Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Automatic Diagnosis of Knee Osteoarthritis Based on Kellgren-Lawrence Grades Using the ResNeXt Model

Haewon Byeon

Department of AI software, Inje University, Gimhae, 50834, Republic of Korea.
Medical Big Data Research Center, Inje University, Gimhae, 50834, Republic of Korea

Osteoarthritis (OA) is a prevalent chronic joint disease that significantly affects the knee joint and diminishes the quality of life for individuals worldwide. The Kellgren-Lawrence (KL) grading system is widely used for diagnosing knee OA; however, its reliance on subjective interpretation leads to variability and inconsistency in diagnostic outcomes. This study aimed to explore the potential of deep learning in automating the classification of KL grades of knee OA, thereby enhancing diagnostic accuracy and consistency. We employed the ResNeXt-50-32×4d architecture, known for its superior feature extraction capabilities through parallel path structures, to classify KL grades based on knee X-ray images. The study utilized a comprehensive dataset from Kaggle, which was divided into training, validation, and test sets to ensure balanced distribution across KL grades. We applied advanced data preprocessing and augmentation techniques to optimize model performance. As a result, the ResNeXt-50 model demonstrated exceptional performance, achieving the highest accuracy of 76.93% among the tested models, along with superior precision, recall, F1-score, and AUC metrics. These results underscore the model's capability to effectively learn complex patterns in medical images, offering a reliable automated diagnostic tool for clinical applications. In conclusion, the ResNeXt-50-32×4d model outperformed traditional ResNet models, highlighting its potential in automating knee OA diagnosis and supporting medical professionals. Future research should focus on evaluating the model's generalization across diverse datasets and enhancing interpretability through visualization techniques, thereby broadening its applicability in medical image analysis.

Keywords: Osteoarthritis Diagnosis, Kellgren-Lawrence Grading, Deep

Learning Automation, ResNeXt Architecture, Medical Image Analysis, Model Performance Metrics.

1. Introduction

Osteoarthritis is a prevalent chronic joint disease among the adult population worldwide, particularly affecting the knee joint and significantly diminishing individuals' quality of life [1, 2]. This condition results in pain, swelling, and stiffness due to the wear of joint cartilage, and in severe cases, it can hinder daily activities [3]. One of the most widely used methods for diagnosing knee osteoarthritis and assessing its progression is the Kellgren-Lawrence (KL) grading system, which systematically evaluates structural changes in the joint through radiographic imaging [4].

The traditional KL grading approach relies on subjective interpretation, leading to inter-observer variability in diagnostic outcomes. This variability poses a limitation in the consistency and accuracy of diagnoses, and the process is time-consuming and costly, especially when handling large datasets [5]. These methodological limitations underscore the need for an automated diagnostic system utilizing artificial intelligence.

Recent advancements in artificial intelligence and deep learning have accelerated the automation of medical image analysis, offering the potential to enhance diagnostic accuracy and reduce the workload of medical professionals [6-8]. Deep Neural Networks, in particular, have demonstrated strong capabilities in complex pattern recognition and feature extraction, successfully applied to various medical image processing tasks [9]. In this context, advanced neural network architectures like ResNeXt are gaining attention for their efficient feature extraction and classification capabilities [10].

This study developed a system that automatically predicts KL grades based on knee X-ray images using the ResNeXt model. The unique characteristics of ResNeXt are expected to outperform existing models such as ResNet-18, ResNet-34, and ResNet-50. Additionally, the predictive capabilities of the model were evaluated using various performance metrics, including accuracy, precision, recall, F1 score, and AUC.

This paper is structured as follows: The second section discusses related work and recent trends in medical image analysis using deep learning. The third section describes the dataset and preprocessing methods used in the study. The fourth section details the architecture and training strategy of the proposed ResNeXt model. The fifth section presents the experimental results and analyzes the model's performance. Finally, the conclusion summarizes the main findings of the study and suggests directions for future research.

2. Related Work

Research on the automatic diagnosis and grading of Knee Osteoarthritis (OA) has evolved significantly with the incorporation of deep learning methodologies, moving beyond traditional image analysis techniques. Traditionally, the diagnosis of knee OA has relied on the Kellgren-Lawrence (KL) grading system, which is inherently subjective and prone to variability among observers. This subjectivity poses challenges for consistent and reliable diagnosis across different healthcare professionals [11].

To address these limitations, recent research has introduced deep learning-based diagnostic systems, which offer more objective and consistent diagnostic outcomes by automating the classification process. For instance, novel deep learning models have been developed to predict and classify the severity of knee OA using MRI and other imaging modalities, achieving high levels of accuracy and consistency comparable to human experts [12]. These models significantly enhance the diagnostic workflow by providing consistent assessments of OA severity, thus supporting healthcare providers in making more informed decisions [13].

