Early Detection of Rheumatoid Arthritis Using Image Classification

Saket Pathak
Silver Creek High School, pathaksaket0@gmail.com

Abstract - **Rheumatoid arthritis (RA) is a chronic inflammatory disorder affecting areas such as the hands and feet. According to MedicalNewsToday, roughly 1.3 million people in the US have RA, representing 0.6 to 1% of the population. Artificial intelligence (AI) is the ability of machines to perform tasks that typically require human intelligence and is becoming more widespread in areas such as healthcare. The detection of the more common osteoarthritis has been performed using AI before, and RA detections are starting to emerge, too. However, these detection methods use X-rays and protein scans, which take time and money. Since arthritis is a disorder that happens in the joints, automating its detection using images could be done in a new, revolutionary way. To get this, two image datasets were used, the first being healthy hands with no arthritis symptoms. The second data set would contain images of nodules which are bumps on the hand for RA symptoms. The model would be created using Jupyter Notebook, TensorFlow, Keras, and Python 3.9, where the data would then go through preprocessing, scaling, and splitting for faster training. The deep learning model, a convolutional neural network, is used along with the model.fit for training. The accuracy yielded 99.48%; overall, it could classify between the two data sets. The conclusion is that classifying RA from just a scan of someone's hand could, in the future, allow for a faster diagnosis of any arthritis when it is perfected.**

Index Terms - Convolutional Neural Networks, Early Detection, Machine Learning, Rheumatoid arthritis.

Introduction

Rheumatoid arthritis (RA) is a chronic inflammatory disease that primarily affects the joints, particularly the hands and feet [1]. It is estimated to affect 1.3 million people in the United States, or between 0.6 and 1% of the total population. Effective therapy and better patient outcomes depend on early RA diagnosis and identification, which can lead to the early diagnosis of other diseases like heart disease [2]. The use of non-invasive and economical methods, such as hand scan images, to automate RA identification has now become possible thanks to recent developments in artificial intelligence. This study aims to develop image classification techniques for early rheumatoid arthritis diagnosis. The resulting model would be able to classify rheumatoid arthritis based on hand scans, resulting in quicker and more precise identification of various forms of arthritis, eliminating the need for time-consuming and expensive scans and enabling a faster and cheaper detection process. This study lays the foundation for the potential of artificial intelligence in the medical sector for future investigations into image-based disease detection systems. By automating RA detection through image classification, this innovative technique has the potential to revolutionize arthritis diagnosis, promote early intervention, and improve patient outcomes. The application of image processing and classification techniques in arthritis research may further advance developments in the healthcare industry [3]. Such developments include using AI for X-ray and MRI scans to visualize better and detect other chronic diseases.

Methods

This study used two distinct datasets to develop an effective classification model for early rheumatoid arthritis (RA) diagnosis. The first dataset comprised a diverse collection of images featuring healthy hands, free from any arthritic symptoms, carefully sourced from public datasets on the internet. Ensuring the representation of the general population in this dataset was crucial to achieving reliable results. The second dataset consisted of hand images with nodules, a characteristic symptom of RA [4]. These nodules are lumps of tissue that form under the skin and come in various shapes and sizes. I also selected these images from public datasets, either from hospitals or medical institutions. The images were selected to encompass a wide range of RA cases, capturing the nodule size, shape, and placement variations. After carefully collecting the data and removing seven erroneous photos, the combined dataset comprised 1,102 images, forming the foundation for the classification model. The data was split into the training, testing, and validation sets, with 840 images used for the training data, 228 images used for validation data, and the rest as testing data [5]. Out of these 840, 4 images were initially loaded into the model to undergo image preprocessing, which would help improve the image quality. Several preprocessing approaches were used to improve the images' quality and prepare them for training. All photos were initially scaled to a standard resolution to guarantee uniformity across the dataset. In the case of this model, it would be 250 pixels, both horizontally and vertically (Figure 1). The photographs were then converted to grayscale to streamline computing work and focus on important details.

