Advanced Classification of Knee Osteoarthritis Using Artificial Intelligence Technologies

Facundo Segura,[*,,#,##] Florencio P. Segura,[*,**,#,##] María Paz Lucero Zudaire,[*,**,##] Florencio V. Segura,[*,#,##] Rocío Mendía,[**] Lucía Ribotta Falco,[**] Paula S. Zalazar,[**] Daniel E. Sequeira[**]**

[*]*Centro Privado Ortopedia y Traumatología Segura, 2nd. Chair of Orthopedics and Traumatology, Universidad Nacional de Córdoba, Córdoba, Argentina.*

[**]*Instituto Conci Carpinella, Córdoba, Argentina*

[#]*Instituto Médico Victoria, Córdoba, Argentina*

[##]*Instituto Modelo de Cardiología, Córdoba, Argentina*

ABSTRACT

Introduction: Knee osteoarthritis is a prevalent and debilitating musculoskeletal condition, particularly in the elderly. Early detection and accurate classification are crucial for improving patient outcomes. **Objective:** To investigate the application of artificial intelligence (AI) and computer vision for the automated detection and classification of knee osteoarthritis based on the Kellgren-Lawrence (KL) scale. Additionally, to develop and evaluate an automated system capable of accurately classifying the severity of the disease. **Materials and Methods:** A public dataset of radiographic knee images pre-classified according to the KL scale was used. The images were processed with LandingLens software, using the ConvNext architecture, a convolutional neural network. The model was trained with 995 images and was used to evaluate 240 trial images. **Results:** The model achieved an overall accuracy of 92.55% in classifying knee osteoarthritis according to the KL scale, with a sensitivity of 93.33%. Per-class accuracy was as follows: 97.87% for grade 0, 79.74% for grade 1, 88.68% for grade 2, 94.04% for grade 3, and 99.42% for grade 4. **Conclusions:** This study confirms the efficacy of AI and computer vision technologies in the automated detection of knee osteoarthritis. Integrating these technologies into clinical practice can enhance the efficiency and consistency of patient evaluations, ultimately leading to improved clinical outcomes and more personalized care.

Keywords: Artificial intelligence; osteoarthritis; knee classification; computer vision.

Level of Evidence: II

Clasificación avanzada de la artrosis de rodilla utilizando tecnologías de Inteligencia Artificial

RESUMEN

Introducción: La artrosis de rodilla es una enfermedad osteoarticular prevalente y debilitante, especialmente en adultos mayores. Su detección temprana y la clasificación precisa son cruciales para mejorar los resultados clínicos. **Objetivos:** Investigar el uso de la inteligencia artificial y la visión por computadora para la detección y clasificación automatizada de la artrosis de rodilla según la escala de Kellgren-Lawrence. Desarrollar un sistema automatizado y evaluar su precisión para clasificar la gravedad de la enfermedad. **Materiales y Métodos:** Se utilizó un conjunto de datos públicos con imágenes radiográficas de rodillas clasificadas según la escala de Kellgren-Lawrence. Las imágenes fueron procesadas con el programa LandingLens, empleando la arquitectura ConvNext, una red neuronal convolucional. El modelo fue entrenado con 995 imágenes y evaluado con 240 imágenes de prueba. **Resultados:** El modelo alcanzó una precisión global del 92,55% en la clasificación de la artrosis de rodilla, con una sensibilidad del 93,33%. La precisión por clase fue del 97,87% para el grado 0; 79,74% para el grado 1; 88,68% para el grado 2; 94,04% para el grado 3 y 99,42% para el grado 4. **Conclusiones:** El estudio confirma la eficacia de la inteligencia artificial y la visión por computadora en la detección automatizada de la artrosis de rodilla. La integración de estas tecnologías en la práctica clínica podría mejorar la eficiencia, la consistencia en la evaluación de los pacientes y los resultados clínicos, y así favorecer una atención médica más personalizada.

Palabras clave: Inteligencia artificial; artrosis; clasificación; visión por computadora.

