CNN-based DL models have maintained their popularity through the year 2023. A model (Subhash & Kureshi, 2023) intended to automatically categorize hand x-rays of RA patients achieved a classification accuracy of 94.64%. Validated by medical professionals, the model identified normal and aberrant conditions. Compared with other CNN methods and classifiers demonstrate an improvement in accuracy, emphasizing the model's enhanced performance and precision. Another research team (KS et al., 2023) distinguished subtypes of RA using image processing systems based on CNN. Their dataset included images representing four distinct subtypes of RA. Study group claimed that key to this efficacy was the ability of deep learning to understand the structure of the underlying data.

3. Results and Discussion

3.1 Rheumatologic Diseases and Artificial Intelligence Union

One of the aims of this review was to inform about the latest trends in using AI in image processing in rheumatological diseases. Academic studies involving ML and DL participation are increasing rapidly. The most popular AI applications are classification processes based on scoring deformities in the relevant tissue and recognizing a particular rheumatic disease. Most ML studies focused on cartilage segmentation, detecting grates of RA and OA, and detecting bone edema and spondyloarthritis (Figure 6). The most used AI technique among them was DL networks with CNN structure (Figure 6). Also, according to the study objective and rheumatologic disease, varying images obtained from different modalities were used (Figure 6).

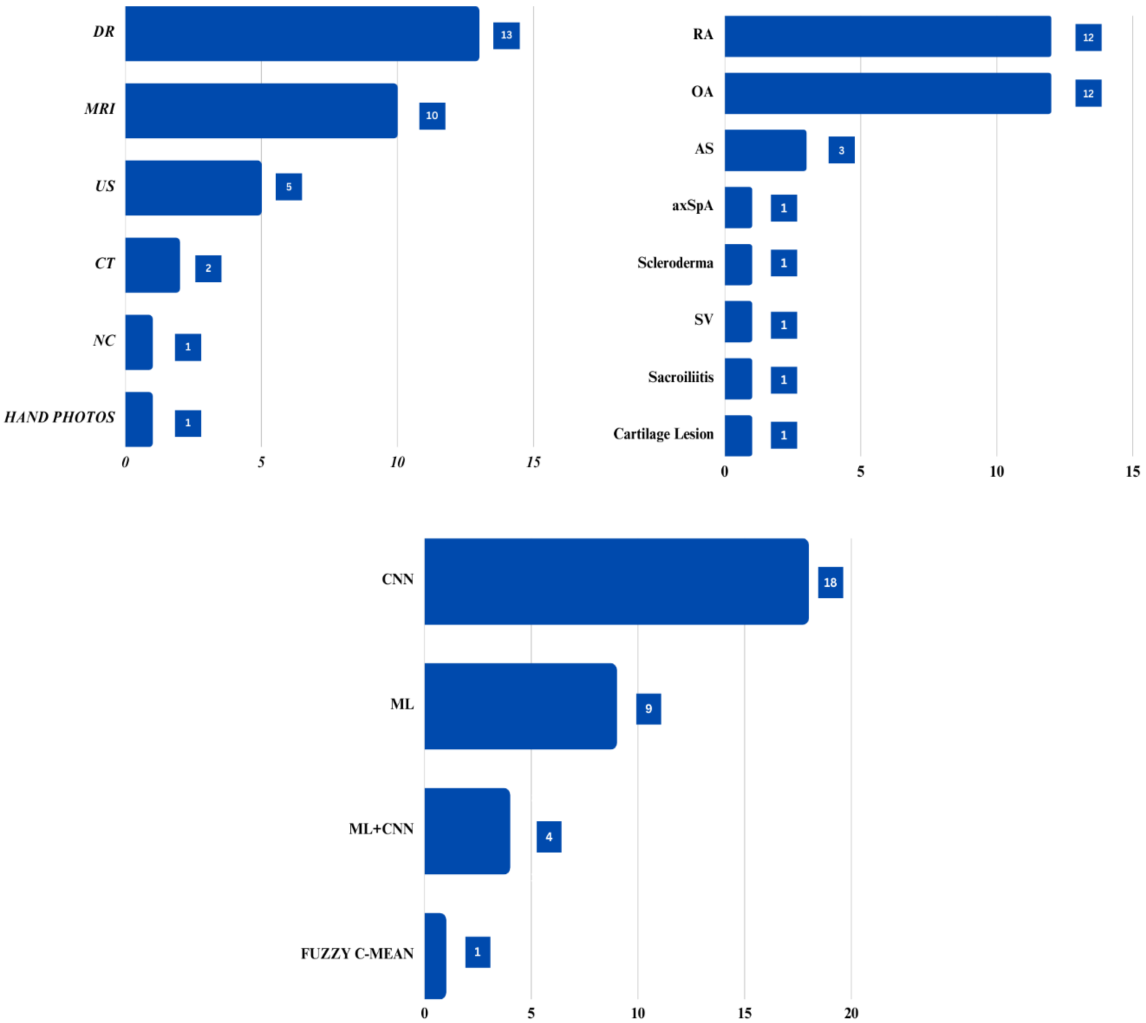


Figure 6: Article numbers according to the medical image source, studied disease, and used AI method, respectively (DR: Digital X-ray, MRI: Magnetic Resonance Imaging, US: Ultrasound, CT: Computed Tomography, NC: nail fold capillaroscopy)

