Deep Learning-Based Framework for Automated Classification of Knee Osteoarthritis Severity and Detection of Joint Space Width in X-Ray Imaging

Logeshwari Alavanthar[1*], Jayashree Stalin [2] and Jasmine Mystica K[3]

[1, 2, 3] Department of Electronics and Communications Engineering, St.Joseph's College of Engineering (affiliation to Anna University), OMR, Semmanchery, Chennai-600119, Tamil Nadu, India.
logeshwarialavanthar7@gmail.com

Abstract. Knee osteoarthritis (OA) is a widespread degenerative condition impacting millions, predominantly older adults, and frequently resulting in chronic pain and a compromised quality of life. Timely and precise diagnosis is crucial for effective disease management, mitigating its progression, and enabling the implementation of targeted therapeutic interventions. In this paper we present a novel framework that employs deep learning for analysis of digital X-ray images to identify and classify knee OA. The method solves two problems, the measurement of Knee Joint Space Width (JSW) an essential marker for OA, and determining the severity of an OA knee using the conventional Kellgren-Lawrence (KL) grading method. To do so, we utilize the YOLOv8 model, which is known for its rapid and accurate detection for locating and segmenting the JSW region from X-ray images. Upon detection of the JSW, a VGG16 neural network augmented through transfer learned is applied. To classify the OA severity according to KL grades the system automatically delivers useful diagnostic information to clinicians to enhance the meaningfulness of OA progression assessments and to facilitate earlier and more successful intervention. Our framework harnesses the power of advanced deep learning techniques, to support improved patient outcomes, and increase diagnostic precision.

Keywords: Knee osteoarthritis severity classification, Deep Learning, YOLOv8, VGG16, KL grading.

1 Introduction

Knee osteoarthritis (OA) is one of the major degenerative joint diseases causing progressive breakdown of knee cartilage, resulting in pain, stiffness, and impaired function that adversely affect the quality of life of hundreds of millions of people, especially those aged 65 years and above. According to the World Health Organization (WHO), more than 300 million people are affected worldwide by osteoarthritis, where the knee joint is one of the most commonly involved. Knee OA is one of the most common causes of disability in elderly adults and increases in prevalence with age, which necessitates timely and precise diagnosis for efficient targeted therapy and improved quality of life. Traditionally, diagnosing knee OA has relied on clinicians manually interpreting radiological images, often using the widely recognized Kellgren-Lawrence (KL) grading system. The system categorizes the knee OA into 5 severity levels based on visual markers such as joint space narrowing, osteophyte formation, Structural alterations within the joint, encompassing the deterioration of articular cartilage. Thickening of the synovial membrane, and the formation of osteophytes. As the cartilage erodes, the underlying bone becomes exposed and may

© The Author(s) 2025
S. Kannadhasan et al. (eds.), Proceedings of the International Conference on Sustainability Innovation in Computing and Engineering (ICSICE 24), Advances in Computer Science Research 120,
https://doi.org/10.2991/978-94-6463-718-2_34

undergo sclerosis, a hardening process that can alter its texture and density. While effective, manual the systematic process of assessing, and determining the value, quality, or effectiveness of something based on established criteria or standard time consuming and subjective, leading to inconsistencies between different radiologists. Such variability in assessments can delay accurate diagnoses and affect timely patient care. In recent years, automated diagnostic technologies have emerged to address these challenges.

Such systems typically apply conventional image processing methods such as edge detection and template matching to identify metrics like joint space width (JSW) from X-ray imageries. In order to address these challenges, we propose a deep learning-based A methodology for diagnosing and classifying knee osteoarthritis using digital X-ray imaging. Our approach builds on top of modern convolutional neural networks (CNNs) to automatically perform two important tasks: first, identifying the JSW region, an important indicator of OA, and second, quantifying that information in KL (Kellgren-Lawrence) grading. In detail, the YOLOv8 model is used to quickly and accurately locate the region of the joint space width (JSW) in X-ray photos.

