Review

Application of deep learning for diagnosis of shoulder diseases in older adults: a narrative review

Sung Min Rhee

Shoulder & Elbow Clinic, Department of Orthopaedic Surgery, College of Medicine, Kyung Hee University Hospital, Seoul, Korea

Corresponding author: Sung-Min Rhee, Department of Orthopaedic Surgery, College of Medicine, Kyung Hee University Hospital, 23, Kyungheedae-ro, Dongdaemun-gu, 02447, Seoul, Korea, E-mail: minrhee77@gmail.com

Abstract

Shoulder diseases pose a significant health challenge for older adults, often causing pain, functional decline, and decreased independence. This narrative review explores how deep learning (DL) can address diagnostic challenges by automating tasks such as image segmentation, disease detection, and motion analysis. Recent research highlights the effectiveness of DL-based convolutional neural networks (CNNs) and machine learning frameworks in diagnosing various shoulder pathologies. Automated image analysis facilitates the accurate assessment of rotator cuff tear size, muscle degeneration, and fatty infiltration in magnetic resonance imaging or computed tomography scans, frequently matching or surpassing the accuracy of human experts. CNN-based systems are also adept at classifying fractures and joint conditions, enabling the rapid identification of common causes of shoulder pain from plain radiographs. Furthermore, advanced techniques like pose estimation provide precise measurements of the shoulder joint's range of motion and support personalized rehabilitation plans. These automated approaches have also been successful in quantifying local osteoporosis, utilizing machine learning-derived indices to classify bone density status. DL has demonstrated significant potential to improve diagnostic accuracy, efficiency, and consistency in the management of shoulder diseases in older patients. Machine learning-based assessments of imaging data and motion parameters can help clinicians optimize treatment plans and improve patient outcomes. However, to ensure their generalizability, reproducibility, and effective integration into routine clinical workflows, large-scale, prospective validation studies are necessary. As data availability and computational resources increase, the ongoing development of DL-driven applications is expected to further advance and personalize musculoskeletal care, benefiting both healthcare providers and the aging population.

Keywords: Aged; Computer neural networks; Deep learning; Rotator cuff injuries; Shoulder pain

Introduction

Background

Shoulder diseases pose a significant health burden on the aging population, affecting millions of individuals worldwide [1-3]. Common conditions such as rotator cuff tears, impingement syndrome, osteoarthritis, and adhesive capsulitis not only cause pain but also significantly impair the daily lives of patients by restricting their mobility and independence [1,4-8]. Timely and accurate diagnosis of these conditions is crucial for optimizing treatment outcomes and enhancing patient quality of life. However, traditional diagnostic tools, such as X-rays, magnetic resonance imaging (MRI), and ultrasound, face challenges including variability in interpretation and limited availability in resource-constrained environments [9]. Furthermore, these methods struggle to accurately and objectively measure joint range of motion, which further compromises their effectiveness in diagnosing musculoskeletal conditions [10].

Recent advances in artificial intelligence (AI), especially in the area of deep learning (DL), have revolutionized the diagnosis of shoulder diseases [11-14]. DL algorithms leverage artificial neural networks, modeled after the human brain, to process and analyze vast amounts of data with exceptional accuracy [15]. These algorithms can detect subtle patterns in medical images that may be overlooked by even experienced radiologists. They also analyze complex movements and postures through pose estimation techniques. By minimizing diagnostic errors, improving consistency, and facilitating detailed motion analysis, DL algorithms are widely applicable in imaging and movement assessment, transforming sectors like healthcare, rehabilitation, and biomechanics.

Objectives

This paper aims to explore recent studies on the application of DL in diagnosing shoulder diseases in older adults.

Ethics statement

As this study is a literature review, it did not require institutional review board approval or individual consent.

The analysis of rotator cuff muscles/tendons and fatty infiltrations using AI

In 2020, Taghizadeh et al. introduced an AI model specifically designed to automatically assess rotator cuff muscle degeneration by analyzing both atrophy and fatty infiltration in computed tomography (CT) images [14]. This model utilized a convolutional neural network (CNN) to automatically evaluate degeneration, including atrophy and fatty infiltration, in preoperative shoulder CT scans of patients with glenohumeral osteoarthritis. The CNN was tested on retrospective data from 103 CT scans and achieved Dice similarity coefficients that were comparable to those of manual radiologist segmentations. It demonstrated high accuracy in measuring atrophy ($R^2 = 0.87$), fatty infiltration ($R^2 = 0.91$), and overall degeneration ($R^2 = 0.91$). These findings highlight the potential of DL to provide efficient and reliable evaluations of rotator cuff muscles preoperatively.