Furthermore, advancements in convolutional neural networks (CNNs) have also demonstrated effectiveness in approaching the automated grading of knee OA, highlighting improvements in diagnostic accuracy even amidst visual disturbances typical in radiographic images. This practical application underscores the potential of deep learning in enhancing the reliability and efficiency of knee OA assessments [14]. By leveraging these technologies, healthcare systems can benefit from standardized, rapid, and accurate assessments of osteoarthritis, reducing the manual workload of clinicians and potentially improving patient outcomes through timely intervention [15].

In the early stages of research on KL grade prediction using deep learning, satisfactory outcomes were achieved by employing pre-trained CNN models through transfer learning. Antony et al. approached the KL grade prediction problem as a regression task using pre-trained CNNs, demonstrating performance surpassing traditional methods. Subsequent studies proposed training CNNs from scratch by combining cross-entropy loss and mean squared error loss, further improving performance.

Recent research has introduced ordinal loss as a novel loss function for KL grade prediction, replacing the traditional cross-entropy loss and enhancing performance. This approach formulates the KL grade prediction problem as an ordinal regression task, contributing to the optimization of neural network performance. Additionally, studies utilizing various CNN architectures have shown that predictive performance can be further enhanced not only by individual model performance but also through model ensembles.

In the early stages of research on KL grade prediction using deep learning, satisfactory outcomes were achieved by leveraging pre-trained Convolutional Neural Networks (CNNs) through transfer learning. Antony et al. were among the pioneers who approached the KL grade prediction problem as a regression task using pre-trained CNNs, demonstrating performance that surpassed traditional methods by efficiently capturing intricate patterns in the radiographic data [16]. This foundational work proved critical in setting the stage for further innovations.

Subsequent studies have advanced this field by proposing the training of CNNs from scratch, integrating cross-entropy loss with mean squared error loss to improve classification performance. Such studies have shown significant advancements in multi-class accuracy and modeling capabilities, where models not only predict the outcomes accurately but also provide insights into the reasoning behind their predictions using attention maps [17].

Recent developments in the field have introduced ordinal loss as a novel approach for KL grade prediction, replacing traditional cross-entropy loss. This has enabled the reformation of the KL grade prediction problem as an ordinal regression task, a methodological advancement that optimizes neural network performance by accounting for the ordered nature of KL grades [18]. This approach has proven effective in improving the generalization and predictive capabilities of models, further validated by the successful implementation of attention mechanisms within CNN architectures designed for KL grade prediction [19]. Moreover, various CNN architectures, when used in ensemble models, have demonstrated enhanced predictive performance. Model ensembles benefit from the strengths of individual models, resulting in robust systems capable of handling the complexities inherent in medical imaging tasks such as KL grade prediction [20].

Previous studies using deep learning for knee OA diagnosis based on the Kellgren-Lawrence (KL) grading system [21-26] are presented in Table 1. These studies demonstrate the positive impact of deep learning on knee OA diagnosis. In particular, the introduction of advanced neural network architectures like ResNeXt holds the potential to offer improved performance. Against this backdrop, this study aims to develop a KL grade prediction model using ResNeXt and to perform a performance comparison with existing ResNet models. This research is expected to make a significant contribution to the automation and accuracy of knee OA diagnosis.

Table 1. Results of previous studies using deep learning to diagnose knee OA based on the KL grading system

Study	Methodology	Key Findings	Challenges/Limitations
Tiulpin & Saarakkala [21]	CNNs	High accuracy in binary classification (OA vs. non-OA); difficulty in intermediate grades (KL2-3).	Limited interpretability, class imbalance
Nehrer et al. [22]	Transfer Learning	Improved accuracy and efficiency; effective for smaller datasets.	Dependency on pre-trained models
Ahmed & Mstafa [23]	Multi-Task Learning	Enhanced model robustness; simultaneous feature extraction.	Increased model complexity
Leung et al. [24]	Ensemble Methods	Higher accuracy, especially for higher KL grades.	Increased computational cost
Binvignat et al. [25]	Explainable AI	Enhanced model transparency; improved clinician trust.	Complexity in integrating interpretability
Nguyen et al. [26]	Semi-Supervised Learning	Comparable performance with less labeled data; cost-effective.	Potential limitations in generalizability

3. Methods

3.1. Dataset

In this study, we utilized the Knee X-ray dataset from Kaggle (https://www.kaggle.com/datasets/ahmedadel002581/knee-x-ray-5-classes) for the diagnosis and KL grade classification of knee osteoarthritis. This dataset contains various images for knee joint detection and KL grade classification, categorized into a total of 5 KL grades

(Grade 0: Normal, Grade 4: Severe) as shown in Fig 1. The total number of images is 7,020. The dataset was divided into three distinct subsets to facilitate model training and evaluation. The training set comprised 70% of the total data, equating to 5,740 images. The validation set accounted for 20% of the data, consisting of 1,404 images. Lastly, the test set included 10% of the dataset, totaling 702 images.