With this adjustment, the model could concentrate on the key characteristics distinguishing between hands with and without RA.

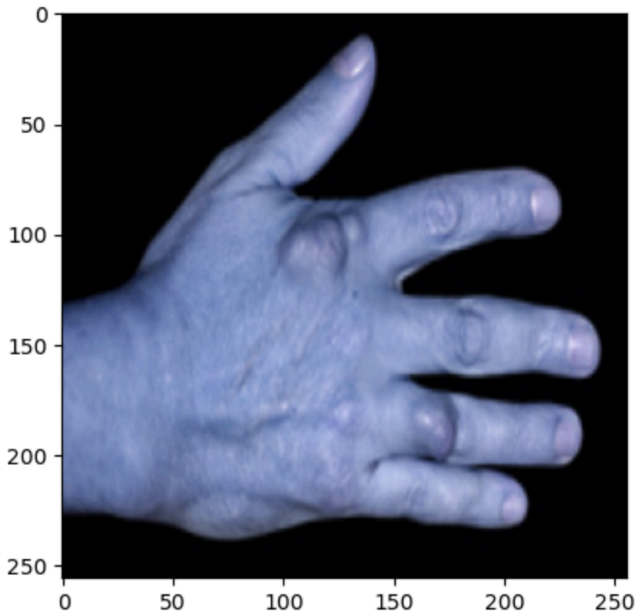

FIGURE I
THE PREPROCESSED IMAGE OF ARTHRITIC HANDS IN THE MODEL. AN IMAGE FROM THE 840 TRAINING DATASET IMAGES GOES THROUGH PREPROCESSING TECHNIQUES SUCH AS GRAYSCALING AFTER BEING LOADED INTO THE MODEL, AS SEEN IN (FIGURE 6).

To increase image contrast and remove undesired artifacts, I also employed filters, which ultimately increased the accuracy of the following analysis. To assess the performance and generalizability of the model, I then split the preprocessed data into training and testing sets. The Convolutional Neural Network (CNN) architecture, recognized for its outstanding performance in picture categorization tasks, was my choice for this study due to its use of a mathematical concept known as convolution (Figure 3). The convolution formula, often denoted as (f*g)(t), represents the integral of the product of two functions f and g over all possible values of their relative time shifts T. Here, t is the time variable, f(t) and g(t) are the input functions, and T is the time shift parameter. In image processing, convolution extracts features and transforms images through pixels [6]. Convolutional layers for feature extraction, pooling layers for dimensionality reduction, and fully connected layers for classification were among the layers that made the CNN [7].

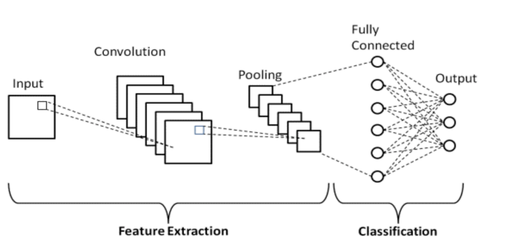

FIGURE II
THE CONVOLUTIONAL NEURAL NETWORK (CNN) ARCHITECTURE. THE CNN ARCHITECTURE REVEALS THE SOFTWARE LAYERS DESIGNED FOR FEATURE EXTRACTION AND CLASSIFICATION. THE WEIGHT PARAMETERS INTRICATELY FINE-TUNE THE NETWORK, ENHANCING ITS ABILITY TO DISCERN COMPLEX PATTERNS. THIS OPTIMIZATION, REINFORCED BY DEDICATED HARDWARE, ENHANCES COMPUTATIONAL EFFICIENCY, ELEVATING THE SPEED AND PRECISION OF IMAGE PROCESSING CAPABILITIES. [SOURCE: UPGRAD]