A list of sources ordered by rheumatologic diseases was presented in Table 2. It was observed that DL methods, especially CNN-based networks, have found a wide application area and dominance in the AI applications in rheumatology (Figure 6). In addition, in some DL studies, it has been observed that instead of designing the system from scratch, it is used in transfer learning methods from pre-trained CNN models like Resnet-50, RetinaNet, InceptionV3, and VGG-16. Previously trained systems are adapted to rheumatologic imaging studies. Thus, a validated system is adapted to solve the relevant problem. Among the DL models, it has been shown by studies that especially the CNN system called U-net gives efficient results in medical segmentation processes. U-Net is a convolutional neural network developed at the Department of Computer Science at the University of Freiburg for segmentation in image processing studies in biomedical fields. The network architecture is based on the fully convolutional network and has been modified and expanded to work with fewer training images and provide more precise partitioning (Ronneberger, Fischer, & Brox, 2015). In terms of both rapid operation and good performance, it is expected to be employed more frequently in future research studies (Kayalibay, Jensen, & van der Smagt, 2017).

Table 2: List of filtered studies ordered by target rheumatologic disease and practiced methodology

Rheumatoid Arthritis	Fuzzy C-mean	(Fiorentino et al., 2019)
	ML	(Hirano et al., 2019)
		(Maziarz et al., 2021)
	CNN	(Carano et al., 2004)
		(Tripoliti et al., 2007)
		(Aizenberg et al., 2018)
		(Murakami et al., 2018)
		(Orange et al., 2018)
		(Wu et al., 2022)
		(Subhash & Kureshi, 2023)
		(KS et al., 2023)
	ML+CNN	(Chaturvedi, 2021)
Osteoarthritis	ML	(Brahim et al., 2019)
		(Hemalatha et al., 2019)
		(Pedoia et al., 2019)
	CNN	(Prasoon et al., 2013)
		(Antony et al., 2016)
		(Ashinsky et al., 2017)

		(Xue et al., 2017)
		(Tiulpin et al., 2018)
		(Norman et al., 2019)
		(Hirvasniemi et al., 2019)
		(Ribas et al., 2022)
	ML+CNN	(Brui et al., 2020)
Ankylosing Spondylitis	ML	(Faleiros et al., 2020)
	CNN	(Han et al., 2021)
	ML+CNN	(Castro-Zunti et al., 2020)
Axial Spondyloarthritis	CNN	(Bressem et al., 2020)
Cartilage Lesion	ML	(F. Liu et al., 2018)
Scleroderma	ML	(Schaefer et al., 2013)
Sacroiliitis	ML+CNN	(Shenkman et al., 2019)
Synovitis	ML+CNN	(Segen et al., 2015)