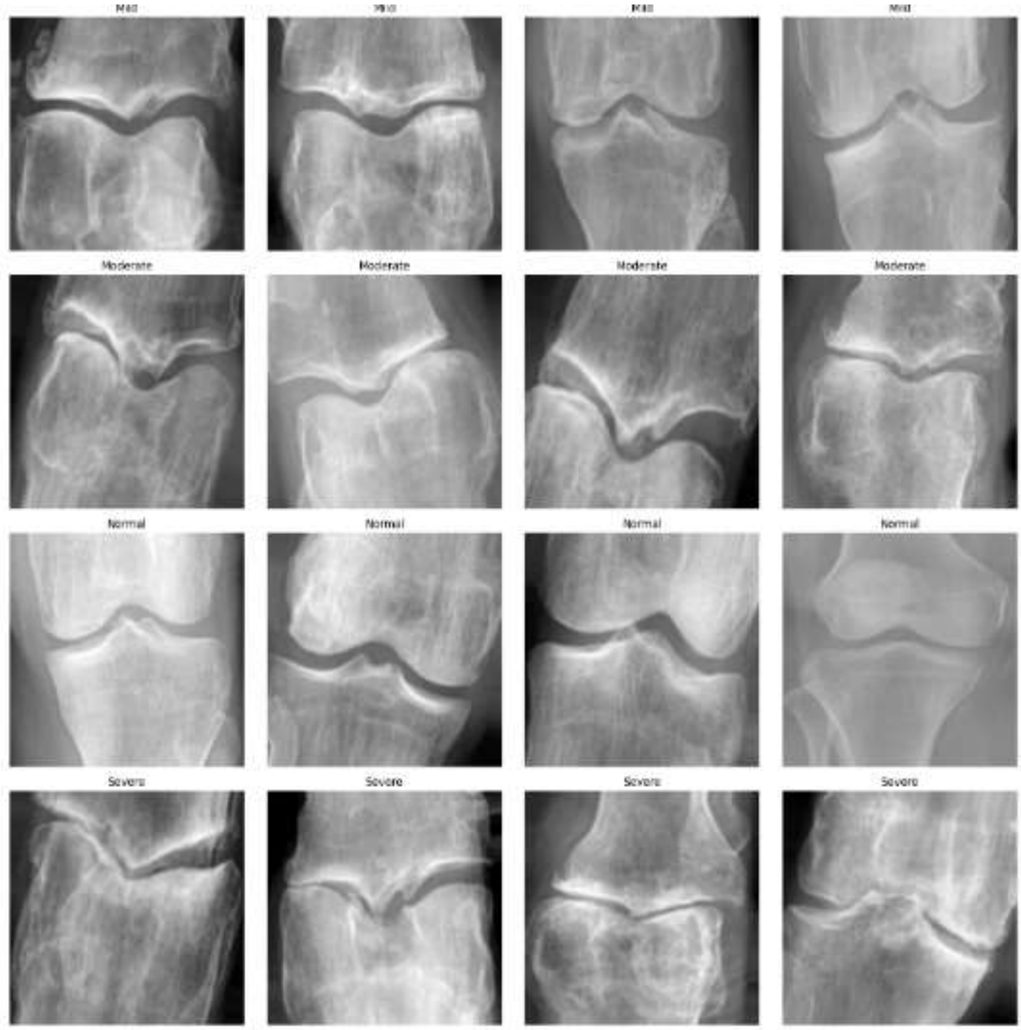

Figure 1. Examples of X-ray images by KL grade included in the dataset

Table 2. KL Grade Composition

Grade	Description	Train	Validation	Test
0	Normal	820	235	118
1	Doubtful: Possible joint space narrowing and osteophyte formation	820	235	117
2	Mild: Definite osteophytes and possible joint space narrowing	820	235	117
3	Moderate: Multiple osteophytes, definite joint space narrowing, some sclerosis	820	235	117
4	Severe: Large osteophytes, marked joint space narrowing, severe sclerosis	820	235	117

3.2. Data Preprocessing

To effectively analyze knee X-ray images, we implemented the following data preprocessing and augmentation techniques.