Nivel de Evidencia: II

Received on June 16th, 2024. Accepted after evaluation on July 21st, 2024 • Dr. FACUNDO SEGURA • facusegura@gmail.com https://orcid.org/0009-0000-7101-9145

How to cite this article: Segura F, Segura FP, Lucero Z MP, Segura FV, Mendía R, Ribotta Falco L, Zalazar PS, Sequeira DE. Advanced Classification of Knee Osteoarthritis Using Artificial Intelligence Technologies. *Rev Asoc Argent Ortop Traumatol* 2024;89(5):462-469. https://doi.org/10.15417/issn.1852-7434.2024.89.5.1993. https://doi.org/10.15417/issn.1852-7434.2024.89.5.1993

INTRODUCTION

Osteoarthritis is the most common joint disorder among adults over 60 years of age in the United States. The prevalence of symptomatic knee osteoarthritis is approximately 10% in men and 13% in women.[1] The knee is undoubtedly one of the most affected joints, with estimates indicating that around 30% of people over 45 years old have radiographic evidence of this condition, and about half of them experience clinical symptoms.[2]

Common symptoms include joint pain, stiffness, and limited function, which negatively impact quality of life and are often associated with comorbidities such as cardiovascular disease, diabetes, hypertension, falls, fractures, and depression.[3] Managing this disease requires a multidisciplinary approach, including pain management, exercise, physiotherapy, and, in severe cases, surgical intervention to improve clinical outcomes and enhance patients' quality of life.[4]

In recent years, artificial intelligence (AI), particularly deep learning and convolutional neural networks (CNNs), has emerged as a powerful tool for improving the accuracy of knee osteoarthritis diagnosis and classification. These AI models can process large volumes of imaging data and learn complex features, enabling more accurate and objective classification of knee radiographs.[5,6]

Several CNN-based approaches have been developed for classifying and detecting knee osteoarthritis. These models have been trained to identify and classify X-ray images of the knee according to the Kellgren-Lawrence (KL) classification system, the most widely used method for determining osteoarthritis severity. Studies using various CNN architectures, such as VGG16, VGG19, ResNet50, YOLOv3, and EfficientNet-B5, have demonstrated high accuracy in classifying the severity of osteoarthritis.[7,8]

LandingLens, developed by Landing AI, is an advanced computer vision platform designed to simplify the creation, implementation, and management of AI models, even for users with no prior experience in AI or machine learning. This platform is particularly useful in industrial environments for tasks such as quality inspection and defect detection.

The software provides an intuitive interface that guides users through image uploading, labeling, training, and model deployment. It allows the development of computer vision models without requiring deep technical knowledge. Additionally, it supports continuous learning, enabling models to automatically update with new data. This ensures that models adapt to changing conditions and improve over time.

Radiographic grading systems for knee osteoarthritis vary in terms of reliability and sensitivity in detecting the severity of the disease. The most commonly evaluated systems include the Kellgren-Lawrence, International Knee Documentation Committee (IKDC), Ahlbäck, and Fairbank scales (Table 1).[9]

The Kellgren-Lawrence system is widely used and has proven reproducible, especially when assessing radiographic features such as osteophytes and joint space narrowing. The IKDC system, meanwhile, is known for its high interobserver and intraobserver reliability, making it one of the most dependable methods for evaluating knee osteoarthritis.

The Fairbank system, on the other hand, has the lowest reliability among the rating scales reviewed.[10,11]

While the Ahlbäck grading system shows good interobserver agreement, it lacks reliability, particularly when used without clinical or arthroscopic examination support.[12]

AI and computer vision provide opportunities to more effectively monitor disease progression, allowing for timely and personalized interventions. Early identification of knee osteoarthritis is critical to mitigating its effects. Early detection, coupled with patient education on exercise and weight management, can significantly reduce the symptoms of the disease.[13]

The integration of AI in the diagnosis and treatment of knee osteoarthritis marks a significant advancement in the field.

This paper aims to demonstrate how deep learning and computer vision techniques can be used to develop a model for detecting and classifying knee osteoarthritis based on the Kellgren-Lawrence classification.

Table 1. Knee osteoarthritis classification systems and their characteristics.