The features extracted are then classified by the VGG16 model that uses transfer learning techniques to classify OANN severity. Such an automated system can drastically change the face of knee OA management by helping clinicians arrive at more uniform and reproducible diagnoses, facilitating earlier interventions, thereby improving the outcome of the patient. This work is organized Section II as review the literature survey. Section III describes the methodology in terms of features and functionality. Section IVcontains results and discussion, analysing the efficacy of the system. Finally, §V concludes with the key findings and future implications.

2 Literature survey

Osteoarthritis (OA) is a progressive degenerative condition that profoundly impacts joint function, leading to diminished mobility and a reduced quality of life. In response to the growing prevalence of OA, numerous studies have been dedicated to developing effective techniques for its early detection and classification, aiming to enhance therapeutic interventions and improve patient prognosis. This review consolidates the latest advancements, approaches, and methodologies within this field, offering a comprehensive overview of current diagnostic tools, treatment strategies, and research trends designed to optimize the management of OA and mitigate its long-term effects on affected individuals.

Anifahetal. [1] presented a technique combining Contrast Limited Adaptive Histogram Equalization grade the severity. The approach yielded high correctness in predicting early KOA stages, with classification rates of 93.8% for KL for each OA is grade 0, indicating no visible 70% to grade 1, signifying doubtful; 4% to grade 2, which marks minimal; 10% to grade 3, identifying moderate; and 88.9% to grade 4, representing severe OA. Although accurate in identifying both mild (grade 0) and advanced stages (grade 4), the method showed lower efficacy in identifying intermediate stages, particularly grade2.

In their study, Chen et al. [2] worked the YOLOv2 for automated recognition of knee joints in X-ray images, exploring a variety of classification networks fine-tuned to identify KOA severity, including VGG, ResNet, and DenseNet. Their model's Mean

Absolute Error (MAE) was 0.344, and its maximum exactness was 69.7%. This research provided insights into fully automated knee joint detection and emphasized the network's flexibility across different architecture choices, showcasing moderate correctness in early-stage OA detection.

Thomas et al. [3] established a CNN to grade osteoarthritis using a large radiograph dataset consisting of 32,116 training images, 4,074 tuning images, and 4,090 testing images. This model attained a percentage of 71% F1 score of 70% indicates a balanced performance of the model, where it effectively manages both precision and recall. This means the model accurately discerns 70% of the pertinent cases. The F1 score reflects the model's overall ability to accurately classify data, with a score of 70% demonstrating reasonable effectiveness in the given task. Underscoring its strength in learning features related to KOA severity. The study illustrated how large datasets contribute to effective model training and highlighted the potential of CNNs in addressing variability across X-ray images of knees affected by different OA severities.

Another study [4] applied ResNet combined with a Convolution Block Attention Module (CBAM) for knee OA grading, leveraging the Initiative (OAI) X-ray image dataset. The model's quadratic kappa score was 0.88, its MSE was 0.36, and its total veracity was 74.81%. For KL grade 1, the classification authenticity was 48.66%, whereas for KL grade 4, it was 90.21%. Notably, these results suggest that while the model effectively identifies severe OA (grades 3 and 4), further refinements could improve classification in intermediate stages.

A different approach examined the impact of including clinical information in KOA classification, where researchers [5] The research examined a convolutional neural network (CNN) architecture integrated with six Squeeze-and Excitation Residual Network (SE-ResNet) modules. Using a private dataset with 3,464 training images, 516 testing images, and 386 validation images, the team conducted two experiments. The first utilized only imaging data, achieving AUC values of 0.91, 0.80, 0.69, 0.86, and 0.96 across KL grades 0 to 4. The second incorporated both imaging and clinical data, yielding improved AUCs of 0.97, 0.85, 0.75, 0.86, and 0.95 for each grade. Findings indicated that KL grade 2 was particularly challenging to predict due to high variability within this grade. This study illustrated that additional clinical data improved model performance, offering a more comprehensive approach to KOA severity prediction.