- Image Resizing:
  - Images were resized to 224x224 pixels. This size is optimized to meet the input requirements of the ResNeXt model, contributing to maximizing the model's training and prediction performance.
- Data Augmentation:
  - Rotation: Applied random rotations within a range of -10 to +10 degrees to generate images at various angles.
  - Horizontal Flip: Applied a horizontal flip with a 50% probability to increase data diversity.
  - Zoom: Applied random zooming within a range of 90% to 110%.
  - Translation: Translated images horizontally and vertically by up to 10% to enhance data diversity.
- Normalization:
  - Normalized pixel values to the [0, 1] range to enhance the stability of model training.

3.3. Handling Class Imbalance

The dataset used in this study has an even distribution across KL grades, thus avoiding class imbalance issues. However, to maximize the model's generalization capability, we employed a stratified fivefold cross-validation technique for dataset partitioning and training.

4. Proposed Architecture and Training Strategy of the ResNeXt Model

In this study, we utilized the ResNeXt-50-32×4d architecture to automatically predict the KL grade of knee osteoarthritis. ResNeXt is an advanced convolutional neural network (CNN) that enhances feature extraction efficiency through parallel paths, demonstrating exceptional performance in various image classification tasks.

4.1. Architecture

The ResNeXt-50-32×4d model fundamentally comprises 50 layers, each extracting features through multiple parallel paths, enhancing the model's representational power without increasing network depth. In this study, we adopted the ResNeXt architecture with the following structural features.

- Cardinality (C): This is a key parameter of ResNeXt, representing the number of parallel paths in each layer. We set the cardinality to 32 to optimize information flow across different paths.

- Bottleneck Structure: To enhance network efficiency, a bottleneck structure was employed using 1x1, 3x3, and 1x1 convolutions in each block.

- Softmax Output Layer: The final output layer uses the softmax activation function to calculate probabilities for each of the 5 KL grades (0 to 4): [P(y_i = k | x_i) = {exp (z_k)}{∑_{j = 1}^{K} exp (z_j)}] where (z_k) is the output of the linear transformation for class (k), and (K) is the total number of classes.

4.2. Training Strategy

To optimize the performance of the ResNeXt model, the following training strategies were employed.

- Transfer Learning: We performed fine-tuning on knee X-ray images based on the pre-trained ResNeXt-50 model on the ImageNet dataset.

- Layer-wise Training: Initially, only the fully connected layer responsible for classification was trained, followed by gradual training of all layers.

- Learning Rate Scheduling: The learning rate was initially set at 0.01, then gradually decreased to 0.001, and finally to 0.0001 to enhance training stability and accuracy.

- Loss Function and Optimization Algorithm: Cross-entropy loss was used, and the Adam optimizer was employed to optimize the model's training: [$L(y, \hat{y}) = -\sum_{i=1}^{N} y_i \log(\hat{y_i})$] where (y_i) is the true class label and ($\hat{y_i}$) is the predicted probability.

4.3. Model Comparison and Performance Evaluation

The performance of the proposed ResNeXt model was evaluated through comparison with ResNet-18, ResNet-34, and ResNet-50 models. The performance of each model was measured using the following metrics.

- Accuracy: The ratio of correctly classified data to the total data, serving as an indicator of the model's overall performance.

- Precision: The ratio of true positive predictions to all positive predictions, assessing the accuracy of the model's positive predictions.

- Recall: The ratio of correctly predicted positive observations to all actual positives, evaluating the model's detection capability.

- F1-score: The harmonic mean of precision and recall, evaluating model performance while considering data imbalance: [$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$]

- AUC (Area Under the Curve): The area under the ROC (Receiver Operating Characteristic) curve, assessing model performance across various classification thresholds.

5. Experimental Results and Performance Analysis

In this study, experiments were conducted using the ResNeXt-50-32×4d model to classify the KL grades of knee osteoarthritis. These experiments were based on the Knee X-ray dataset from Kaggle, and the performance of each model was evaluated using various metrics, including accuracy, precision, recall, F1-score, and AUC. The results of the experiments are as follows.

5.1. Experimental Setup and Data

The dataset was divided into training, validation, and test sets, with each set evenly distributed across KL grades. Each model was trained with optimized image sizes to achieve optimal performance.

5.2. Performance Analysis Results

The performance of each model is summarized in Table 3 and Figure 2. The ResNeXt-50 model achieved the highest accuracy of 76.93%, outperforming other ResNet models. In terms of precision, ResNeXt-50 also led with a score of 0.7876, indicating superior accuracy in positive predictions. The recall was measured at 0.7525, reflecting the model's excellent detection capability. The F1-score, shown in Figure 2, was the highest at 0.7665, serving as a metric for evaluating overall model performance while considering data imbalance.

