Predicting Knee Osteoarthritis Grades using Deep Learning - A Extensive Examination

V. Vijaya Kishore[1*], G. Hema padmini[2], I. Sudharshan[3], K. Tejaswini[4], M. Roopa[5,] Rambabu Inaganti[6]

[1] Prof, Dept of ECE, Mohan Babu University (Erstwhile Sree Vidyanikethan Engineering College), Tirupati, India
*kishiee@rediffmail.com
[2,3,4,5] Dept of ECE, Erstwhile Sree Vidyanikethan Engineering College, Tirupati, India
ganjihemu102@gmail.com,sudharshani369@gmail.com
tejareddykodathala@gmail.com,maruveniroopa@gmail.com
[6]Smith & Nephew, Pittsburgh, PA 15108. , USA
Rambabu8@gmail.com

Abstract. A major global health concern is knee osteoarthritis, which is frequently identified by conventional radiographic grading schemes like the Kellgren-Lawrence scale. Reliance on X-ray pictures, however, may cause a delayed diagnosis. Convolutional neural networks (CNNs), in particular, are deep learning techniques that have been the subject of recent research aimed at improving diagnostic efficiency and accuracy. Eight CNN-based adaptive neural network models were examined for the diagnosis of knee osteoarthritis. These models were trained and verified using a large dataset of knee X-rays, and then their ability to classify the severity of osteoarthritis in the knee was thoroughly examined. According to our findings, the best-performing model outperformed the others, with an astounding accuracy of 98.73%. This study highlights how deep learning models, in particular CNNs, can increase knee osteoarthritis diagnosis accuracy and speed.

Keywords: Deep Learning, X-rays, Adaptive neural network, Gradient descent optimization, Leaky ReLU

1 Introduction

The Osteoarthritis (OA) of the knee, a prevalent musculoskeletal disease, necessitates accurate diagnostic and therapeutic approaches due to its significant impact on individuals and healthcare resources. Recent developments in machine learning (ML) and artificial intelligence (AI), particularly deep learning (DL), present promising paths for accurate disease evaluation, while older methods rely on subjective assessments and qualitative tests. Based on radiographic imaging, the Kellgren-Lawrence (KL) grading system has long been used to determine the severity of osteoarthritis[1] [2]. However, its reliance on subjective interpretation and qualitative assessments poses limitations,

K. R. Madhavi et al. (eds.), *Proceedings of the International Conference on Computational Innovations and Emerging Trends (ICCIET 2024)*, Advances in Computer Science Research 112,
https://doi.org/10.2991/978-94-6463-471-6_71

leading to diagnostic inaccuracies and suboptimal treatment decisions. To address these challenges, our research aims to explore DL strategies for accurately classifying OA severity using knee joint radiographs. We evaluate various DL architectures, including VGG16, EfficientNetB7, DenseNet169, MobileNet, VGG19, NASNetlarge, Xception, and InceptionV3 leveraging annotated datasets and rigorous model training. Our goal is to develop a robust predictive model capable of objectively categorizing different levels of knee OA severity. By harnessing DL-based diagnostic tools, clinicians can gain valuable insights into disease progression, treatment responses, and patient outcomes. Early detection facilitated by these tools enables personalized interventions and proactive management strategies, ultimately improving patient care and optimizing resource allocation. Although the KL grading system primarily focuses on structural changes observed in X-rays, it may not always correlate perfectly with clinical symptoms. Thus, healthcare professionals must consider additional factors, such as patient symptoms and functional status, when managing knee OA comprehensively.

The Kellgren-Lawrence (KL) grading system, developed in 1957, is widely used to classify osteoarthritis (OA) severity in the knee based on X-ray imaging. It ranges from Grade 0 (no signs of OA) to Grade 4 (severe OA), assessing structural changes like osteophytes and joint space narrowing. However, it may not always correlate perfectly with patients' clinical symptoms [9]. Healthcare professionals primarily use this system for knee OA but may adapt it for other affected joints. While valuable for monitoring disease progression and guiding treatment decisions, comprehensive management requires considering additional factors such as patient symptoms and functional status.

our research represents a significant contribution to the intersection of DL-driven healthcare and musculoskeletal medicine. By working together, we hope to close the gap that exists between contemporary technology and clinical practice, improving knee OA diagnosis, treatment, and prognosis in the process.