3.2 Diagnosing benefits of using Artificial Intelligence in Rheumatology

In the future, DL and ML applications can be used to detect diseases such as AS with a long diagnostic time. Early diagnosis means that the disease can be treated early. Thus, individuals can be protected from the adverse side effects of such conditions, and their living standards can increase positively. Rheumatic diseases also have a socio-economic impact on the patient, health system, and society. These types of disorders have adverse effects both clinically and economically. For instance, AS-related direct costs in the first-year amount to $1,775 vs. $2,674 of direct health costs for all causes; indirect costs are about $4,945 (Ward, 2002). According to research (Mennini et al., 2018) rheumatoid arthritis patients are also concerned about their capacity to work, social contacts, and family life. Reduced physical function is the most significant cost element from an economic standpoint (Annelies Boonen, Brinkhuizen, Landewé, van der Heijde, & Severens, 2010; A. Boonen & Mau, 2009).

3.3 Human vs Machine

As emphasized in the reviewed studies, AI systems have an essential collaborative potential to support specialists in rheumatology. So, what makes these advanced systems important for this field of application? First of all, with the implementation of electronic healthcare databases for medical images, reports, and electronic health records, AI-based systems can quickly establish relationships between different kinds of data that are difficult for humans to perceive (Bidgood

Jr, Horii, Prior, & Van Syckle, 1997). The success of the results depends on the statistical bases and the quality of the data used to train the relevant systems. This feature is their greatest strength and allows experts to provide unique information for their decision-making. On the contrary, unlike medical professionals, ML and DL technologies produce the output decisions they make purely for the limited application areas they are designed to work with. For this reason, although the developed systems perform as well as the experts, their success is limited only by the diversity of the data set they are trained on (Knight, 2017). Experts and machines have some advantages over each other. For example, while devices can work 24/7, the working time is relatively shorter by giving people full attention (Kansagra et al., 2016). Observer exhaustion is a natural part of radiology practice, and it's especially problematic in screening tests when the odds of getting a true positive are minimal. In this example, an AI system with a high negative predictive value may choose an "enriched" sample of cases for early review, likely to contain true-positive cases (Thrall et al., 2018). However, while machines must be trained with large datasets to produce accurate results when people are acquainted with small datasets, they can develop different strategies to solve extensive problems (Kansagra et al., 2016). In addition, experts can communicate more successfully and discuss the accuracy of the results they find while sharing the results they obtained with other experts and patients (Kushner & Lucey, 2005; Pahade et al., 2012). Thus, although artificial intelligence systems provide successful results, it is evident that the decisions they produce alone can only be used to support medical professionals' final decisions.

4. Conclusion

In the field of rheumatology, image processing studies with DL and ML approaches have mainly focused on Rheumatoid Arthritis and Osteoarthritis. However, the number of studies on diseases that negatively affect the quality of life of patients such as Ankylosing Spondylitis, especially in its advanced stages, is quite limited. Examining the studies on rheumatologic image processing reveals that the CNN approach has got more attention and has proven to be more effective than other approaches. It is predicted that researchers who start image processing studies in rheumatological diseases can develop more efficient classification systems if they try to approach the solution with such convolutional DL methods.

Although some AI-based diagnosing systems achieved human-level performance results (Lee et al., 2017),one of the fundamental reasons AI technologies cannot find a place in clinical use is that experts do not know precisely what the system has learned, especially in DL-based systems. There is no definite consensus on how these developed models reach the relevant conclusion during operation (Avramidis, Avramidou, & Papakostas, 2022). Moreover, professionals' conditioning about human social behavior and the capabilities of inherent, underlying science and technology might be regarded as the basis for their opposition to exploiting the potential of AI in health (Thrall et al., 2018). As a result, a new way is required to adequately represent how DL models develop their outputs and decision-making processes.