Table 3. Model Performance: Accuracy, Precision, Recall

Model	Accuracy	Precision	Recall
ResNeXt-50	76.93%	0.7876	0.7525
ResNet-18	74.70%	0.7641	0.7463
ResNet-34	75.41%	0.7557	0.7351
ResNet-50	76.33%	0.7810	0.7513

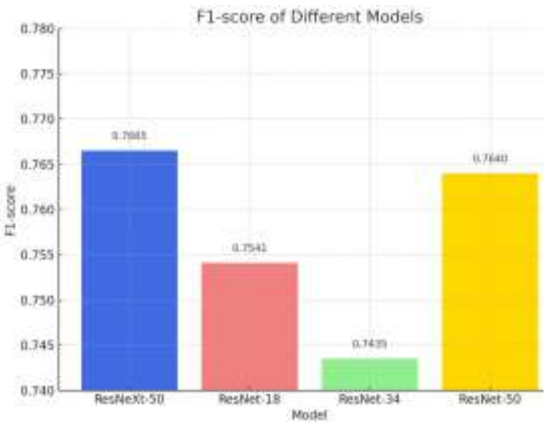

Figure 2. Comparison of F1-scores of Each Model

Figure 3 compares the AUC (Area Under the Curve) values of ResNeXt-50 and various ResNet models using ROC (Receiver Operating Characteristic) curves. The ROC curve serves as a tool for visually evaluating the performance of classification models across different thresholds, with larger AUC values indicating superior overall classification ability.

The ResNeXt-50 model recorded an AUC value of 0.85, demonstrating outstanding classification performance that surpassed other ResNet models. This indicates that ResNeXt-50 has the capability to accurately distinguish between various KL grades. The ResNet-50 model showed close performance with an AUC value of 0.84, whereas ResNet-18 and ResNet-34 recorded relatively lower AUC values of 0.81 and 0.82, respectively. These results suggest that the ResNeXt-50 model exhibits excellent performance in classifying KL grades of knee osteoarthritis, highlighting its potential as a reliable tool for clinical applications.

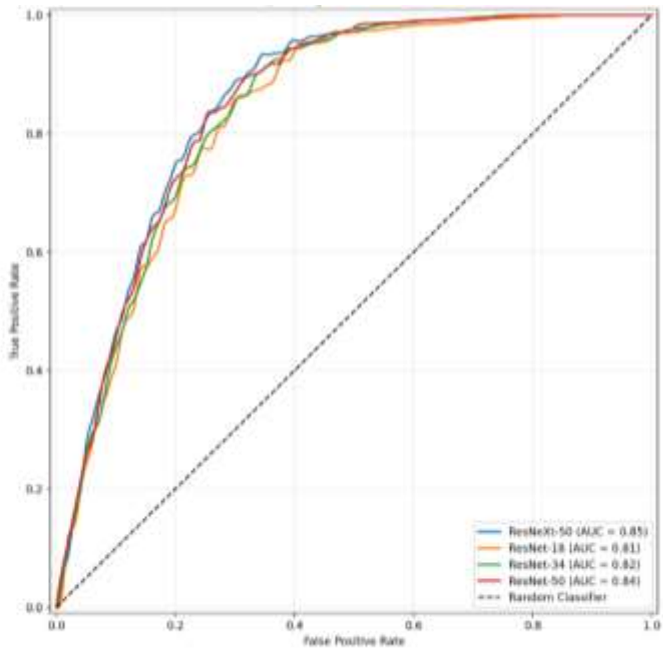

Figure 3. ROC Curves of Each Model's Prediction Performance

5.3. Model Loss

Figure 4 presents a comparison of the training and validation loss values of the ResNeXt-50 and various ResNet models. Loss values are crucial indicators for evaluating the stability and efficiency of model training, with lower loss values indicating better model fit to the given data.

The ResNeXt-50 model recorded a training loss of 0.35 and a validation loss of 0.42, exhibiting lower loss values than other models. This suggests that ResNeXt-50 effectively learned from both training and validation data, demonstrating superior generalization performance. The ResNet-50 model showed the second-lowest loss values, with a training loss of 0.37 and a validation loss of 0.43. In contrast, ResNet-18 and ResNet-34 recorded relatively higher loss values, with training losses of 0.40 and 0.38 and validation losses of 0.46 and 0.44, respectively. These results indicate that the ResNeXt-50 model demonstrates excellent learning efficiency in classifying KL grades of knee osteoarthritis, providing stable performance across diverse datasets.