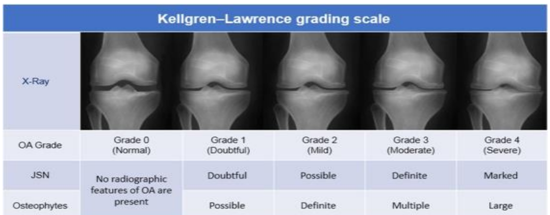


Fig. 1. Kellgren-Lawrence Grading System.

2 Literature Review

The diagnosis and treatment of knee osteoarthritis have been utilizing deep learning (DL) approaches more and more, since they demonstrate great promise for improving early detection, intervention, and treatment planning. With the Kellgren-Lawrence grading system for early detection and intervention, DL models have shown a revolutionary impact on knee OA diagnosis and management. The studies extensive datasets of clinical and knee X-ray images demonstrate the robustness and generalizability of the established models, which mark advancements in improving healthcare outcomes. Beyond the limitations of typical machine learning techniques, a thorough assessment of DL models for KOA severity prediction has been carried out. These papers note important challenges, such as low dataset diversity and the hazards of overfitting, despite encouraging results, highlighting the necessity for ongoing research efforts to improve DL models for KOA prediction.

DL models, like as Convolutional Neural Networks (also known as CNNs) and recurrent neural network models (RNNs), have been used in KOA for detection and classification in addition to severity prediction. While highlighting the impressive sensitivity and accuracy shown in earlier research, these also highlight ongoing issues including interpretability and the need for big, diverse datasets to support the resilience and generalizability of the model. A novel federated learning approach has been introduced to address class imbalance, thereby enhancing the accuracy and reliability of multiclass classification tasks. This innovative methodology signifies a step forward in refining DL models for KOA diagnosis and management by tackling critical gaps in existing approaches. Further innovation is seen in a knee cartilage segmentation method utilizing Generative Adversarial Networks (GANs) with transfer learning. Addressing the challenges associated with heterogeneous clinical MRI data, this approach presents a significant breakthrough in improving segmentation accuracy, thereby advancing diagnosis and treatment planning for knee OA. The role of advanced MRI techniques and artificial intelligence in OA assessment has been explored, advocating for improved accuracy and efficiency in diagnosis and monitoring. This discussion lays the groundwork for future research aiming to harness AI capabilities for enhanced clinical decision-making in knee OA management. A comprehensive overview of image segmentation techniques for knee OA research highlights the benefits and limitations of both classical and DL-based methods, serving as a valuable resource for researchers. This holistic perspective facilitates informed decision-making in method selection based on specific research requirements, contributing to the interdisciplinary collaboration between medical research and artificial intelligence for transformative advancements in knee OA diagnosis and management. Variations in accuracy and loss measures are observed during training and validation when comparing deep learning models for knee osteoarthritis (KOA)[12]. For each classification task, a large number of photos must be manually classified by an expert to aid in model training. This study lays the groundwork for an orthopedic department to transition to a complete automated suggestion system. Such a system would provide OA surgeons with vital inputs that would enable them to prioritize patient care based on the model's output, plan patient therapy, and evaluate surgical options. All things considered, these results demonstrate how revolutionary KOA diagnosis and treatment decision-making procedures may be using MobileNet and related models[2]. A number of significant changes can be made to improve the precision of our knee osteoarthritis (KOA) categorization method. Batch normalization, first of all, helps stabilize and expedite the training process by normalizing the input data layer by modifying and scaling the activations. Gradient descent optimization makes it possible to suit the KOA dataset's unique properties more precisely. Furthermore, by including more layers in the neural network architecture, it may be possible to identify more intricate patterns and features in the photos, which would enhance the model's capacity to correctly categorize KOA severity levels. Our goal is to greatly improve our classification model's efficacy and accuracy in identifying KOA by extending and modifying its architecture. This will allow for better treatment choices improved patient outcomes.

3 Methodology

A neural network, employing the KL method, analyzes radiography data to diagnose knee osteoarthritis (KOA) and assess its severity. It classifies OA grades based on learned patterns from radiographic images.

3.1 Input Data

A dataset of 1650 photographs was split into the following grades for the purpose of knee osteoarthritis (OA) detection: 514 level 0, 477 level 1, 232 level 2, 221 level 3, and 206 level 4. Subsets of the dataset were separated out to be used for train (70%), test (10%), and validating (20%) in order to optimize the parameters of the deep learning model, avoid overfitting, and adjust hyperparameters. By preventing hyperparameter customization based on testing data, this segregation guarantees optimal performance on unseen data.