Classification	Grade and characteristics				
Kellgren-Lawrence	0: No JSN or reactive changes	1: Doubtful JSN, possible osteophytic lipping	2: Definite osteophytes, possible JSN	3: Moderate osteophytes, definite JSN, some sclerosis, possible bone-end deformity	4: Large osteophytes, marked JSN, severe sclerosis, definite bone ends deformity
IKDC	A: No JSN	B: >4 mm joint space; small osteophytes, slight sclerosis, or femoral condyle flattening	C: 2-4 mm joint space	D: <2 mm joint space	
Fairbank	0: Normal	1: Squaring of tibial margin	2: Flattening of femoral condyle, squaring and sclerosis of tibial margin	3: JSN, hypertrophic changes, or both	4: All of the characteristics at left, to a more severe degree
Brandt et al.	0: <25% JSN without secondary features (subchondral sclerosis, geodes, and osteophytes)	1: <25% JSN with secondary features or 25%-50% JSN without secondary features	2: 25%-50% JSN with secondary features or 50%-75% JSN without secondary features	3: 50%-75% JSN with secondary features or >75% JSN without secondary features	4: >75% JSN with secondary features
Ahlbäck	0: Normal	1: JSN† (with or without subchondral sclerosis).	2: Obliteration of joint space	3: Bone defect/ loss <5 mm	4: Bone defect and/or loss 5-10 mm
Jäger-Wirth	0: No osteoarthritis	1: Initial osteoarthritis, small osteophytes, minimal JSN	2: Moderate osteoarthritis, about 50% JSN	3: Medium-grade osteoarthritis	4: Heavy osteoarthritis

*JSN = joint space narrowing, IKDC = International Knee Documentation Committee.
†Joint space narrowing is <3 mm of the joint space or <50% of the other compartment.
From: Kohn MD, Sassoon AA, Fernando ND. Classifications in Brief: Kellgren-Lawrence Classification of Osteoarthritis. Clin Orthop Relat Res 2016;474(8):1886-93. https://doi.org/10.1007/s11999-016-4732-4.

OBJECTIVE

To develop an automated system for the detection and classification of knee osteoarthritis using the Kellgren-Lawrence scale and the IKDC questionnaire, using a computer vision-based AI program.

MATERIALS AND METHODS

A publicly available dataset containing medical images of patients with varying degrees of knee osteoarthritis (https://www.kaggle.com/datasets/shashwatwork/knee-osteoarthritis-dataset-with-severity) was used to develop a predictive model for the disease. These images had already been classified according to the Kellgren-Lawrence scale.

The images were uploaded to the LandingAI platform as supervised learning data.

To develop the model based on the Kellgren-Lawrence classification, 1,195 knee osteoarthritis images were used, divided into five groups according to their classification: 328 grade 0, 153 grade 1, 212 grade 2, 329 grade 3, and 173 grade 4 images. The training set consisted of 995 images, while 240 images were reserved for testing.

A machine learning model based on the ConvNext convolutional neural network architecture, with 16 million parameters, was implemented to perform the prediction. The model was trained for 15 epochs. To evaluate the model, a confusion matrix, accuracy, sensitivity for each class, and the F1 score were used.

RESULTS

The performance of the knee osteoarthritis classification model was evaluated using the Kellgren-Lawrence scale, which comprises five levels of severity: 0, 1, 2, 3, and 4. A dataset of 1,195 images was used to train and evaluate the model.

The model achieved an overall accuracy of 92.55%, demonstrating its ability to correctly classify the presence and severity of knee osteoarthritis in the majority of cases. This high level of accuracy suggests the model is reliable in distinguishing between different grades of osteoarthritis according to the Kellgren-Lawrence scale (Figure 1).

When evaluating per-class accuracy, the model achieved 97.87% accuracy for grade 0, correctly classifying 321 out of 328 images in this category. For grade 1, the accuracy was 79.74%, with 122 out of 153 images correctly classified. A higher accuracy was observed for grade 2 at 88.68%, with 188 out of 212 images correctly classified. For grade 3, the model attained an accuracy of 94.04%, correctly classifying 303 out of 329 images. Finally, for grade 4, the model achieved 99.42% accuracy, correctly classifying 172 out of 173 images (Figure 2).