Finally, in rheumatology, it would be incorrect to predict that AI technologies will only bring innovations in image processing. In particular, there are ML-based studies for editing electronic health records and creating new study protocols by combining different types of health data (Rothenberg, Patel, & Herscu, 2016), ANN models to reconstruct CT images to lower ionic radiation doses (Nathan M. Cross, 2017), programs that shorten MRI scanning times (Golkov et al., 2016; Hammernik et al., 2018), and studies for medical image improvement (Esses et al., 2018; Lakhani et al., 2018).

Ethics in Publishing

There are no ethical issues regarding the publication of this study

Acknowledgements

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

The suggested essay is based on a dissertation titled "Development of a Decision Support System in the Early Diagnosis of Ankylosing Spondylitis Using Image Processing and Deep Learning Methods." On 17.03.2021, the board of directors of Marmara University, Institute of Pure and Applied Sciences authorized the above-mentioned dissertation project, which is currently ongoing.

Author Contributions

Author contribution roles were presented according to CRediT author statement

Zehra Aysun Altikardes: Conceptualization, Validation, Investigation, Writing -Original Draft, Project Administration

Emre Canayaz:, Validation, Investigation, Writing -Review-Editting,

Alpraslan Unsal: Conceptualization, Validation, Supervision, Writing -Review-Editting

References

Aizenberg, E., Roex, E. A., Nieuwenhuis, W. P., Mangnus, L., van der Helm-van Mil, A. H., Reijnierse, M., . . . Stoel, B. C. (2018). Automatic quantification of bone marrow edema on MRI of the wrist in patients with early arthritis: a feasibility study. *Magnetic resonance in medicine, 79*(2), 1127-1134.

Algan, G., & Ulusoy, I. (2021). Image classification with deep learning in the presence of noisy labels: A survey. *Knowledge-Based Systems, 215*, 106771.

Antony, J., McGuinness, K., O'Connor, N. E., & Moran, K. (2016). *Quantifying radiographic knee osteoarthritis severity using deep convolutional neural networks.* Paper presented at the 2016 23rd International Conference on Pattern Recognition (ICPR).

Ashinsky, B. G., Bouhrara, M., Coletta, C. E., Lehallier, B., Urish, K. L., Lin, P. C., . . . Spencer, R. G. (2017). Predicting early symptomatic osteoarthritis in the human knee using machine learning classification of magnetic resonance images from the osteoarthritis initiative. *Journal of Orthopaedic Research, 35*(10), 2243-2250.

Avramidis, G. P., Avramidou, M. P., & Papakostas, G. A. (2022). Rheumatoid Arthritis Diagnosis: Deep Learning vs. Humane. *Applied Sciences, 12*(1), 10.

Becker, A. (2019). Artificial intelligence in medicine: What is it doing for us today? *Health Policy and Technology, 8*(2), 198-205.

Bengio, Y., & LeCun, Y. (2007). Scaling learning algorithms towards AI. *Large-scale kernel machines, 34*(5), 1-41.

Bidgood Jr, W. D., Horii, S. C., Prior, F. W., & Van Syckle, D. E. (1997). Understanding and using DICOM, the data interchange standard for biomedical imaging. *Journal of the American Medical Informatics Association, 4*(3), 199-212.

Boonen, A., Brinkhuizen, T., Landewé, R., van der Heijde, D., & Severens, J. L. (2010). Impact of ankylosing spondylitis on sick leave, presenteeism and unpaid productivity, and estimation of the societal cost. *Annals of the Rheumatic Diseases, 69*(6), 1123-1128. doi:10.1136/ard.2009.116764

Boonen, A., & Mau, W. (2009). The economic burden of disease: comparison between rheumatoid arthritis and ankylosing spondylitis. *Clinical and experimental rheumatology, 27*(4 Suppl 55), S112-117.

Brahim, A., Jennane, R., Riad, R., Janvier, T., Khedher, L., Toumi, H., & Lespessailles, E. (2019). A decision support tool for early detection of knee OsteoArthritis using X-ray imaging and machine learning: Data from the OsteoArthritis Initiative. *Computerized Medical Imaging and Graphics, 73*, 11-18.