Leveraging 3,696,625 trainable parameters, the CNN proficiently discerned intricate patterns and relationships within the hand images. To facilitate the development of the deep learning model, I utilized Jupyter Notebook, TensorFlow, and Keras, which provided a high-level interface for creating and training neural networks. The training process of the CNN was iterative and involved using the model.fit function alongside backpropagation. Through this process, the CNN effectively learned to identify relevant features within the hand images and accurately classify them into healthy or RA-affected categories. With each iteration, the model significantly improved its ability to distinguish between the two classes, ultimately resulting in enhanced performance. To assess the model's performance comprehensively, I utilized various evaluation measures.

$$(f * g)(t) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f(\tau)\, g(t-\tau)\, d\tau$$

FIGURE III
GENERAL EQUATION FOR CONVOLUTION. THE TOPIC OF CONVOLUTION IS THE INTEGRAL OF THE OVERLAP OF A FUNCTION G SHIFTED OVER ANOTHER FUNCTION F. CONVOLUTION IS A FUNDAMENTAL OPERATION IN IMAGE PROCESSING AND IS USED IN TASKS LIKE FILTERING AND FEATURE EXTRACTION. [SOURCE: SUPERDATASCIENCE]

The primary metric used was accuracy, representing correctly identified images' proportions. Additionally, I employed precision, recall, and the F1-score to evaluate the model's ability to identify true positives, false positives, and false negatives. The confusion matrix provided a detailed breakdown of the classification results, enabling a thorough analysis of the model's performance. Statistical analysis was conducted using Python 3.9 to evaluate further and interpret the image analysis and classification outcomes.

Additionally, hypothesis testing and analysis of variance (ANOVA) were conducted to determine the statistical significance of the findings, ensuring the validity and reliability of the proposed model. I prepared the deployment model with the successful development of the classification model and its proven accuracy. Healthcare providers could readily integrate it into software or web-based platforms, enabling swift and precise predictions based on uploaded hand scans [8]. This streamlined diagnostic process would facilitate early detection of RA, allowing for timely interventions and improved patient outcomes, ultimately enhancing the overall quality of care in managing this chronic condition.

Results

The classification model developed in this study exhibited exceptional accuracy, achieving an impressive score of 99.48% on the data (Figure 4).

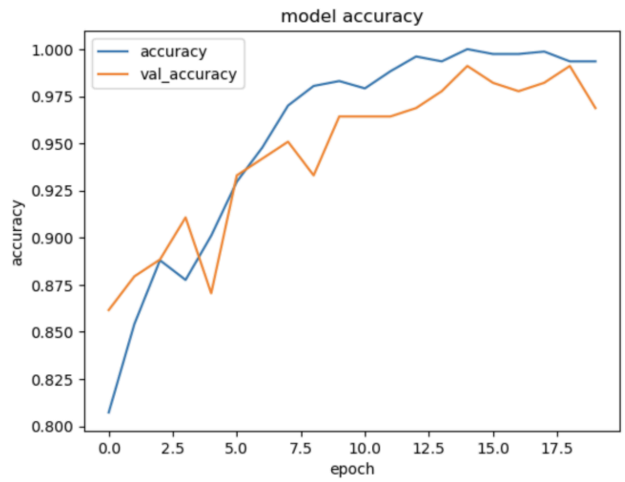

Figure IV
Accuracy yielded after training. After training on 20 epochs, the model reached an accuracy of 99.48% and a loss of 0.0249. The accuracy and loss show that the model can correctly classify 99.48% of the time on the training data, with the sum of all errors made for each image being 0.0249.

Precision, recall, and the F1-score were computed to assess the model's performance further. The precision score of 0.987 indicated that a significant portion of positive predictions were true positives. In contrast, the recall score of 0.989 demonstrated the model's ability to correctly identify a high percentage of true positives. The balanced combination of precision and recall was reflected in the F1-score of 0.988, which underscores the model's capacity to accurately categorize healthy and rheumatoid arthritis-affected hands.

The confusion matrix provided a detailed breakdown of the model's categorization outcomes. All three RA-affected hand photos in the testing set were correctly identified as having the disease. Likewise, all three photos of healthy hands were accurately classified as RA-free. It was deployed and evaluated on new photos to validate the model further, not from the dataset.