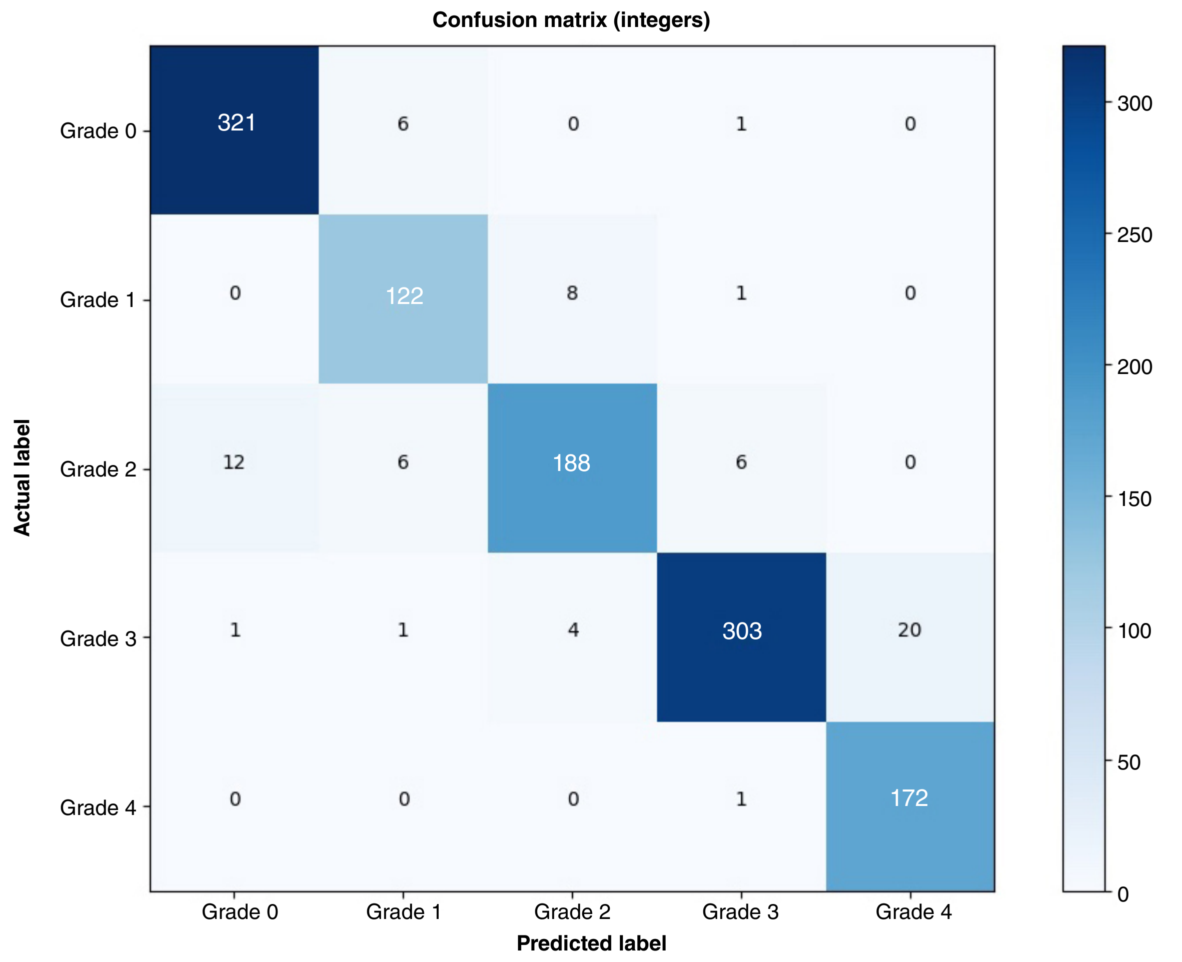

Figure 1. Confusion matrix for multi-class classification with integers. The matrix shows high classification accuracy, with 321/328, 122/153, 188/212, 303/329, and 172/173 images correctly classified for grades 0, 1, 2, 3, and 4, respectively. Misclassifications mainly occurred between adjacent classes.