In another investigation [6], U-Net was used to localize left and right knee joints within radiographs from the OAI dataset, followed by classification using Dense Net with demographic data, including age, gender, and BMI. With 25,873 training images, 7,779 validation images, and 5,941 test images, the model achieved sensitivities of 83.7% for normal, 70.2% for mild, 68.9% for moderate, and 86.0% for severe OA, with specificities of 86.1% (normal), 83.8% mild cases, 97.1% moderate cases, and 99.1% severe cases. KL grade 1 was excluded from the analysis due to classification challenges, and inter-observer reliability varied between 0.51 and 0.89, indicating inconsistencies, especially in borderline KOA stages. This study highlighted the importance of differentiating adjacent KL grades, as these were often misclassified.

Finally, a trasfer learning approach with a pre-trained ResNet-34 model was proposed to predict KL grades and OA progression [7]. The study leveraging patient data and knee radiographs from 728 OAI participants created predictive models for KOA

severity and development based on age, ethnicity, BMI, and gender. This transfer learning method achieved AUCs of 0.93, 0.80, 0.88, 0.96, and 0.99 for KL grades 0 - 4, respectively. This model exhibited high predictive reliability, outshined in severe OA stages and was the first one proposing transfer learning models to embrace the classification robustness in medical imaging assignments by leveraging pre-trained architectures.

Descriptors2 Genetic markers9 Genetic markers have been investigated to assist with early osteoarthritis diagnosis especially in high-risk groups. [8] Badshah et al. studied genetic markers in the Pakistani population and discovered traits that contribute to early detection. Their insights suggest the potential for tools to target diagnosis and treatment for individuals with similar genetic backgrounds. The findings of this study emphasize the importance of genetic profiling in the management and prevention of osteoarthritis and emphasize the need for local studies, as genotypes differ from population to population. This will ensure that evidence remains updated and leads to better healthcare approaches considering the regional and ethnic demands.

New methods of diagnosing osteoarthritis use advances in machine vision and image analysis. Gornale et al. [9] detected osteoarthritis using a small retrospective cohort of knee X-ray images. The study further contributed to the extraction and analysis of significant features in relevant classes within X-ray images which were found to improve the accuracy in recognising osteoarthritic transformations using machine vision techniques. Such research highlights how automated systems can minimise diagnostic subjectivity, and allow early detection. Our work underscores the contribution offered by computational approaches to medical imaging, illustrating how these emerging tools can assist clinicians in potentially more timely and cost-effective identification of osteoarthritis.

Diagnostic and prognostic revolution in knee osteoarthritis: imaging and machine-learning techniques. Teoh et al.

Compared to signs-based approaches, such as manual grading of the imaging and machine learning methods, which identify critical diagnostic features [10]. Their study emphasized machine learning's potential to automate the identification of characteristics used to predict and improve prediction accuracy. It highlighted the pros and cons of both traditional and computational methods and provided suggestions for integrating advanced technologies into clinical workflows. The results showed that almost all articles demonstrated the transformative nature of ML in OA diagnosis in terms of improving diagnostic accuracy, decreasing subjectivity of the reads and allowing tailored management of OA, which provided more efficient and accurate healthcare solutions.

3 Methodology

This comprehensive pipeline for knee OA detection and classification utilizes a triple approach utilizing state-of-the-art deep learning techniques to enhance the verification of diagnosis and patient monitoring and care. The first step is to gather a diverse database of digital X-ray images of knees from the same medical source. The dataset encompasses diverse knee conditions and varying image qualities to enable the model to perform well in real-life situations. Detection of Joint Space Width (JSW) Region.

The next object to detect is the multiple features in the JSW, which can be a key factor in detecting OA. In this task, YOLOv8 model is used which is used for quick and clear object detection. Annotated X-ray images emphasizing JSW regions, and the corresponding model trained on this labeled data. Fig. Overall proposed methodology for knee osteoarthritis classification is shown in fig1.