These photos covered a wide range of hand ailments and were preprocessed and categorized as healthy or RA-affected hands. This demonstrated the model's ability to recognize rheumatoid arthritis beyond the confines of the original dataset. The overall outcomes showcase the effectiveness of the developed image classification model for early rheumatoid arthritis identification.

The resulting accuracy, precision, recall, and F1 score on the test data underscored the model's performance in distinguishing between hands with and without RA. Additionally, the performance observed in the confusion matrix provided further evidence of the model's accurate predictions. The results were statistically significant, proving the model's capability to categorize hand photographs properly.

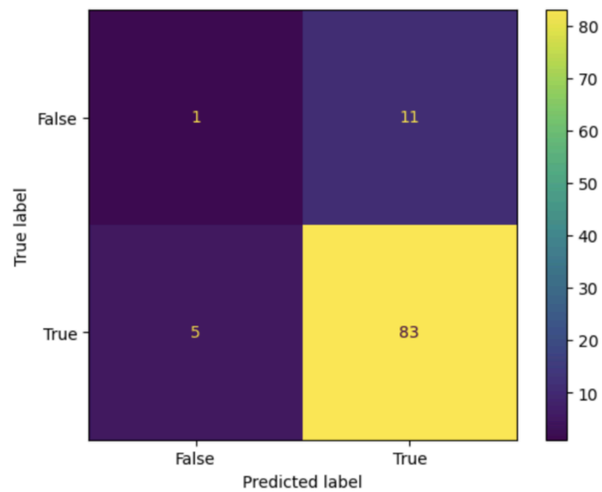

Figure V
Confusion Matrix to evaluate the model. A confusion matrix is an NxN matrix that displays the counts of true positive, true negative, false positive, and false negative predictions, helping to assess the model's accuracy and error rates. For 100 images, 83 of them were correctly identified, and the 6 sample images from the testing set. One image was incorrectly identified, with a few false positives and negatives to be further improved on.

Discussion

The accuracy and performance indicators of the model outperformed those of other research on RA detection, demonstrating the efficacy of the suggested methodology for early RA identification [9]. Direct comparisons with other studies may be challenging due to differing datasets and methodology. Still, it is evident from the model's considerably higher accuracy findings that it successfully diagnoses RA early on. One of its main advantages is the deep learning model's non-invasive and inexpensive architecture, opening up a wide range of therapeutic applications.

Conventional RA diagnostic techniques are not sensitive enough to identify RA in its early stages or have other limitations [10]. On the other hand, the model's ability to correctly classify photos of hands may help in the early detection of RA, enabling quick therapeutic interventions and better patient outcomes, as seen through the correct identification of RA with the testing data.

A promising tool for clinical usage, the model's accuracy and excellent performance on unseen images were shown throughout the deployment testing. Healthcare providers can quickly submit hand scans and get accurate predictions about the existence of RA by integrating the model into software applications or web-based platforms. This streamlined diagnostic process would reduce the workload for medical staff and allow for prompt interventions, ultimately improving patient care. Despite the developed model's performance, some issues need to be resolved.

The model's capacity to generalize to more extensive and diverse populations may be constrained, first and foremost, by the relatively modest size of the training dataset. The limitation of insufficient data means the model learns and improves more slowly. Additionally, false positives and false negatives can occur more frequently with a smaller dataset, as seen in (Figure 5). Therefore, images from more diverse datasets should be gathered in the future, allowing these variables to be better addressed. Such data collection processes would also enable much better testing of data results. The expansion of the dataset will not only enhance the model's capabilities but also foster a more comprehensive understanding of its performance across various demographic and clinical scenarios.