Similarly, Ro et al. developed a DL framework that utilizes MRI to evaluate factors such as the occupation ratio and fatty infiltration in the supraspinatus muscle of patients with rotator cuff tears [12]. This study employed a deep-learning framework to analyze the occupation ratio and fatty infiltration in the supraspinatus muscle using shoulder MRI. A full CNN facilitated rapid and precise segmentation of the supraspinatus muscle and fossa, achieving high Dice similarity coefficients (0.97 for the fossa and 0.94 for the muscle) along with excellent sensitivity and specificity. Fatty infiltration was quantified using a region-based Otsu thresholding method, which revealed significant differences across Goutallier grades ($p < 0.0001$) [16] and demonstrated a moderate negative correlation with the occupation ratio ($\varrho = -0.75$, $p < 0.0001$) [17]. These findings indicate that integrating DL with automated thresholding techniques offers an objective and efficient means of quantifying key indices in shoulder MRI, thereby enhancing diagnostic accuracy and consistency.

Detection of shoulder pathologies including rotator cuff tears and fractures

Recently, DL technology has been employed to automate the segmentation and detection of rotator cuff tears using MRI.

Lee et al. developed a DL model utilizing a 3D U-Net CNN to detect, segment, and visualize rotator cuff tear lesions in three dimensions using MRI data from 303 patients [18]. The model, trained and validated on labeled MRI datasets, demonstrated robust performance. It achieved a Dice coefficient of 94.3%, a sensitivity of 97.1%, a specificity of 95.0%, a precision of 84.9%, an F1-score of 90.5%, and a Youden index of 91.8% (Figure 1).