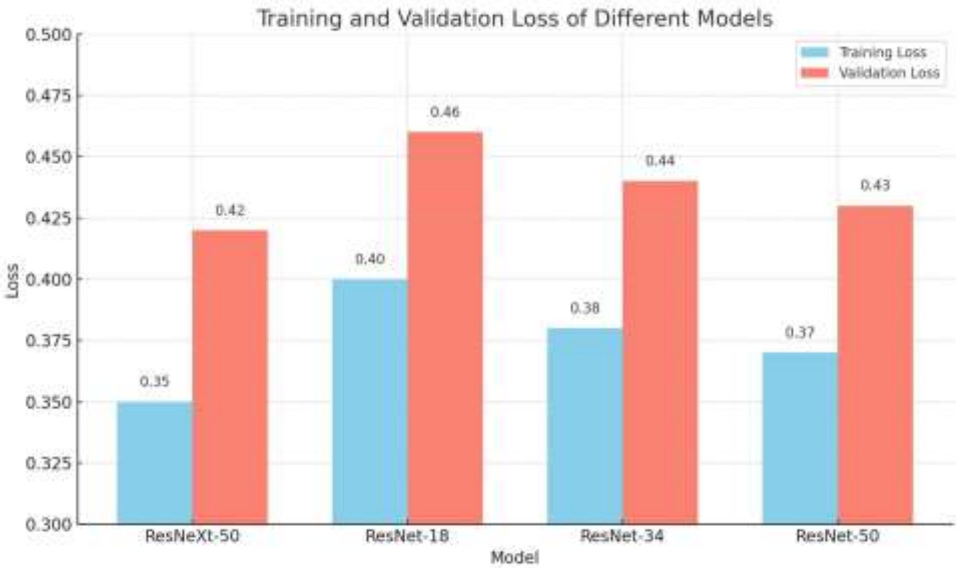

Figure 4. Training and Validation Loss of Each Model

6. Discussion

This study explored the potential of deep learning in classifying the KL grades of knee osteoarthritis using the ResNeXt-50-32×4d model. The ResNeXt model demonstrated overall superior performance compared to traditional ResNet models, particularly excelling in various evaluation metrics such as accuracy, precision, recall, F1-score, and AUC. These results suggest that the parallel path architecture and efficient training strategy of ResNeXt significantly contributed to its ability to effectively learn the complex patterns in knee X-ray images.

Deep learning models such as ResNeXt-50, known for their enhanced feature extraction capabilities and flexible architecture, have shown substantial promise in the automated diagnosis of knee osteoarthritis. The use of such sophisticated models enables unprecedented levels of accuracy and stability in predicting KL grades, which are critical for assessing the severity of osteoarthritis. These models utilize residual connections and an array of transformation modules, allowing them to manage large-scale images with improved computational efficiency, which is particularly beneficial in processing high-resolution medical images [27].

The potential of the ResNeXt-50 model extends beyond mere classification; it serves as a powerful tool that medical professionals can leverage to augment their diagnostic procedures. By integrating such technologies, healthcare facilities can significantly enhance the speed and objectivity of diagnoses, thus facilitating a quicker decision-making process in clinical settings. This becomes particularly crucial when dealing with large patient datasets, where manual analysis may be impracticably slow and prone to human error [28].

Moreover, the adoption of automated systems like the ResNeXt-50 model contributes to the democratization of healthcare. In resource-limited settings or where specialized expertise is scarce, these systems enable a level of diagnostic reliability that can be equated to centers with abundant resources. This can lead to improved patient outcomes by enabling early intervention and tailored treatment plans based on precise and consistent diagnostic data [29].Future research will focus on evaluating the generalization performance of the model using diverse datasets and exploring the applicability of the ResNeXt model in other medical image analysis domains. Additionally, to enhance model interpretability, visualization techniques such as Grad-CAM could be employed to analyze the image regions that the model focuses on when predicting specific grades.

The limitations of this study are as follows. First, the dataset used was limited to a specific Kaggle dataset, necessitating evaluation of the model's generalization performance on other datasets. Second, the impact of image resolution and preprocessing methods on model performance was not thoroughly explored. Third, the application and analysis of visualization techniques to improve model interpretability were limited. Addressing these limitations, future research will aim to expand the applicability of the model through performance validation under more comprehensive datasets and various conditions.

7. Conclusion

This study demonstrated that the ResNeXt-50-32×4d model can outperform existing ResNet models in classifying KL grades of knee osteoarthritis. The parallel path architecture and efficient training strategy of ResNeXt show promise in effectively processing complex medical imaging data and suggest its applicability as an automated diagnostic system in clinical environments. Future studies should evaluate the model's generalization performance across various datasets and conditions and further explore methods to enhance interpretability.