Bressem, K. K., Vahldiek, J. L., Adams, L. C., Niehues, S. M., Haibel, H., Rodriguez, V. R., . . . Poddubnyy, D. (2020). Detecting radiographic sacroiliitis using deep learning with expert-level accuracy in axial spondyloarthritis. *medRxiv*, 2020.2005.2019.20105304. doi:10.1101/2020.05.19.20105304

Brui, E., Efimtcev, A. Y., Fokin, V. A., Fernandez, R., Levchuk, A. G., Ogier, A. C., . . . Bendahan, D. (2020). Deep learning-based fully automatic segmentation of wrist cartilage in MR images. *NMR in Biomedicine, 33*(8), e4320.

Carano, R. A., Lynch, J. A., Redei, J., Ostrowitzki, S., Miaux, Y., Zaim, S., . . . Genant, H. K. (2004). Multispectral analysis of bone lesions in the hands of patients with rheumatoid arthritis. *Magnetic resonance imaging, 22*(4), 505-514.

Castro-Zunti, R., Park, E. H., Choi, Y., Jin, G. Y., & Ko, S.-b. (2020). Early detection of ankylosing spondylitis using texture features and statistical machine learning, and deep learning, with some patient age analysis. *Computerized Medical Imaging and Graphics, 82*, 101718. doi:https://doi.org/10.1016/j.compmedimag.2020.101718

Chaturvedi, N. (2021). Deepra: Predicting joint damage from radiographs using cnn with attention. *arXiv preprint arXiv:2102.06982*.

Deng, L., Li, J., Huang, J.-T., Yao, K., Yu, D., Seide, F., . . . Williams, J. (2013). *Recent advances in deep learning for speech research at Microsoft.* Paper presented at the 2013 IEEE international conference on acoustics, speech and signal processing.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Esses, S. J., Lu, X., Zhao, T., Shanbhogue, K., Dane, B., Bruno, M., & Chandarana, H. (2018). Automated image quality evaluation of T2-weighted liver MRI utilizing deep learning architecture. *Journal of Magnetic Resonance Imaging, 47*(3), 723-728.

Faleiros, M. C., Nogueira-Barbosa, M. H., Dalto, V. F., Júnior, J. R. F., Tenório, A. P. M., Luppino-Assad, R., . . . de Azevedo-Marques, P. M. (2020). Machine learning techniques for computer-aided classification of active inflammatory sacroiliitis in magnetic resonance imaging. *Advances in Rheumatology, 60*, 1-10.

Fiorentino, M. C., Moccia, S., Cipolletta, E., Filippucci, E., & Frontoni, E. (2019). *A learning approach for informative-frame selection in US rheumatology images.* Paper presented at the International Conference on image analysis and processing.

Golkov, V., Dosovitskiy, A., Sperl, J. I., Menzel, M. I., Czisch, M., Sämann, P., . . . Cremers, D. (2016). Q-space deep learning: twelve-fold shorter and model-free diffusion MRI scans. *IEEE transactions on medical imaging, 35*(5), 1344-1351.

Hamet, P., & Tremblay, J. (2017). Artificial intelligence in medicine. *Metabolism, 69*, S36-S40.

Hammernik, K., Klatzer, T., Kobler, E., Recht, M. P., Sodickson, D. K., Pock, T., & Knoll, F. (2018). Learning a variational network for reconstruction of accelerated MRI data. *Magnetic resonance in medicine, 79*(6), 3055-3071.

Han, Q., Lu, Y., Han, J., Luo, A., Huang, L., Ding, J., . . . Liang, Q. (2021). Automatic quantification and grading of hip bone marrow oedema in ankylosing spondylitis based on deep learning. *Modern Rheumatology*.