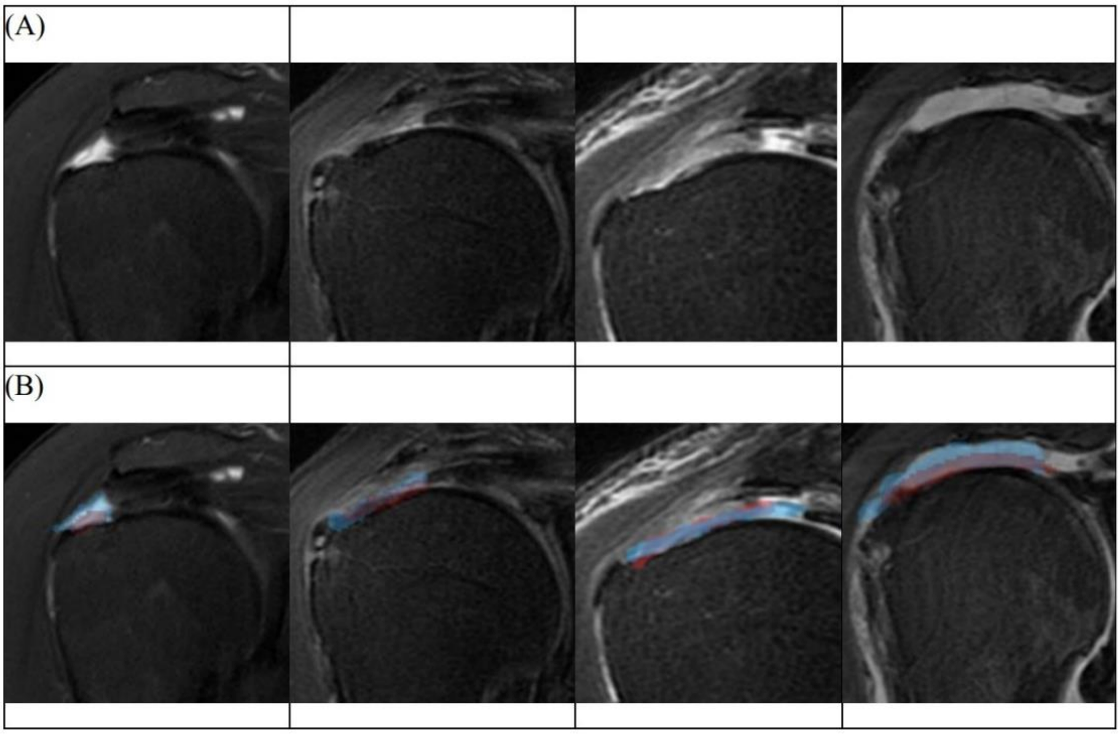

Fig. 1. Segmentation results corresponding to the rotator cuff tear site. (A) Original MRI images displaying the presence of a rotator cuff tear. (B) The red region represents the area manually labeled by shoulder specialists, while the blue region indicates the area segmented by the proposed deep learning model. This figure has been used with the author's permission [18].

Hashimoto et al. assessed the diagnostic capabilities of a CNN in detecting and classifying rotator cuff tears, using 1,169 anteroposterior shoulder radiographs. These were categorized into four groups: intact, small, medium, and large-to-massive tears [19]. In binary classification tasks, the CNN achieved a sensitivity of 92%, a specificity of 69%, an accuracy of 86%, and an area under the receiver operating curve (AUC) of 0.88. The CNN outperformed orthopedic surgeons in both detection and classification accuracy, demonstrating its potential as a reliable tool for diagnosing rotator cuff tears from plain radiographs.

A recent meta-analysis demonstrated that AI could perform comparably to clinicians in detecting fractures, highlighting its potential for broader applications in orthopedics. Magneli et al. developed and evaluated a CNN for classifying fractures in shoulder radiographs, focusing on proximal humeral fractures (PHF) based on the AO/OTA classification system, with secondary objectives for diaphyseal humerus, clavicle, and scapula fractures [20]. The CNN, trained on a dataset of 6,172 examinations, achieved an overall AUC of 0.89 for fracture classification. Notably, the AUC for PHF classes exceeded 0.90. The model also demonstrated excellent AUCs for diaphyseal humerus (0.97) and clavicle fractures (0.96), and a good performance for scapula fractures (0.87). Furthermore, Grauhan et al. developed a model capable of identifying a variety of common causes of shoulder pain on radiographs, extending beyond fractures to include conditions such as proximal humeral fractures, dislocations, periarticular calcifications, osteoarthritis, osteosynthesis, and joint prostheses [11]. This study utilized the ResNet-50 architecture to detect common causes of shoulder pain—such as fractures, dislocations, osteoarthritis, periarticular calcifications, osteosynthesis, and endoprosthesis—from plain radiographs. Trained on 2,700 radiographs and evaluated on a separate annotated dataset, the model demonstrated high accuracy. The CNN achieved excellent performance, with AUC values of 0.871 for fractures, 0.896 for joint dislocations, 0.945 for osteoarthritis, and 0.800 for periarticular calcifications. It also detected osteosynthesis and endoprosthesis with high accuracy, achieving AUC values of 0.998 and 1.0, respectively. Sensitivity and specificity varied by condition, with values of 0.75 and 0.86 for fractures, 0.95 and 0.65 for joint dislocations, 0.90 and 0.86 for osteoarthritis, and 0.60 and 0.89 for calcifications. These results underscore the potential of CNNs to aid clinicians by prioritizing worklists and improving diagnostic efficiency in high-workload settings.

Detection of local osteoporosis in the proximal humerus

Li et al. developed a diagnostic method using machine learning to assess local osteoporosis in the proximal humerus by analyzing demographic data, bone density, and X-ray ratios. The study involved a cohort of 97 patients (76 females and 21 males with an average age of 73 years), categorized into groups based on bone density: normal (25 patients), osteopenia (35 patients), and osteoporosis (37 patients). Utilizing the modified Tingart index [21], a decision tree was employed to identify critical diagnostic indicators, including the humeral shaft medullary cavity ratio (M2/M4), age, and sex. An M2/M4 ratio below 1.13 was indicative of local osteoporosis, whereas a ratio of 1.13 or higher, when analyzed alongside age and sex, helped differentiate between osteoporosis, osteopenia, and normal bone density. The decision tree achieved accuracies of 76.27% in the training set and 78.95% in the validation set. Additionally, multinomial logistic regression validated significant associations of M2/M4, age, and sex with osteoporosis.