3.1 Data Collection

Data collection forms the backbone of this study. High-quality X ray images datasets were collected from public repositories, such as Kaggle. Data collection is a crucial step in developing an effective diagnostic system for OA. Digital X-ray images of knee joints were sourced from multiple reliable repositories, including publicly available medical imaging databases and institutional archives. These datasets were selected to ensure a diverse representation of knee conditions, including healthy joints and varying stages of OA. Efforts were made to include data from patients of different ages, genders, and ethnicities to enhance the model's generalizability. Images were also gathered from a range of imaging equipment and techniques, capturing variations in resolution, quality, and positioning, which reflect real-world clinical scenarios.

3.2 Preprocessing

Data preprocessing is the first step to ensure that actually, the dataset has the ideal structure to apply for further analysis or modelling. The raw MRI scans were standardized using a set of transformations. All the images are uniformly resized into 224 by 224 pixels because, like all other Vision Transformers, it utilizes a standard dimension for a resolution. Intensity normalization accounted for variations in scanner settings and ensured comparability of the values in pixel intensity. Preprocessing is a crucial step in preparing medical images for analysis, particularly when applying deep learning models to tasks like identifying and classifying osteoarthritis. Preprocessing techniques were carefully implemented to standardize and enhance the quality of knee joint X-ray images, ensuring the models performed at their best. These steps are crucial in enhancing the level of detail and robustness of deep learning algorithms Finally, the preprocess data was divided into training (70%), validation set of 15%, and testing set of 15% to allow for training and evaluation of the proposed models without overfitting.

3.3 Knee Joint Space Width (JSW) Region Detection using YOLOv8

This module uses the high performance deep learning mode YOLOv8 to detect and localize the Knee Joint Space Width (JSW) region that exists in knee X-ray image. The first step of the process is to manually annotate the JSW region, which is the area of interest in the Radiographic knee images. These annotations provide the ground truth for training the task model, which enables the model to correctly predict and localize the JSW at inference. After annotation, the dataset is formatted for YOLOv8 via Robo flow, which is immensely helpful in simplifying data preprocessing and integration with the model.

The JSW landmarks serve as ground truth for task model training, which allows the model to learn to predict and localize the JSW during inference. After annotation, the dataset is prepared to YOLOv8 compatible formatting through Robo flow, which makes the data preprocessing and the inputing on the model really easy.

This method consists of merging four images into one at training time, leading the model to learn to detect objects in different conditions like other placements of the object, partial occlusion from the object, and surrounding pixels. It is worth pointing out that the mosaic augmentation is done online, i.e., during each epoch of training, on the fly. It enables the model to generalize better to new and unseen images, making it more robust. These improvements and techniques work together to make sure that the YOLOv8 model can accurately detect and localize the JSW region, which is important for diagnosing knee osteoarthritis.