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

Availability of data and materials. Detailed information can be found at: https://www.kaggle.com/datasets/ahmedadel002581/knee-x-ray-5-classes.

Author's Contribution. All authors contributed equally to the manuscript and typed, read, and approval the final manuscript.

Acknowledgement. We are thankful to the editors and the anonymous reviewers for many valuable suggestions to improve this paper.

Funding: This research Supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF- RS-2023-00237287, NRF-2021S1A5A8062526), local government-university cooperation-based regional innovation projects (2021RIS-003).

Ethical approval. This study was exempt from research ethics approval because it utilized secondary data from public sources.

References

[1] V. Gouttebarge, H. Inklaar, F. Backx, & G. Kerkhoffs, Prevalence of osteoarthritis in former elite athletes: a systematic overview of the recent literature, Rheumatology International, 35 (2015), 405-418. https://dx.doi.org/10.1007/s00296-014-3110-3

[2] J. Y. Reginster, J. P. Pelletier, J. Martel-Pelletier, Y. Henrotin, L. Crasborn, J. Pouchot, ... & F. Guillemin, Impact of osteoarthritis on quality of life, Osteoarthritis: Clinical and Experimental Aspects, (1999), 331-355.

[3] K. Sinusas, Osteoarthritis: diagnosis and treatment, American Family Physician, 85 (2012), 49-56. https://www.aafp.org/afp/2012/0101/p49.html

[4] M. D. Kohn, A. A. Sassoon, & N. D. Fernando, Classifications in brief: Kellgren-Lawrence classification of osteoarthritis, Clinical Orthopaedics and Related Research®, 474 (2016), 1886-1893. https://dx.doi.org/10.1007/s11999-016-4732-4

[5] H. J. Park, S. S. Kim, S. Y. Lee, N. H. Park, J. Y. Park, Y. J. Choi, & H. J. Jeon, A practical MRI grading system for osteoarthritis of the knee: association with Kellgren–Lawrence radiographic scores, European Journal of Radiology, 82 (2013), 112-117. https://dx.doi.org/10.1016/j.ejrad.2012.09.020

[6] H. Byeon, YOLO v10-Based Brain Tumor Detection: An Innovative Approach in CT Imaging, Nanotechnology Perceptions, (2024), 113-125.

[7] H. Byeon, Leveraging AdaNet for In-Depth Analysis of Hypertension Medication Adherence in 1,512 Middle-Aged Male Korean Workers, Nanotechnology Perceptions, (2024), 393-404.

[8] H. Byeon, M. Shabaz, J. V. N. Ramesh, A. K. Dutta, R. Vijay, M. Soni, ... & P. P. Singh, Feature fusion-based food protein subcellular prediction for drug composition, Food Chemistry, 454 (2024), 139747. https://dx.doi.org/10.1016/j.foodchem.2023.139747

[9] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, & J. Liang, Convolutional neural networks for medical image analysis: Full training or fine tuning?, IEEE Transactions on Medical Imaging, 35 (2016), 1299-1312. https://dx.doi.org/10.1109/TMI.2016.2535302

[10] Y. Luo, Y. Xu, C. Wang, Q. Li, C. Fu, & H. Jiang, ResNeXt-CC: a novel network based on cross-layer deep-feature fusion for white blood cell classification, Scientific Reports, 14 (2024), 18439. https://dx.doi.org/10.1038/s41598-024-18439-7

[11] S. Olsson, E. Akbarian, A. Lind, A. Razavian, & M. Gordon, Automating classification of osteoarthritis according to Kellgren-Lawrence in the knee using deep learning in an unfiltered adult population, BMC Musculoskeletal Disorders, (2021). https://dx.doi.org/10.1186/s12891-021-04722-7

[12] J. S. Yoon, C.-J. Yon, D. Lee, J. J. Lee, C. H. Kang, S.-B. Kang, N.-K. Lee, & C. B. Chang, Assessment of a novel deep learning-based software developed for automatic feature extraction and grading of radiographic knee osteoarthritis, BMC Musculoskeletal Disorders, (2023). https://dx.doi.org/10.1186/s12891-023-06951-4

[13] J.-Z. Yang, Q. Ji, M. Ni, G. Zhang, & Y. Wang, Automatic assessment of knee osteoarthritis severity in portable devices based on deep learning, Journal of Orthopaedic Surgery and Research, (2022). https://dx.doi.org/10.1186/s13018-022-03429-2