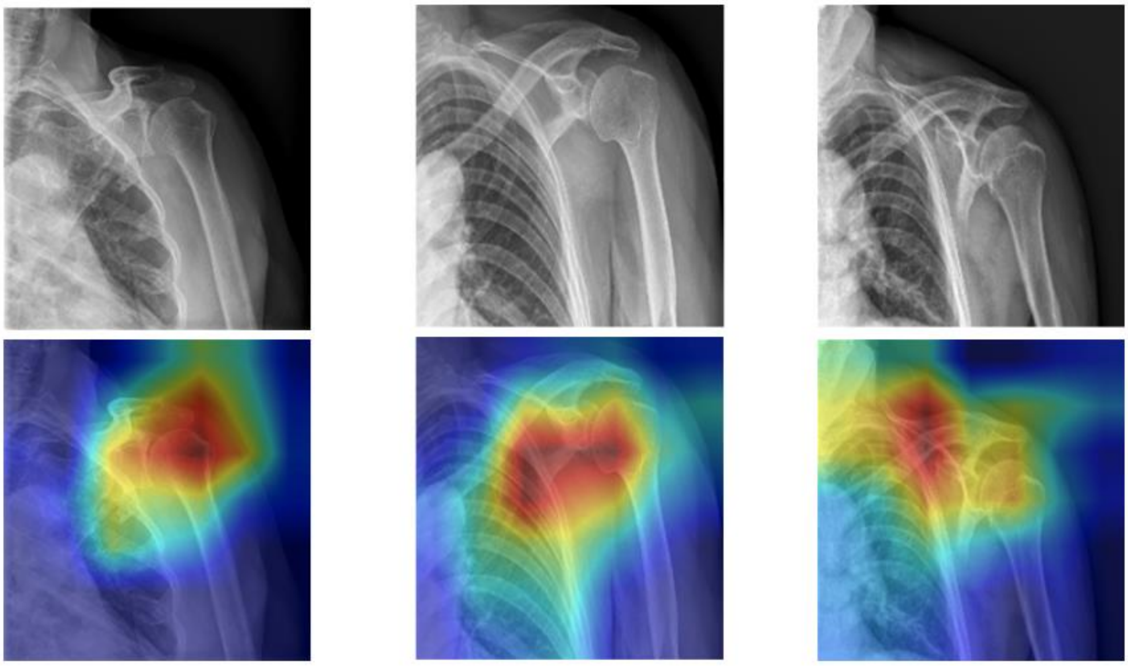

Fig. 2. The Grad-CAM visualization [22] shows which regions the AI focused on when analyzing the bone mineral density of the shoulder.

Analysis of shoulder range of motion using machine learning

Measuring shoulder joint angles accurately has been challenging due to the complexity of shoulder motion and its intricate rotational axes. Recently, pose estimation, a computer vision technique that utilizes machine learning, has garnered significant attention [23,24]. This technology predicts the positions and orientations of human joints or key points from images or videos, enabling detailed analysis of movements and postures [25]. In a recent study, the integration of pose estimation AI with machine learning has demonstrated a promising approach to estimating the range of motion of the shoulder with remarkable precision, paving the way for advancements in sports biomechanics and rehabilitation.

Fig. 3. A company utilizes machine learning-based pose estimation technology to measure a patient's range of motion, analyze the patient's current condition based on the results, and assign the most suitable rehabilitation exercises. This figure is used with permission from Itphy, Inc.

Takigami et al. employed pose estimation AI in conjunction with a machine learning model to estimate the internal and external rotation angles of the shoulder [26]. They processed videos of 10 healthy male volunteers (average age 37.7 years) into 10,608 images to develop parameters for training the model. Using smartphone angle measurements as the ground truth, the AI model demonstrated a correlation coefficient of 0.971 and a mean absolute error (MAE) of 5.778 using linear regression. With Light GBM, it achieved a correlation coefficient of 0.999 and an MAE of 0.945. This method offers a precise and efficient way to measure shoulder rotation angles, showing great potential for applications in sports biomechanics and rehabilitation.