3.2 Pre-Processing

Convolutional neural networks, also referred to as CNNs, are widely utilized in orthopedic osteoarthritis analysis using deep learning to extract information from medical images such as X-rays. The process involves training the CNN with labeled data to learn osteoarthritic patterns, validating its performance on a separate dataset, and finally testing its effectiveness on unseen data. Regularization techniques and hyperparameter tuning are essential for optimizing the model's performance.

3.3 Convolutional Layers

Convolutional Neural Networks (CNNs) evaluate images using filters, small matrices with weights, through convolution. This process filters the input image, calculates dot products in overlapping parts, and generates a feature map highlighting identified features, allowing the network to identify patterns, edges, and textures. CNNs often include additional layers like pooling and fully connected layers to further improve and interpret features beyond initial extraction. Through a hierarchical approach, the network gradually comprehends complex representations, enhancing its ability to classify and comprehend images comprehensively without duplication.

3.4 Activation Layer

Leaky ReLU is an activation function employed in neural networks to introduce a slight gradient for negative inputs, unlike traditional ReLU which sets negative values to zero. This characteristic mitigates "dying neurons" and the vanishing gradient problem in deep networks, enhancing learning stability. Leaky ReLU efficiently addresses ReLU's dead neuron limitation while maintaining computational efficiency, contributing to its widespread adoption in deep learning. Its inclusion in knee osteoarthritis (KOA) analysis with Kellgren-Lawrence (KL) grading allows for a modest, positive gradient for negative inputs, preventing neurons from becoming dormant. This adjustment improves network stability during training, crucial for accurately estimating KL grades in osteoarthritis cases. By incorporating Leaky ReLU, neural networks effectively handle variations in KL grading data, leading to improved performance and more dependable predictions of disease severity.

3.5 Pooling Layer

In CNNs, the pooling layer typically follows the convolutional layer, down sampling by selecting maximum or average values from specified areas in the feature map. This reduces spatial dimensions while maintaining original content, enhancing model efficacy and reducing overfitting risk.

3.6 Fully Connected Layer

The fully connected layer of a neural network made up of convolutional neurons (CNN) is essential for deciphering characteristics from earlier layers and, eventually, for classifying images. By flattening the output into a one-dimensional vector and passing it through multiple layers with weighted sums and non-linear activation functions like Leaky ReLU, the fully connected layer generates a probability distribution across classes, identifying the predicted class based on the highest probability. This layer is integral to the network's pattern recognition and classification capabilities. Additionally, CNNs excel in image recognition due to their specialized convolution operation, allowing them to scan images continuously to identify specific features.

Our methodology focuses on enhancing the accuracy of convolutional neural networks (CNNs) through adaptive architecture design and Leaky Rectified Linear Unit (ReLU) activation functions. We begin by selecting diverse and relevant datasets, preprocessing them to ensure uniformity and augmenting them to increase variability. Our baseline CNN architecture comprises alternating convolutional and pooling layers with Leaky ReLU activation functions to introduce non-linearity and dropout layers to mitigate overfitting. We progressively increase model depth by adding convolutional blocks, optimizing depth through iterative experimentation while applying regularization techniques data augmentation. Hyperparameters are fine-tuned, and optimization algorithms such as Adam are employed. We explore advanced architectural enhancements like residual connections and attention mechanisms. Performance evaluation is conducted using standard metrics, and statistical analysis validates performance improvements. The trained model is deployed, with ongoing monitoring and periodic retraining to maintain accuracy and adaptability.

4 Results and Discussions

A comparison was conducted among various deep learning models utilized for knee osteoarthritis (KOA) detection, considering both accuracy and loss metrics during training and validation. Results showed a wide range of performance, with different detection rates across these models. NASNetLarge and InceptionV3 emerged as the most accurate models, achieving remarkable accuracy of 98.73%. Notably, we have gradient decent optimization, making our project unique, and improving the accuracy of all models. EfficientNetB7 demonstrated the highest loss. These findings demonstrate the effectiveness of InceptionV3 in creating an image classification system for osteoarthritis in the knee based on KL-grade severity.