Hemalatha, R., Vijaybaskar, V., & Thamizhvani, T. (2019). Automatic localization of anatomical regions in medical ultrasound images of rheumatoid arthritis using deep learning. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine, 233*(6), 657-667.

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*(2), rkz047.

Hirvasniemi, J., Gielis, W. P., Arbabi, S., Agricola, R., van Spil, W. E., Arbabi, V., & Weinans, H. (2019). Bone texture analysis for prediction of incident radiographic hip osteoarthritis using machine learning: data from the Cohort Hip and Cohort Knee (CHECK) study. *Osteoarthritis and cartilage, 27*(6), 906-914.

Kansagra, A. P., John-Paul, J. Y., Chatterjee, A. R., Lenchik, L., Chow, D. S., Prater, A. B., . . . Heilbrun, M. E. (2016). Big data and the future of radiology informatics. *Academic radiology, 23*(1), 30-42.

Kayalibay, B., Jensen, G., & van der Smagt, P. (2017). CNN-based segmentation of medical imaging data. *arXiv preprint arXiv:1701.03056*.

Khan, M. A. (2002). Thoughts concerning the early diagnosis of ankylosing spondylitis and related diseases. *Clinical and experimental rheumatology, 20*(6 Suppl 28), S6-10. Retrieved from http://europepmc.org/abstract/MED/12463439

Kim, K.-J., & Tagkopoulos, I. (2019). Application of machine learning in rheumatic disease research. *The Korean journal of internal medicine, 34*(4), 708-722. doi:10.3904/kjim.2018.349

Knight, W. (2017). The dark secret at the heart of AI'11 April 2017. In: MIT Technology Review https://www. technologyreview. com/s/604087/the-dark ….

Kowsari, K., Brown, D. E., Heidarysafa, M., Meimandi, K. J., Gerber, M. S., & Barnes, L. E. (2017). *Hdltex: Hierarchical deep learning for text classification.* Paper presented at the 2017 16th IEEE international conference on machine learning and applications (ICMLA).

Kräter, M., Abuhattum, S., Soteriou, D., Jacobi, A., Krüger, T., Guck, J., & Herbig, M. (2021). AIDeveloper: deep learning image classification in life science and beyond. *Advanced science, 8*(11), 2003743.

KS, A. S. D. M. D., Selvakumar, M., Sathyamangalam, E., & Nadu, T. (2023). CLASSIFICATION OF DEEP LEARNING ALGORITHM FOR RHEUMATOID ARTHRITIS PREDICTOR.

Kushner, D. C., & Lucey, L. L. (2005). Diagnostic radiology reporting and communication: the ACR guideline. *Journal of the American College of Radiology, 2*(1), 15-21.

Lakhani, P., Prater, A. B., Hutson, R. K., Andriole, K. P., Dreyer, K. J., Morey, J., . . . Itri, J. N. (2018). Machine learning in radiology: applications beyond image interpretation. *Journal of the American College of Radiology, 15*(2), 350-359.

LeCun, Y., & Bengio, Y. (1995). Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks, 3361*(10), 1995.

Lee, H., Tajmir, S., Lee, J., Zissen, M., Yeshiwas, B. A., Alkasab, T. K., . . . Do, S. (2017). Fully automated deep learning system for bone age assessment. *Journal of digital imaging, 30*(4), 427-441.

Liu, F., Zhou, Z., Samsonov, A., Blankenbaker, D., Larison, W., Kanarek, A., . . . Kijowski, R. (2018). Deep learning approach for evaluating knee MR images: achieving high diagnostic performance for cartilage lesion detection. *Radiology, 289*(1), 160-169.

Liu, J., Chang, W.-C., Wu, Y., & Yang, Y. (2017). *Deep learning for extreme multi-label text classification.* Paper presented at the Proceedings of the 40th international ACM SIGIR conference on research and development in information retrieval.

Majumder, N., Poria, S., Gelbukh, A., & Cambria, E. (2017). Deep learning-based document modeling for personality detection from text. *IEEE Intelligent Systems, 32*(2), 74-79.