Ramkumar et al. validated a motion-based machine learning software development kit (SDK) designed to assess shoulder range of motion. They compared its accuracy with that of manual goniometer measurements across four motion arcs: abduction, forward flexion, internal rotation, and external rotation [27]. Utilizing a mobile application, 10 subjects each performed the motions five times. The SDK recorded mean angular differences of less than 5° for all motions ($p > 0.05$), with specific mean differences of -3.7° for abduction, -4.9° for forward flexion, -2.4° for internal rotation, and -2.6° for external rotation.

Conclusion

The use of DL in diagnosing shoulder diseases among older patients has shown considerable promise in several areas. These include analyzing rotator cuff muscle degeneration, detecting pathologies such as rotator cuff tears and fractures, evaluating local osteoporosis in the proximal humerus, and accurately measuring the shoulder's range of motion. DL models, which employ sophisticated architectures like CNNs and incorporate machine learning algorithms, consistently achieve high levels of accuracy, sensitivity, and specificity in medical imaging tasks. These models often outperform traditional diagnostic techniques and expert clinicians.

ORCID

Sung Min Rhee: https://orcid.org/0000-0002-7087-4219

Authors' contributions

All works were done by Sung Min Rhee.

Conflicts of interest

Sung Min Rhee is the Chief Executive Officer of Itphy Inc., the company that provided Fig. 3. Otherwise, there are no conflicts of interest to declare.

Funding

Not applicable.

Data availability

Not applicable.

Acknowledgments

Not applicable.

Supplementary materials

Not applicable

References

1. Gumina S, Kim H, Jung Y, Song HS. Rotator cuff degeneration and healing after rotator cuff repair. *Clin Shoulder Elb* 2023;26(3):323-329.

2. Lucas J, van Doorn P, Hegedus E, Lewis J, van der Windt D. A systematic review of the global prevalence and incidence of shoulder pain. *BMC Musculoskeletal Disorders* 2022;23(1):1073.

3. Hodgetts CJ, Leboeuf-Yde C, Beynon A, Walker BF. Shoulder pain prevalence by age and within occupational groups: a systematic review. *Archives of Physiotherapy* 2021;11(1):24.

4. Kim YT, Kim TY, Lee JB, Hwang JT. Glenohumeral versus subacromial steroid injections for impingement syndrome with mild stiffness: a randomized controlled trial. *Clin Shoulder Elb* 2023;26(4):390-396.

5. Mardani-Kivi M, Hashemi-Motlagh K, Darabipour Z. Arthroscopic release in adhesive capsulitis of the shoulder: a retrospective study with 2 to 6 years of follow-up. *Clin Shoulder Elb* 2021;24(3):172-177.

6. Hones KM, Hao KA, Buchanan TR, Trammell AP, Wright JO, Wright TW, et al. Does preoperative forward elevation weakness affect clinical outcomes in anatomic or reverse total shoulder arthroplasty patients with glenohumeral osteoarthritis and intact rotator cuff? *Clin Shoulder Elb* 2024;27(3):316-326.

7. Rhee SM, Youn SM, Kim CH, Chang GW, Kim SY, Ham HJ, et al. Rotator cuff repairs with all-suture tape anchors: no difference in outcomes between with or without all-suture tape

anchors. *Knee Surg Sports Traumatol Arthrosc* 2023;31(9):4060-4067.

8. Ko YW, Park JH, Youn SM, Rhee YG, Rhee SM. Effects of comorbidities on the outcomes of manipulation under anesthesia for primary stiff shoulder. *J Shoulder Elbow Surg* 2021;30(8):e482-e492.

9. Malavolta EA, Assunção JH, Gracitelli MEC, Yen TK, Bordalo-Rodrigues M, Ferreira Neto AA. Accuracy of magnetic resonance imaging (MRI) for subscapularis tear: a systematic review and meta-analysis of diagnostic studies. *Arch Orthop Trauma Surg* 2019;139(5):659-667.

10. Elrahim RMA, Embaby EA, Ali MF, Kamel RM. Inter-rater and intra-rater reliability of Kinovea software for measurement of shoulder range of motion. *Bulletin of Faculty of Physical Therapy* 2016;21(2):80-87.