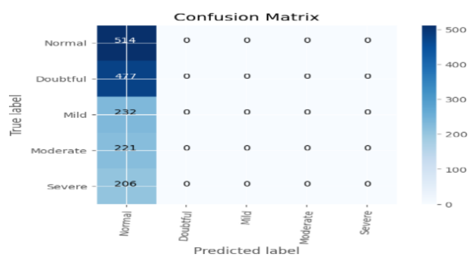
Table 1. Comparison of Eight Distinct Learning Models Validation, Training Accuracy and Loss.

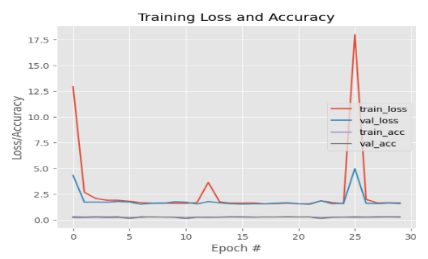
Model Name	Val-Accuracy	Train-Accuracy	Val-Loss	Train-loss
DenseNet169	0.9697	0.9436	0.1041	0.2308
EfficientNetB7	0.3115	0.2703	1.985	1.6602
VGG16	0.9674	0.9248	0.994	0.2082
MobileNet	0.9800	0.9945	0.1113	0.0120
NASNetLarge	0.9873	0.9697	0.0699	0.1031
VGG19	0.9606	0.9248	0.1436	0.1945
InceptionV3	0.9873	0.9824	0.0612	0.0633
Xception	0.9673	0.9042	0.1131	0.2874

First X-ray analysis and decision-making phases can be automated by choosing the best models for classifying OA medical pictures. To support model training, a large number of photos must be manually classified by knowledgeable people. This work provides significant information for OA surgeons to evaluate surgical treatments, patient management, and prioritise care for patients based on model outputs. It also lays the foundation for the implementation of a throughout its entirety device recommendation system in orthopedic departments. It's also important to note that the model's training time has been drastically lowered, with each epoch taking only 20 seconds and each step taking 126 milliseconds. Occupational activities have been identified in previous reviews as possible risk factors for osteoarthritis (OA) of the knee; however, few studies have produced quantitative data, probably because the literature covers a wide variety of risk variables and assessments. therapy decisions can be aided by deep learning (DL) algorithms, which can differentiate between radiographic OA with no symptoms and symptomatic OA that needs therapy. With companies developing in models of prediction for hospitalized patients and patient data management, the incorporation of DL into healthcare has greatly enhanced patient care. Appropriate sample size consideration in relation to model predictors is necessary to ensure dependable predictive performance. Machine learning algorithms' robustness is significantly impacted by the amount and quality of the dataset, particularly with regard to image data. With deep learning (DL) emerging and providing customized solutions for healthcare requirements, the future of the field of healthcare seems bright. For first diagnosis, a dependable and accurate device would be especially helpful in places where radiology specialists are not readily available. To enhance image analysis and facilitate continuous network learning, DL systems can incorporate radiologist reports and image data. AI usage in orthopedics is dominated by deep learning techniques combined with clinical pictures. Transfer learning improves the ability to learn from one dataset in order to generate predictions on other datasets.

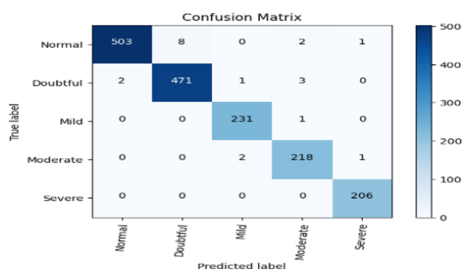
Table 2. Comparison of EfficientNetB7 and NASNetLarge's Confusion Matrix, Training Loss, and Accuracy

EfficientNetB7





NASNetLarge



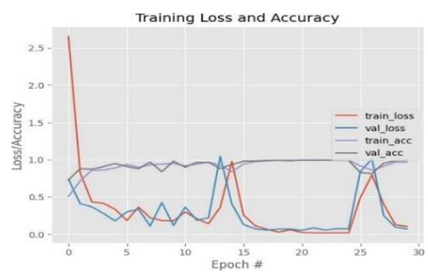
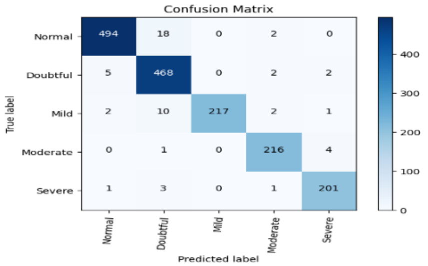
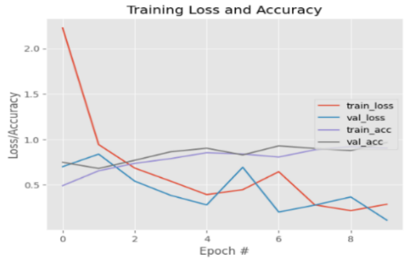