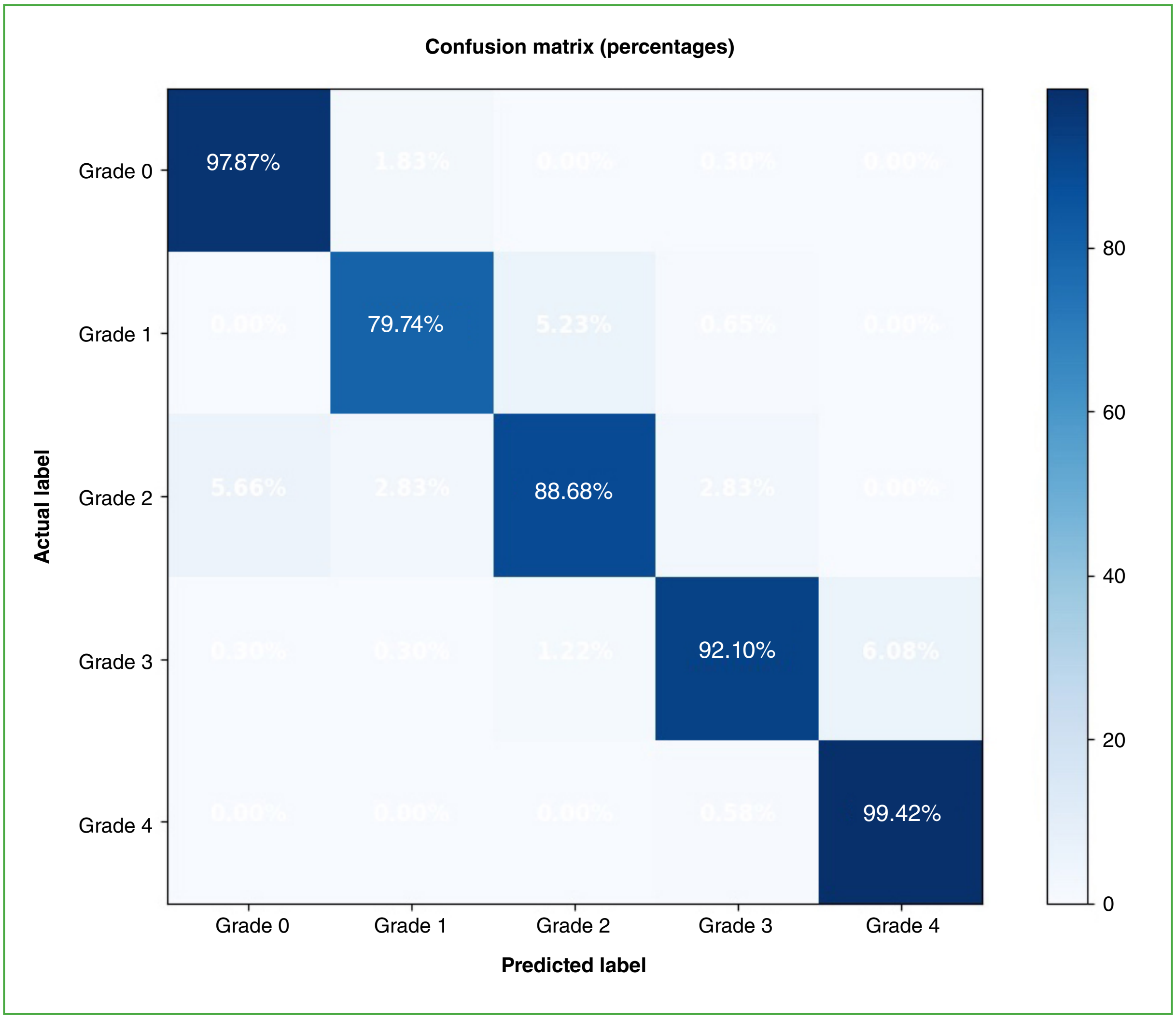

Figure 2. Confusion matrix for multi-class classification with percentage values. The percentage-based matrix indicates high classification accuracy: 97.87% for grade 0, 79.74% for grade 1, 88.68% for grade 2, 94.04% for grade 3, and 99.42% for grade 4. Misclassifications were mainly between adjacent grades.