11. Grauhan NF, Niehues SM, Gaudin RA, Keller S, Vahldiek JL, Adams LC, et al. Deep learning for accurately recognizing common causes of shoulder pain on radiographs. *Skeletal Radiology* 2022;51(2):355-362.

12. Ro K, Kim JY, Park H, Cho BH, Kim IY, Shim SB, et al. Deep-learning framework and computer assisted fatty infiltration analysis for the supraspinatus muscle in MRI. *Scientific Reports* 2021;11(1):15065.

13. Shariatnia MM, Ramazanian T, Sanchez-Sotelo J, Maradit Kremers H. Deep learning model for measurement of shoulder critical angle and acromion index on shoulder radiographs. *JSES Rev Rep Tech* 2022;2(3):297-301.

14. Taghizadeh E, Truffer O, Becce F, Eminian S, Gidoin S, Terrier A, et al. Deep learning for the rapid automatic quantification and characterization of rotator cuff muscle degeneration from shoulder CT datasets. *European Radiology* 2021;31(1):181-190.

15. Alzubaidi L, Al-Dulaimi K, Salhi A, Alammar Z, Fadhel MA, Albahri AS, et al. Comprehensive review of deep learning in orthopaedics: Applications, challenges, trustworthiness, and fusion. *Artif Intell Med* 2024;155:102935.

16. Goutallier D, Postel JM, Bernageau J, Lavau L, Voisin MC. Fatty muscle degeneration in cuff ruptures. Pre- and postoperative evaluation by CT scan. *Clin Orthop Relat Res* 1994 Jul.(304):78-83.

17. Jeong JY, Chung PK, Lee SM, Yoo JC. Supraspinatus muscle occupation ratio predicts rotator cuff reparability. *J Shoulder Elbow Surg* 2017;26(6):960-966.
18. Lee SH, Lee J, Oh KS, Yoon JP, Seo A, Jeong Y, et al. Automated 3-dimensional MRI segmentation for the posterosuperior rotator cuff tear lesion using deep learning algorithm. *PLoS One* 2023;18(5):e0284111.
19. Hashimoto E, Maki S, Ochiai N, Ise S, Inagaki K, Hiraoka Y, et al. Automated detection and classification of the rotator cuff tear on plain shoulder radiograph using deep learning. *J Shoulder Elbow Surg* 2024;33(8):1733-1739.
20. Magnéli M, Ling P, Gislén J, Fagrell J, Demir Y, Arverud ED, et al. Deep learning classification of shoulder fractures on plain radiographs of the humerus, scapula and clavicle. *PLoS One* 2023;18(8):e0289808.
21. Tingart MJ, Apreleva M, von Stechow D, Zurakowski D, Warner JJ. The cortical thickness of the proximal humeral diaphysis predicts bone mineral density of the proximal humerus. *J Bone Joint Surg Br* 2003;85(4):611-617.
22. Rs R, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization2017.
23. van den Hoorn W, Lavaill M, Cutbush K, Gupta A, Kerr G. Comparison of Shoulder Range of Motion Quantified with Mobile Phone Video-Based Skeletal Tracking and 3D Motion Capture-Preliminary Study. *Sensors (Basel)* 2024;24(2).
24. Pereira B, Cunha B, Viana P, Lopes M, Melo ASC, Sousa ASP. A Machine Learning App for Monitoring Physical Therapy at Home. *Sensors (Basel)* 2023;24(1).
25. Sun K, Xiao B, Liu D, Wang J. Deep High-Resolution Representation Learning for Human Pose Estimation. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).2019. p. 5686-5696.
26. Takigami S, Inui A, Mifune Y, Nishimoto H, Yamaura K, Kato T, et al. Estimation of Shoulder Joint Rotation Angle Using Tablet Device and Pose Estimation Artificial Intelligence Model. *Sensors (Basel)* 2024;24(9).
27. Ramkumar PN, Haeberle HS, Navarro SM, Sultan AA, Mont MA, Ricchetti ET, et al. Mobile

technology and telemedicine for shoulder range of motion: validation of a motion-based machine-learning software development kit. *J Shoulder Elbow Surg* 2018;27(7):1198-1204.