[14] T. Cui, R. Liu, Y. Jing, J. Fu, & J. Chen, Development of machine learning models aiming at knee osteoarthritis diagnosing: an MRI radiomics analysis, Journal of Orthopaedic Surgery and Research, (2023). https://dx.doi.org/10.1186/s13018-023-03837-y

[15] M. R. Ibraheem, S. Almuayqil, A. A. Abd El-Aziz, M. A. Tawfeek, & F. M. Talaat, Diagnosis of patellofemoral osteoarthritis using enhanced sequential deep learning techniques, Egyptian Informatics Journal, (2023). https://dx.doi.org/10.1016/j.eij.2023.100

[16] Z. M. Wani, & D. S. Saini, Deep Neural Network-based Knee Osteoarthritis Grading Using X-Rays, International Journal for Research in Applied Science and Engineering Technology, (2022). https://dx.doi.org/10.22214/ijraset.2022.41757

[17] B. Zhang, J. Tan, K. Cho, G. Chang, & C. Deniz, Attention-based CNN for KL Grade Classification: Data from the Osteoarthritis Initiative, International Symposium on Biomedical Imaging (ISBI), (2020). https://dx.doi.org/10.1109/ISBI45749.2020.9098456

[18] J. Serrat, & I. Ruiz, Rank-based ordinal classification, International Conference on Pattern Recognition (ICPR), (2021). https://dx.doi.org/10.1109/ICPR48806.2021.9412846

[19] R. Jain, P. K. Sharma, S. Gaj, A. Sur, & P. Ghosh, Knee Osteoarthritis Severity Prediction using an Attentive Multi-Scale Deep Convolutional Neural Network, Multimedia Tools and Applications, (2021). https://dx.doi.org/10.1007/s11042-023-15484-w

[20] E. Ramasso, R. Teloli, & R. Marcel, On the Condition Monitoring of Bolted Joints through Acoustic Emission and Deep Transfer Learning: Generalization, Ordinal Loss and Super-Convergence, arXiv, (2024). https://dx.doi.org/10.48550/arXiv.2405

[21] A. Tiulpin, & S. Saarakkala, Automatic grading of individual knee osteoarthritis features in plain radiographs using deep convolutional neural networks, Diagnostics, 10 (2020), 932. https://dx.doi.org/10.3390/diagnostics10110932

[22] S. Nehrer, R. Ljuhar, P. Steindl, R. Simon, D. Maurer, D. Ljuhar, Z. Bertalan, H. Dimai, C. Goetz, & T. Paixao, Automated Knee Osteoarthritis Assessment Increases Physicians' Agreement Rate and Accuracy: Data from the Osteoarthritis Initiative, Cartilage, 13 (2019), 957S-965S. https://dx.doi.org/10.1177/1947603519871675

[23] S. M. Ahmed, & R. J. Mstafa, Identifying severity grading of knee osteoarthritis from x-ray images using an efficient mixture of deep learning and machine learning models, Diagnostics, 12 (2022), 2939. https://dx.doi.org/10.3390/diagnostics12122939

[24] K. Leung, B. Zhang, J. Tan, Y. Shen, K. J. Geras, J. S. Babb, ... & 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, 296 (2020), 584-593. https://dx.doi.org/10.1148/radiol.2020201113

[25] M. Binvignat, V. Pedoia, A. J. Butte, K. Louati, D. Klatzmann, F. Berenbaum, ... & J. Sellam, Use of machine learning in osteoarthritis research: a systematic literature review, RMD Open, 8 (2022), e001998. https://dx.doi.org/10.1136/rmdopen-2022-001998

[26] H. H. Nguyen, S. Saarakkala, M. B. Blaschko, & A. Tiulpin, Semixup: In- and Out-of-Manifold Regularization for Deep Semi-Supervised Knee Osteoarthritis Severity Grading From Plain Radiographs, IEEE Transactions on Medical Imaging, 39 (2020), 4346-4356. https://dx.doi.org/10.1109/TMI.2020.3025310

[27] K. He, X. Zhang, S. Ren, & J. Sun, Deep Residual Learning for Image Recognition, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), (2016), 770-778. https://dx.doi.org/10.1109/CVPR.2016.90

[28] A. Krizhevsky, I. Sutskever, & G. Hinton, ImageNet Classification with Deep Convolutional Neural Networks, Advances in Neural Information Processing Systems (NIPS), (2012), 1097-1105.

[29] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, ... & C. I. Sánchez, A survey on deep learning in medical image analysis, Medical Image Analysis, 42 (2017), 60-88. https://dx.doi.org/10.1016/j.media.2017.07.005