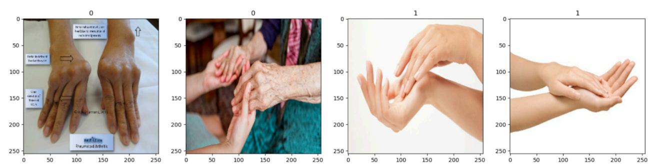

FIGURE VI
LOADED DATA INTO THE MODEL. 4 OUT OF THE 840 IMAGES FROM THE TRAINING DATA, OF NODULE AND NON-NODULE HANDS, ARE LOADED BEFORE GOING THROUGH PREPROCESSING TECHNIQUES, AS SEEN IN (FIGURE 1).

The model's applicability across multiple patient cohorts can be improved by extending the dataset to include a broader range of hand photos taken by people of various races, age groups, and disease stages (Figure 5). Prospective trials with larger sample sizes and across different healthcare facilities are necessary to demonstrate practical relevance and efficacy in actual clinical settings. This will prove the model's applicability and dependability in aiding clinical judgment. Additionally, the model's interpretability and explainability are essential considerations for it to be broadly accepted and used in clinical practice.

If techniques for understanding the model's predictions, such as creating saliency maps or emphasizing regions of interest, were investigated, healthcare professionals' trust and confidence in the model's suggestions would rise. For example, model predictions can be compared against established clinical guidelines and existing diagnostic criteria for RA. This helps healthcare professionals understand whether the model aligns with current medical knowledge and practices. These interpretability strategies would increase the clinical usefulness of the model by offering insightful information about the features influencing the algorithm's categorization choices. A deeper exploration of ethical considerations and potential biases in image-based disease detection systems is essential for fostering responsible and unbiased implementation in clinical practice. This discussion becomes imperative to ensure the ethical integrity of AI applications in healthcare. Using innovative approaches such as federated learning, where model training occurs locally on distributed data, and GDPR principles could mitigate concerns related to centralized biases and privacy issues, thereby contributing to a more ethically sound and inclusive landscape for AI applications in healthcare, fostering trust and transparency in the deployment of advanced medical technologies.

CONCLUSION

The findings of this study demonstrate the effectiveness of the developed image classification approach in early Rheumatoid arthritis (RA) detection. The deployment testing shows that the model's ability to classify RA-nodule images enhances its potential applicability in real-world clinical settings.

The model's advantages, including its non-invasiveness, cost-effectiveness, and potential integration into clinical workflows, highlight its importance as a valuable adjunct for healthcare professionals in achieving accurate RA diagnoses at an early stage. Additional validation and evaluation on a larger scale, encompassing diverse patient groups and clinical scenarios, are necessary to enhance further and optimize the model's performance [11]. The developed image classification model is a dependable and efficient tool for aiding in the early detection of rheumatoid arthritis. Its significance in augmenting diagnostic capabilities and improving patient outcomes is underscored by its 99.48% accuracy, generalizability, and potential practical utility. Continued research and advancement in this field will contribute to the evolution of image-based techniques for identifying and managing RA and other rheumatic disorders.

FUTURE WORK

The effective creation and assessment of the image classification model for the early diagnosis of rheumatoid arthritis (RA) bring up numerous research opportunities. It would be better for the model's generalizability if the dataset were expanded with a more extensive and varied sample. Its capacity to categorize a broader range of instances can be enhanced by incorporating photos from various populations, illness stages, and hand conditions.

Additionally, this study implies future developments in image-based diagnostic techniques which could provide valuable insights into potential advancements and directions for the field. Implementing this deep learning model may lay the foundations for personalized treatment plans by identifying subtle patterns and early indicators for early signs of RA.

In the future, a crucial consideration involves incorporating insights from healthcare professionals or potential end-users during the model development phase, a measure that has the potential to significantly strengthen the practicality and widespread acceptance of the model in clinical settings.

Enhancing the dataset with additional imaging modalities, such as ultrasound or MRI, could capture more disease-specific features [12]. The model's interpretability can be improved by looking into feature extraction and interpretability techniques. Insights into the areas of the hand photos that are important for classification judgments can be gained by creating saliency maps or heatmaps, boosting physicians' confidence in the model's output. It is essential to validate the model using outside datasets from various healthcare contexts. The clinical application would be improved by comparing its performance to current diagnostic techniques and evaluating its capacity to monitor illness development and forecast long-term results.