Maziarz, K., Krason, A., & Wojna, Z. (2021). Deep Learning for Rheumatoid Arthritis: Joint Detection and Damage Scoring in X-rays. *arXiv preprint arXiv:2104.13915*.

Mennini, F. S., Viti, R., Marcellusi, A., Sciattella, P., Viapiana, O., & Rossini, M. (2018). Economic evaluation of spondyloarthritis: economic impact of diagnostic delay in Italy. *Clinicoecon Outcomes Res, 10*, 45-51. doi:10.2147/ceor.S144209

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. *Multimedia tools and applications, 77*(9), 10921-10937.

Nanni, L., Costa, Y. M., Aguiar, R. L., Silla Jr, C. N., & Brahnam, S. (2018). Ensemble of deep learning, visual and acoustic features for music genre classification. *Journal of New Music Research, 47*(4), 383-397.

Nathan M. Cross, M., MS, CIIP, University of Pennsylvania; Jason DeBerry, MD; Daniel Ortiz, MD; Justine Kemp, MD; José Morey, MD. (2017). *Diagnostic Quality of Machine Learning Algorithm for Optimization of Low-Dose Computed Tomography Data* Paper presented at the SIIM (Society for Imaging Informatics in Medicine) 2017 Pittsburgi PA. https://cdn.ymaws.com/siim.org/resource/resmgr/siim2017/abstracts/posters-Cross.pdf

Norman, B., Pedoia, V., Noworolski, A., Link, T. M., & Majumdar, S. (2019). Applying densely connected convolutional neural networks for staging osteoarthritis severity from plain radiographs. *Journal of digital imaging, 32*(3), 471-477.

Orange, D. E., Agius, P., DiCarlo, E. F., Robine, N., Geiger, H., Szymonifka, J., . . . Donlin, L. T. (2018). Identification of Three Rheumatoid Arthritis Disease Subtypes by Machine Learning Integration of Synovial Histologic Features and RNA Sequencing Data. *Arthritis & Rheumatology, 70*(5), 690-701. doi:https://doi.org/10.1002/art.40428

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., . . . Brennan, S. E. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Systematic reviews, 10*(1), 1-11.

Pahade, J., Couto, C., Davis, R., Patel, P., Siewert, B., & Rosen, M. (2012). Reviewing imaging examination results immediately after study completion with a radiologist: Patient preferences and assessment of feasibility in an academic department. *AJR. American journal of roentgenology, 199*(4), 844.

Pedoia, V., Norman, B., Mehany, S. N., Bucknor, M. D., Link, T. M., & Majumdar, S. (2019). 3D convolutional neural networks for detection and severity staging of meniscus and PFJ cartilage morphological degenerative changes in osteoarthritis and anterior cruciate ligament subjects. *Journal of Magnetic Resonance Imaging, 49*(2), 400-410.

Poddubnyy, D., Sieper, J., Akar, S., Muñoz-Fernández, S., Haibel, H., Hojnik, M., . . . Inman, R. D. (2021). Characteristics of patients with axial spondyloarthritis by geographic regions: PROOF multicountry observational study baseline results. *Rheumatology*. doi:10.1093/rheumatology/keab901

Prasoon, A., Petersen, K., Igel, C., Lauze, F., Dam, E., & Nielsen, M. (2013). *Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network.* Paper presented at the International conference on medical image computing and computer-assisted intervention.

Ribas, L. C., Riad, R., Jennane, R., & Bruno, O. M. (2022). A complex network based approach for knee Osteoarthritis detection: Data from the Osteoarthritis initiative. *Biomedical Signal Processing and Control, 71*, 103133. doi:https://doi.org/10.1016/j.bspc.2021.103133

Ronneberger, O., Fischer, P., & Brox, T. (2015). *U-net: Convolutional networks for biomedical image segmentation.* Paper presented at the International Conference on Medical image computing and computer-assisted intervention.