The model's overall sensitivity was 94.23%, reflecting its capability to correctly identify most positive cases of knee osteoarthritis. Its specificity was 98.61%, indicating a strong ability to correctly identify negative cases. The F1 score, which balances accuracy and sensitivity, was 94.21%, confirming a good overall balance between the model's precision and robustness.

In Figure 3A, the model classifies all three radiographs as grade 0 with high accuracy. In Figure 3B, it can be seen how the heat map provides a useful tool for interpreting and visualizing the model's decision-making process, highlighting the areas of the image most relevant to the prediction. These heat maps are essential both for explaining the model's reasoning and for detecting potential issues in the training process or dataset. The same approach applies to other classification grades as well.

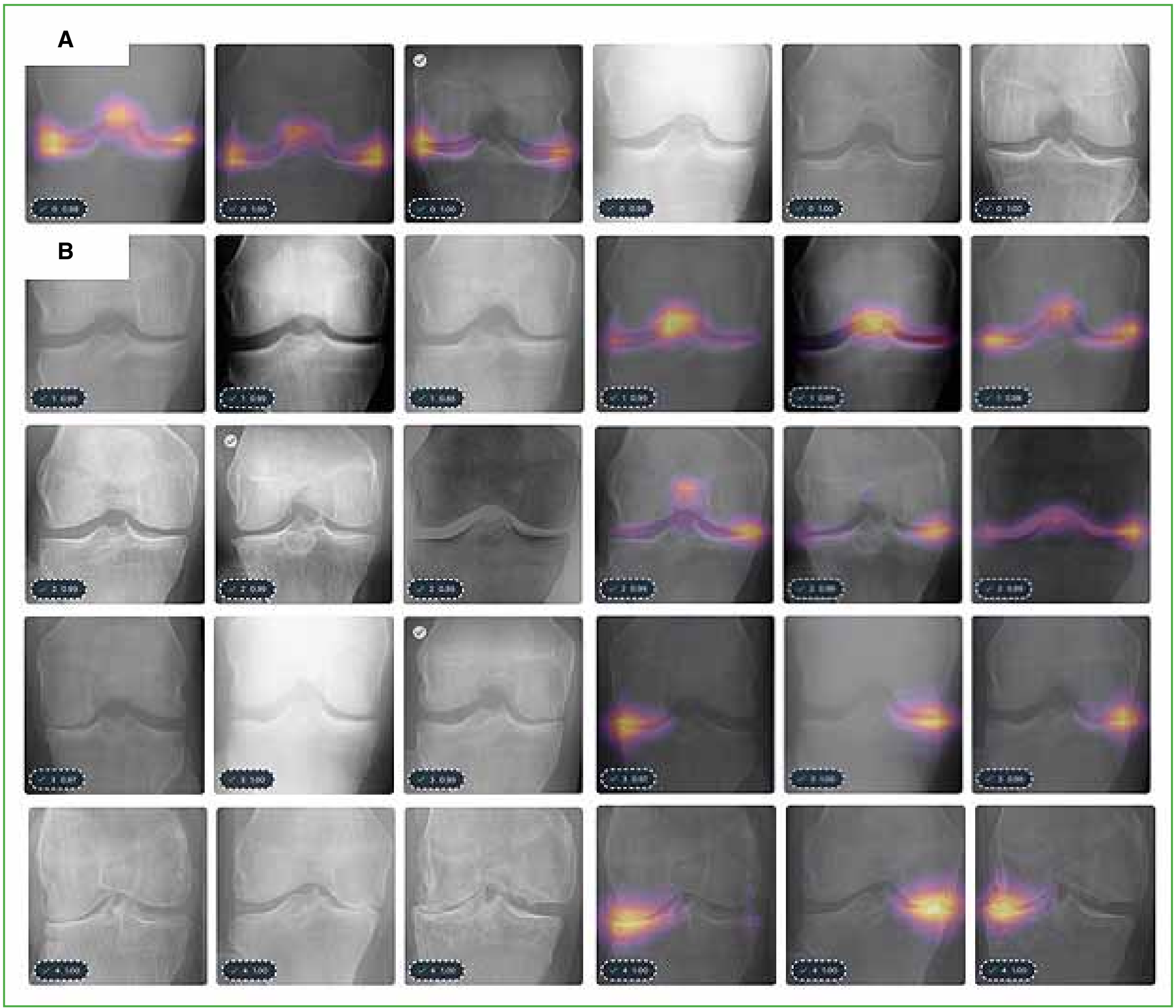

Figure 3. Heat map illustrating how the model detects areas affected by osteoarthritis and classifies the images according to the respective grades.