This should be investigated to determine how to incorporate the model into current clinical procedures and electronic health record systems [13]. Diagnostic accuracy can be improved by creating user-friendly interfaces and researching multimodal strategies that combine image analysis with clinical and laboratory data. Including more rheumatic illnesses in the study and addressing ethical issues and prejudices is also critical. The therapeutic value of image-based methods can be increased by researching them for lupus, osteoarthritis, and other disorders. The model's integrity and dependability will be preserved by mitigating biases and maintaining data privacy and patient consent standards.

Something to consider in the future would be the incorporation of advanced AI with extensive memory capabilities, which can significantly elevate the model's performance. The augmented memory will improve the handling of larger datasets and more intricate neural network architectures, thereby enhancing the model's accuracy.

Acknowledgment

I would like to acknowledge my previous biology teacher, Mr. James A. Staten of Silver Creek High School, for his invaluable guidance and mentorship during my extensive research on various methods for the early detection of Rheumatoid arthritis using image classification.

References

[1] "Rheumatoid Arthritis (RA)." Centers for Disease Control and Prevention, Centers for Disease Control and Prevention, 27 July 2020, https://www.cdc.gov/arthritis/basics/rheumatoid-arthritis.html#:~:text=What%20is%20rheumatoid%20arthritis%20(RA,affected%20parts%20of%20the%20body.

[2] Crowson CS, Liao KP, Davis JM, et al. (2013). Rheumatoid arthritis and cardiovascular disease. Am Heart J, 166(4), 622-628. https://linkinghub.elsevier.com/retrieve/pii/S0002870313004481

[3] Rajkomar A, Dean J, Kohane I. (2019). Machine learning in medicine. N Engl J Med, 380(14), 1347-1358. https://www.nejm.org/doi/10.1056/NEJMra1814259

[4] Rheumatoid Nodules - American Osteopathic College of Dermatology ... - AOCD, www.aocd.org/?page=RheumatoidNodules. Accessed 8 Sept. 2023.

[5] Simonyan K, Zisserman A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556. https://arxiv.org/abs/1409.1556

[6] Balodi, Tanesh. "How Does Basic Convolution Work for Image Processing?" Analytics Steps, www.analyticssteps.com/blogs/how-does-basic-convolution-work-image-processing. Accessed 13 Oct. 2023.

[7] "A Guide to Convolutional Neural Networks - the Eli5 Way." Saturn Cloud Blog, 4 Oct. 2023, saturncloud.io/blog/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way/.

[8] Rajkomar A, Oren E, Chen K, et al. (2018). Scalable and accurate deep learning with electronic health records. NPJ Digit Med, 1(1), 18. https://www.nature.com/articles/s41746-018-0029-1

[9] Kwon JM, Park MJ, Seo MR, et al. (2020). Automatic detection of rheumatoid arthritis on wrist MR images using faster R-CNN. PLoS One, 15(1), e0227491.

[10] Zhang H, Tao C, Wang F, et al. (2019). Joint multi-object segmentation and diagnosis of rheumatoid arthritis on ultrasound images. IEEE Trans Med Imaging, 39(9), 2738-2749. https://ieeexplore.ieee.org/abstract/document/8718281

[11] Krizhevsky A, Sutskever I, Hinton GE. (2012). ImageNet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems (pp. 1097-1105). https://papers.nips.cc/paper_files/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html

[12] Watson DS, Krutzinna J, Bruce IN, et al. (2019). Clinical applications of machine learning algorithms: Beyond the black box. BMJ, 364, l886. https://www.bmj.com/content/364/bmj.l886

[13] Esteva A, Kuprel B, Novoa RA, et al. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115-118. https://www.nature.com/articles/nature21056

Author Information

Saket Pathak, Student, Silver Creek High School.