Rothenberg, S., Patel, J., & Herscu, M. (2016). *Evaluation of a machine learning approach to protocol MRI examinations: initial experience predicting use of contrast by neuroradiologists in MRI protocols.* Paper presented at the Radiology Society of North America, 102nd Scientific Assembly and Annual Meeting.

Samuel, A. L. (2000). Some studies in machine learning using the game of checkers. *IBM Journal of research and development, 44*(1.2), 206-226.

Schaefer, G., Krawczyk, B., Doshi, N. P., & Merla, A. (2013). *Scleroderma capillary pattern identification using texture descriptors and ensemble classification.* Paper presented at the 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).

Segen, J., Kulbacki, M., & Wereszczyński, K. (2015). *Registration of ultrasound images for automated assessment of synovitis activity.* Paper presented at the Asian Conference on Intelligent Information and Database Systems.

Shenkman, Y., Qutteineh, B., Joskowicz, L., Szeskin, A., Yusef, A., Mayer, A., & Eshed, I. (2019). Automatic detection and diagnosis of sacroiliitis in CT scans as incidental findings. *Medical Image Analysis, 57*, 165-175. doi:https://doi.org/10.1016/j.media.2019.07.007

Subhash, M. G., & Kureshi, A. K. (2023). An Efficient CNN for Hand X-Ray Classification of Rheumatoid Arthritis. *Microprocessors and Microsystems*, 104822. doi:https://doi.org/10.1016/j.micpro.2023.104822

Thrall, J. H., Li, X., Li, Q., Cruz, C., Do, S., Dreyer, K., & Brink, J. (2018). Artificial Intelligence and Machine Learning in Radiology: Opportunities, Challenges, Pitfalls, and Criteria for Success. *Journal of the American College of Radiology, 15*(3, Part B), 504-508. doi:https://doi.org/10.1016/j.jacr.2017.12.026

Tiulpin, A., Thevenot, J., Rahtu, E., Lehenkari, P., & Saarakkala, S. (2018). Automatic Knee Osteoarthritis Diagnosis from Plain Radiographs: A Deep Learning-Based Approach. *Scientific Reports, 8*(1), 1727. doi:10.1038/s41598-018-20132-7

Tripoliti, E. E., Fotiadis, D. I., & Argyropoulou, M. (2007). Automated segmentation and quantification of inflammatory tissue of the hand in rheumatoid arthritis patients using magnetic resonance imaging data. *Artificial intelligence in medicine, 40*(2), 65-85.

Ward, M. M. (2002). Functional disability predicts total costs in patients with ankylosing spondylitis. *Arthritis Rheum, 46*(1), 223-231. doi:10.1002/1529-0131(200201)46:1<223::Aid-art498>3.0.Co;2-#

Wesseling, J., Dekker, J., van den Berg, W. B., Bierma-Zeinstra, S. M. A., Boers, M., Cats, H. A., . . . Bijlsma, J. W. J. (2009). CHECK (Cohort Hip and Cohort Knee): similarities and differences with the Osteoarthritis Initiative. *Annals of the Rheumatic Diseases, 68*(9), 1413-1419. doi:10.1136/ard.2008.096164

Wu, M., Wu, H., Wu, L., Cui, C., Shi, S., Xu, J., . . . Dong, F. (2022). A deep learning classification of metacarpophalangeal joints synovial proliferation in rheumatoid arthritis by ultrasound images. *J Clin Ultrasound, 50*(2), 296-301. doi:10.1002/jcu.23143

Xue, Y., Zhang, R., Deng, Y., Chen, K., & Jiang, T. (2017). A preliminary examination of the diagnostic value of deep learning in hip osteoarthritis. *PloS one, 12*(6), e0178992.

Zhang, W., Yang, G., Lin, Y., Ji, C., & Gupta, M. M. (2018). *On definition of deep learning.* Paper presented at the 2018 World automation congress (WAC).

Zharkova, V. (2007). *Artificial intelligence in recognition and classification of astrophysical and medical images* (Vol. 46): Springer Science & Business Media.