DISCUSSION

For many years, anteroposterior radiography has been the most effective method for classifying knee osteoarthritis for diagnostic purposes. The Kellgren-Lawrence classification system is widely used to determine disease progression and severity. However, the interpretation of these images can vary among evaluators, leading to inconsistencies in classification and, consequently, in treatment.[14]

Numerous studies have shown promising results using deep learning techniques and neural networks for the detection and classification of knee osteoarthritis. This field of AI offers an automated, objective alternative to the traditional visual interpretation of radiographic images by medical experts.

In a study by Sikkandar et al., a convolutional neural network was used for the automatic classification of knee osteoarthritis images, achieving an accuracy of 93.2% and a multiclass classification accuracy of 72.01%. This high level of accuracy suggests that convolutional neural networks can be highly effective in detecting and classifying knee osteoarthritis in radiographic images.[6]

Similarly, Deokar et al. developed an automatic knee osteoarthritis detection system based on feature extraction and neural networks, with an accuracy of 98.5% during the training stage and 92% during the testing stage. These results highlight the ability of neural networks to learn and generalize from complex features extracted from medical images.[15]

In our study, we investigated the efficacy of a computer vision-based approach for the automated detection and classification of knee osteoarthritis.

Using a dataset of 1,901 images, the model achieved remarkable accuracy and sensitivity in determining the degree of osteoarthritis. These results surpass those reported in studies by other authors using deep learning techniques and neural networks (Table 2). Without a doubt, these findings suggest that the application of computer vision techniques could offer an effective and practical alternative for evaluating knee osteoarthritis, providing accurate and rapid diagnoses that could significantly improve patient care.

Table 2. Comparison of accuracy, sensitivity and specificity results according to the number of images.

Study	Accuracy	Sensitivity	Specificity	Number of images analyzed
Brahim (2019)	82.98%	87.15%	80.65%	1024
Tiwari (2022)	93.69%	92.53%	92.87%	2068
Pongsakonpruttikul (2022)	81%	85%	85%	1650
Segura (2024)	**92.55%**	**94.23%**	**98.61%**	**1195**

The implementation of these technologies facilitates the automation of radiographic image analysis. Such programs can be trained to automatically identify and classify knee images according to osteoarthritis severity, using the Kellgren-Lawrence classification system as a standard. This not only improves diagnostic efficiency but also enhances accuracy, reducing inter-rater variability and providing clinical decision support.

CONCLUSIONS

The results of this study demonstrate the efficacy of AI and computer vision in the automated classification of knee osteoarthritis, achieving an accuracy of 92.55% and a sensitivity of 94.23%. These findings emphasize the potential of these technologies to support physicians in accurately diagnosing the disease, providing a valuable tool that can enhance both efficiency and consistency in patient evaluations.

Conflict of interest: The authors declare no conflicts of interest.

F. P. Segura ORCID ID: https://orcid.org/0000-0002-2376-4834
M. P. Lucero Zudaire ORCID ID: https://orcid.org/0009-0009-8632-480X
F. V. Segura ORCID ID: https://orcid.org/0009-0004-0424-8334
R. Mendía ORCID ID: https://orcid.org/0009-0003-3114-930X
L. Ribotta Falco ORCID ID: https://orcid.org/0009-0004-9630-7935
P. S. Zalazar ORCID ID: https://orcid.org/0009-0007-2399-0384
D. E. Sequeira ORCID ID: https://orcid.org/0009-0006-8226-8343