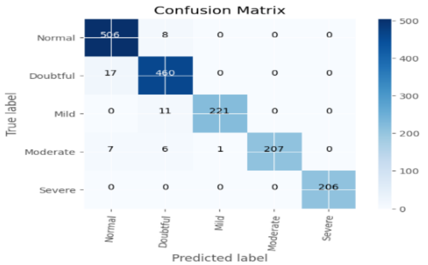
Table 3. Comparison of Xception, DenseNet169, VGG16, and InceptionV3's Accuracy, Confusion Matrix, and Training Loss

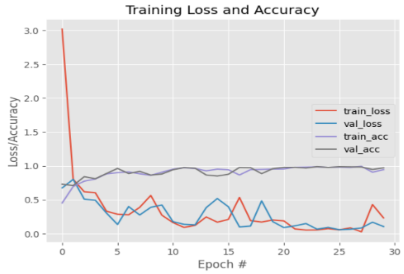
Xception



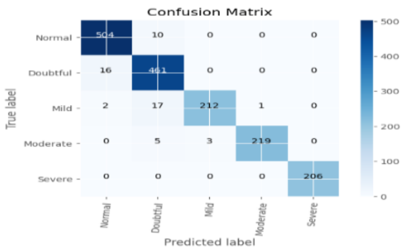


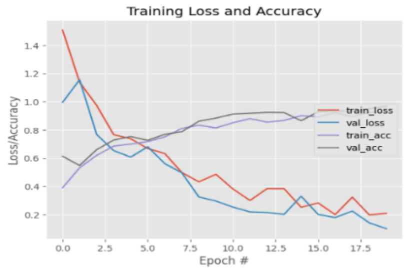
DenseNet169





VGG16





InceptionV3

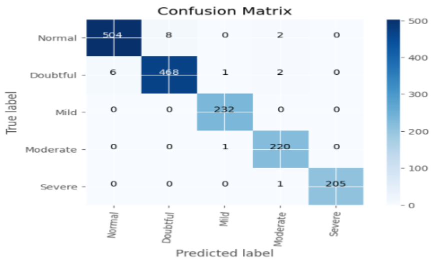
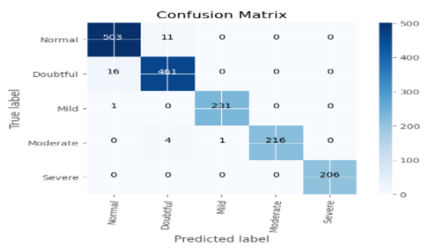
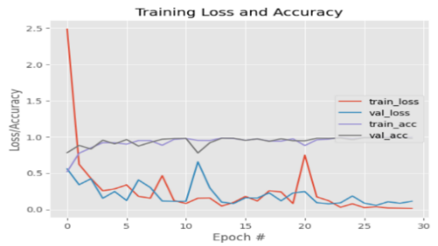


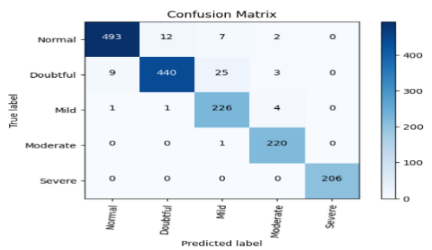
Table 4. Comparison of MobileNet and VGG19 Accuracy, Training Loss, and Confusion Matrix

MobileNet





VGG19



5 Conclusion

Even though our KOA project produced remarkable results with an astounding accuracy rate of 98.73%, further research is still needed to fully understand osteoarthritis (OA) of the knee prediction models. The intricacy of radiographic indicators and risk factors from knee x-rays necessitates continued research to develop and validate prediction models across diverse populations. The integration of such models into everyday clinical practice could greatly enhance clinical decision-making, providing clinicians with valuable tools for early detection and personalized management of KOA. Therefore, it is essential to refine and modify these models and their associated parameters to explore their performance and accuracy levels, particularly concerning the impact of tuning and epochs. Future efforts will focus on advancing these prediction models to ensure their effectiveness across various demographic groups and healthcare settings. By optimizing parameters and refining the models, we aim to empower clinicians with reliable tools for KOA diagnosis and management, ultimately improving patient outcomes and quality of life. This ongoing research reflects our commitment to advancing musculoskeletal medicine and addressing the evolving challenges of KOA diagnosis and management.