3.4 Knee Osteoarthritis Classification using VGG16

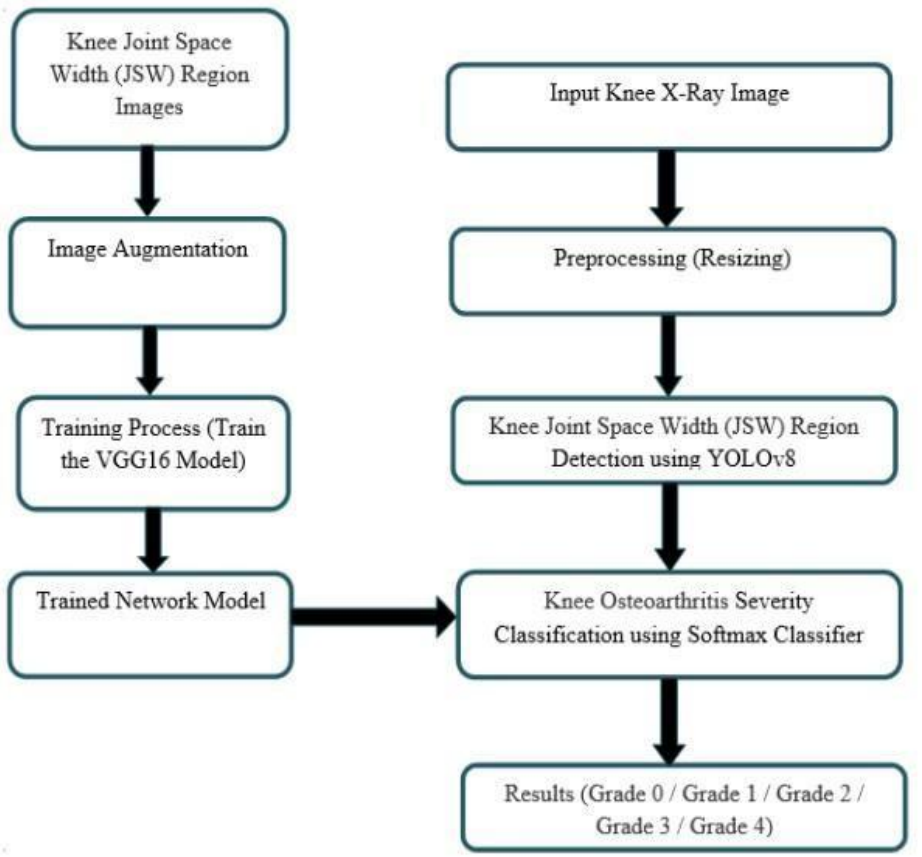

Fig. 1. Overall proposed methodology for knee osteoarthritis classification.

In this module, (OA) is most commonly classified according to its severity, an important consideration for appropriate treatment options and intervention strategies. Deep learning techniques is applied to the data sets to classify OA correctly, VGG16. This process is based on enhancing the VGG16 convolutional neural network (CNN). Fig. 1 demonstrates the configuration of the proposed solution: Pre-processed Knee X-ray images are fed into VGG16 after Knee Joint Space Width (JSW) detection using YOLOv8 model. Feature extraction modules are fine-tuned to classify OA severity, according to the Kellgren Lawrence (KL) scoring system, which ranges from 0 (no OA) to 4 (severe) via 1 (doubtful), and 2 (mild) and 3 (moderate). It helps in accurate diagnosis and also treatment planning.

4 Result and discussion

This data structure reports the Results of an evaluation of knee osteoarthritis (OA) diagnosis and classification. VGG16 model was evaluated to classify OA severity. These measures give insight into the performance of the model in recognizing and categorizing the various stages of OA. Besides the evaluation of classification, we also assessed the performance of YOLOv8 on obtaining the region Knee Joint Space Width (JSW). To evaluate the JSW detection, we used mean Average Perfection (mAP) metrics, along with perfection and revisiting, which measure how precisely the model identifies the JSW area in knee X-ray photos. These findings enable further insight into the performance of the deep learning knee OA and emphasize their potential utility to inform clinical decision making. The dataset used in this work included images of five grades of severity, and was obtained from a public Kaggle repository. Table 1: Performance YOLOv8 and metrics are shown: mAP, correctness and reminding mAP, correctness and reminding mAP [1].

This stage shows results of organ segmentation; it clearly shows the convincing performance of detection of the Knee Joint Space Width (JSW) region. In the results, we see that the YOLOv8 model accurately detected the JSW of knee X-ray with an MAP score of 0.995 at an Intersection over Union (IOU) threshold, also known as the IOU criterion or IOU cut-off. Or, restating it, the overlap ratio threshold or the overlap threshold for IoU=0.5. Fig. illustrates this whole process from input image through JSW region detection and final classification output. 2. Table 2 summarizes the quantitative performance metrics, providing key insights into the VGG16 model efficacy. The VGG16 achieved a precision of 92.5%, recall of 90.8%, accuracy of 91.3% and F1-score of 91.0%. As shown in Fig. 2, the ROC curve analysis indicates that the EF and TGC scores retained the same predictive capacity for destruction effect in the train test set. 3 gives a graphical overview of the performance in dividing rigidity at different threshold settings. The AUC of our model was 0.9427, indicating that the model can effectively discriminate among the different severity classes of knee OA in our analysis. This analysis reinforces the ability of the proposed learning model accurately to ascertain and classify knee OA severity, highlighting their potential clinical utility in guiding both timely and appropriate treatment practice.

Table. 1. Performance metrics.