REFERENCES

1. Zhang Y, Jordan JM. Epidemiology of osteoarthritis. *Clin Geriatr Med* 2010;26(3):355-69. https://doi.org/10.1016/j.cger.2010.03.001
2. Lawrence RC, Felson DT, Helmick CG, Arnold LM, Choi H, Deyo RA, et al. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States: Part II. *Arthritis Rheum* 2008;58(1):26-35. https://doi.org/10.1002/art.23176
3. Zheng S, Tu L, Cicuttini F, Zhu Z, Han W, Antony B, et al. Depression in patients with knee osteoarthritis: risk factors and associations with joint symptoms. *BMC Musculoskelet Disord* 2021;22(1):40. https://doi.org/10.1186/s12891-020-03875-1
4. Bannuru RR, Osani MC. Vaysbrot EE, Arden NK, Bennell K, Bierma-Zeinstra SMA, et al. OARSI guidelines for the non-surgical management of knee, hip, and polyarticular osteoarthritis. *Osteoarthritis Cartilage* 2019:27(11):1578-89. https://doi.org//10.1016/j.joca.2019.06.011
5. Pongsakonpruttikul N, Angthong C, Kittichai V, Chuwongin S, Puengpipattrakul P, Thongpat P, et al. Artificial intelligence assistance in radiographic detection and classification of knee osteoarthritis and its severity: a cross-sectional diagnostic study. *Eur Rev Med Pharmacol Sci* 2022;26(5):1549-58. https://doi.org/10.26355/eurrev_202203_28220
6. Sikkandar MY, Begum SS, Alkathiry AA, Alotaibi MSN, Manzar MD. Automatic detection and classification of human knee osteoarthritis using convolutional neural networks. *Computers, Materials & Continua* 2022;70(3):4279-91. https://doi.org/:10.32604/cmc.2022.020571
7. Schwartz AJ, Clarke HD, Spangehl MJ, Bingham JS, Etzioni DA, Neville MR. Can a convolutional neural network classify knee osteoarthritis on plain radiographs as accurately as fellowship-trained knee arthroplasty surgeons? *J Arthroplasty* 2020;35(9):2423-8. https://doi.org/10.1016/j.arth.2020.04.059
8. Mahmoudian A, Lohmander LS, Mobasheri A, Englund M, Luyten FP. Early-stage symptomatic osteoarthritis of the knee — time for action. *Nat Rev Rheumatol* 2021;17(10):621-32. https://doi.org/10.1038/s41584-021-00673-4
9. Kohn MD, Sassoon AA, Fernando ND. Classifications in Brief: Kellgren-Lawrence Classification of Osteoarthritis. *Clin Orthop Relat Res* 2016;474(8):1886-93. https://doi.org/10.1007/s11999-016-4732-4
10. Wing N, Van Zyl N, Wing M, Corrigan R, Loch A, Wall C. Reliability of three radiographic classification systems for knee osteoarthritis among observers of different experience levels. *Skeletal Radiol* 2021;50(2):399-405. https://doi.org/10.1007/s00256-020-03551-4
11. Eckersley T, Faulkner J, Al-Dadah O. Inter- and intra-observer reliability of radiological grading systems for knee osteoarthritis. *Skeletal Radiol* 2021;50(10):2069-78. https://doi.org/10.1007/s00256-021-03767-y
12. Galli M, De Santis V, Tafuro L. Reliability of the Ahlbäck classification of knee osteoarthritis. *Osteoarthritis Cartilage* 2003;11(8):580-4. https://doi.org/ 10.1016/s1063-4584(03)00095-5
13. Kessler S, Guenther KP, Puhl W. Scoring prevalence and severity in gonarthritis: the suitability of the Kellgren & Lawrence scale. *Clin Rheumatol* 1998;17(3):205-9. https://doi.org/10.1007/BF01451048
14. Ahmed HA, Mohammed EA. Using Artificial Intelligence to classify osteoarthritis in the knee joint. *NTU Journal of Engineering and Technology* [Internet] 2022;1(3):31-40. Available at: https://www.iasj.net/iasj/download/fc5a99f585e6bbda
15. Deokar DD, Patil CG. Effective feature extraction based automatic knee osteoarthritis detection and classification using neural network. *International Journal of Engineering and Techniques* [Internet] 2015;1(3):134-9. Available at: http://www.ijetjournal.org/Volume1/Issue3/IJET-V1I3P22.pdf