6 References

1. V. Vijaya Kishore, et al.: Knee Osteoarthritis Prediction Driven By Deep Learning And The Kellgren-Lawrence Grading. Proceedings On Engineering Sciences, vol. 05, no. 3, pp. 475-484, 2023.
2. Kalpana, V., and Hemanth Kumar G.: Evaluating The Efficacy Of Deep Learning Models For A knee Osteoarthritis Prediction Based On Kellgren-Lawrence Grading System, ePrime-Advances in Electrical a Engineering, Electronics and Energy, 2023.
3. Tariq, Z. Suhail and Z. Nawaz.: Knee Osteoarthritis Detection and Classification Using X-Rays, IEEE Access, vol. 11, pp. 48292-48303, 2023.
4. J. Kaur & C.C., Alvaro & M. Kermani & R. Azarderakhsh, A comprehensive survey on the implementations, attacks, and countermeasures of the current NIST lightweight cryptography standard (2023).
5. S.C. Sangapu, K.S.N. Prasad, R.J. Kannan, et al.:Impact of class imbalance in VeReMi dataset for misbehavior detection in autonomous vehicles, Softa Comput., 2023.
6. C.C., Alvaro & J. Kaur & M.K., Mehran & R. Azarderakhsh, Algorithmic security is insufficient: a comprehensive survey on implementation attacks haunting post quantum security (2023)
7. A. Tiwari, M. Poduval, V. Bagaria.: Evaluation of artificial intelligence models for osteoarthritis of the knee using deep learning algorithms for orthopedic radiographs, World J. Orthop. 13 (6) (2022) 603–614.
8. Hirvasniemi J., et al.:The Knee OsteoArthritis Prediction(KNOAP2020) challenge: An a image analysis a challenge to predict incident symptomatic radiographic knee osteoarthritis from MRI and X-ray images, Osteoarthritis and Cartilage, vol. 24, pp. S10-S11, 2022.
9. Y. Himeur, M. Elnour, F. Fadli, et al., AI-big data analytics for building automation and management systems: a survey, actual challenges and future perspectives, Artif. Intell. Rev. 56 (2023) 4929–5021.
10. M. Yang, C. Colak, K.K. Chundru, et al.: Automated knee cartilage segmentation for heterogeneous clinical a MRI using generative adversarial networks with transfer learning, Quant Imaging Med Surg., vol. 12, no. 5, pp. 2620–2633, 2022.
11. S.S. Abdullah, M.P. Rajasekaran.:Automatic detection and classification of knee osteoarthritis using deep a learning approach, Radiol Med., vol. 127, no. 4, pp. 398–406, 2022.
12. Abd Ghani, M. K., Mohammed, M. A., Arunkumar, N., Mostafa, S. A., Ibrahim, D. A., Abdullah, M. K., ... & Burhanuddin, M. A.:Decision-level fusion scheme for nasopharyngeal carcinoma identification using machine learning techniques, Neural Computing and Applications, vol. 32, pp. 625-638, 2020.
13. Eckstein, F., Wirth, W., & Culvenor, A. G. (2021). Osteoarthritis year in review 2020: imaging. Osteoarthritis and Cartilage, 29(2), 170-179.
14. Christodoulou, E., Moustakidis., S., Papandrianos, N., Tsaopoulos, D., & Papageorgiou, E.: Exploring deep learning capabilities in knee osteoarthritis case study for classification, 2019 10th International Conference on Information, Intelligence, Systems and Applications (IISA), pp. 1-6, 2019.
15. A. Brahim, R. Jennane, R. Riad, T. Janvier, L. Khedher, H. Toumi, and E. Lespessailles.: A decision support tool for early detection of knee osteoarthritis using X-ray imaging and machine learning: Data from the Osteoarthritis Initiative, Computerized Medical Imaging and Graphics, vol. 73, pp. 11-18, Jan. 2019.