Metrics	Performance
MAP	0.995
Precision	0.988
Recall	1.000

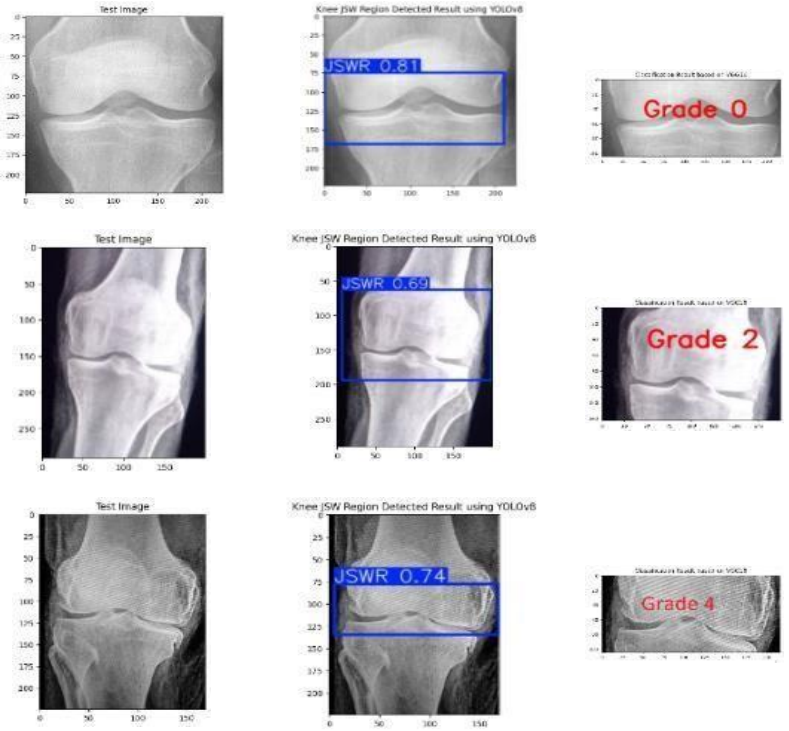

Fig. 2. Knee OA Severity Classification Results.

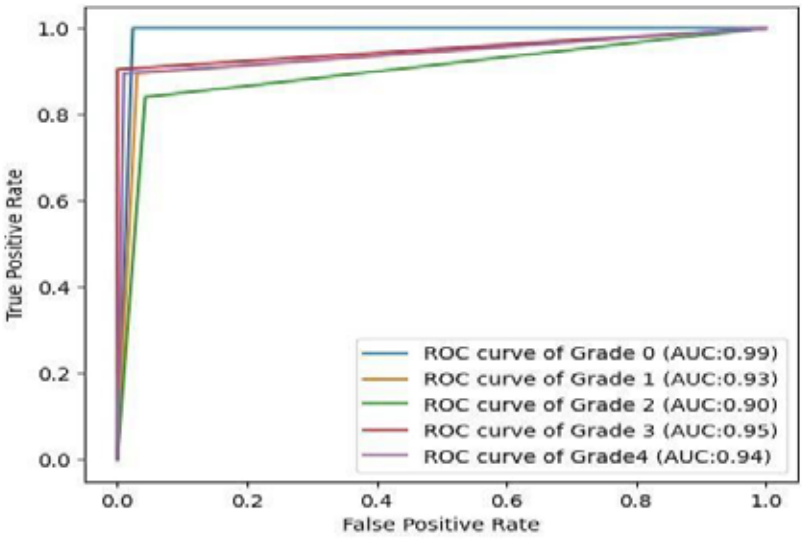

Fig. 3. ROC curve analysis for knee OA classification using VGG16 model.

Table 2. Performance analysis of the VGG16 model for knee OA classification.

Metrics	Performance
Accuracy	0.9153
Recall	0.9143
Precision	0.9236
F1-score	0.9123

5 Conclusion

This study proposed deep learning-based enhanced artificial intelligence (AI) approach for (OA) detection and categorization from digital X-ray images. By leveraging the vision of the YOLOv8 model, we successfully localized the former, which is a key metric of OA severity, the JSW area. Well, the YOLOv8 model was immune to this missing JSW area, we achieved a mAP of 0.995 with this model. Similarly, a transfer learning based approach to classify the severity of knee osteoarthritis using the Kellgren Lawrence (KL) algorithm grading framework was implemented using VGG16 model. The performance metrics of the VGG16 model were impressive as authenticity was 91.53%, memory was 91%, exactness was 92%, and F1 score was 91%. The ROC curve analysis also confirmed the discriminative ability of the model to discriminate between individuals of different severity of knee OA. This has tremendous implications for the improvement of patient outcomes or early interventions if implemented in clinical settings. Future research will focus on increasing the dataset, improving the models and testing other imaging types to further improve the robustness of knee OA classification.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. L. Anifah, I. K. E. Purnama, M. Hariadi, and M. H. Purnomo, "Osteoarthritis classification using self -organizing map based on Gabor kernel and contrast-limited adaptive histogram equalization," Open Biomed. Eng. J., vol. 7, no. 1, pp. 18–28, Feb. 2013.
2. P. Chen, L. Gao, X. Shi, K. Allen, and L. Yang, "Fully automatic knee osteoarthritis severity grading using deep neural networks with a novel ordinal loss," Computerized Med. Imag. Graph., vol. 75, pp. 84–92, Jul. 2019.
3. K. A. Thomas, L. Kidzinski, E. Halilaj, S. L. Fleming, G. R.Venkataraman, E. H. G. Oei, G. E. Gold, and S. L. Delp, "Automated classification of radiographic knee osteoarthritis severity using deep neural networks," *Radiol., Artif. Intell.*, vol. 2, no. 2, Mar. 2020, Art. no. e190065.

4. B. Zhang, J. Tan, K. Cho, G. Chang, and C. M. Deniz, "Attention-based CNN for KL grade classification: Data from the osteoarthritis initiative," in Proc. IEEE 17th Int. Symp. Biomed. Imag. (ISBI), Apr. 2020, pp. 731–735.
5. H. Kim, K. J. Lee, D. Choi, J. I. Lee, H. G. Choi, and Y. S. Lee, "Can additional patient information improve the diagnostic performance of deep learning for the interpretation of knee osteoarthritis severity," J. Clin. Med., vol. 9, no. 10, p. 3341, Oct. 2020.
6. B. Norman, V. Pedoia, A. Noworolski, T. M. Link, and S. Majumdar, "Applying densely connected convolutional neural networks for staging osteoarthritis severity from plain radiographs," J. Digit. Imag., vol. 32, no. 3, pp. 471–477, Jun. 2019.
7. K. Leung, B. Zhang, J. Tan, Y. Shen, K. J. Geras, J. S. Babb, K. Cho, G. Chang, and C. M. Deniz, "Prediction of total knee replacement and diagnosis of osteoarthritis by using deep learning on knee radiographs: Data from the osteoarthritis initiative," Radiology, vol. 296, no. 3, pp. 584– 593, Sep. 2020.
8. Y. Badshah, M. Shabbir, H. Hayat, Z. Fatima, A. Burki, S. Khan, and S. U. Rehman, "Genetic markers of osteoarthritis: Early diagnosis in susceptible Pakistani population," J. Orthopaedic Surgery Res., vol. 16, no. 1, pp. 1–8, Dec. 2021.
9. S. S. Gornale, P. U. Patravali, and R. R. Manza, "Detection of osteoarthritis using knee X-ray image analyses: machine vision based approach," Int. J. Comput. Appl., vol. 145, no. 1, pp. 20–26, 2016.
10. Y. X. Teoh, K. W. Lai, J. Usman, S. L. Goh, H. Mohafez, K. Hasikin, P. Qian, Y. Jiang, Y. Zhang, and S. Dhanalakshmi, "Discovering knee osteoarthritis imaging features for diagnosis and prognosis: Review of manual imaging grading and machine learning approaches," J. Healthcare Eng., vol. 2022, pp. 1–19, Feb. 2022.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.