



KNEE OSTEOARTHRITIS PREDICTION AND CLASSIFICATION FROM X-RAY IMAGES USING DEEP LEARNING AIML

¹Likhithasree Raparathi, ²Ch. Aravind Kumar(Assistant professor),

³Dr.N.Satyavathi(Associate professor)

^{1,2}CSE, Vaagdevi College of Engineering(Bollikunta, Warangal)

Email id: raparthilikhithasree594@gmail.com, aravindkumar_ch@vaagdevi.edu.in,
headcse@vaagdevi.edu.in

Abstract: Kellgren–Lawrence (KL) grading system is a radiographic-based grading system used to figure out how bad knee osteoarthritis (KOA) is. KOA is a degenerative joint disease that gets worse over time and changes the structure of X-ray pictures. We made a big collection of KOA X-ray pictures by combining four public datasets and adding 110,232 new samples with the help of a deep convolutional generative adversarial network (DCGAN). to support the prompt detection and promote the accuracy of the diagnosis. The images were enhanced using advanced pre-processing methods like adaptive histogram equalization (AHE) and quick non-local means filtering and uniform scaling. The first classification model was developed based on a modified compact convolutional transformer architecture, which is called KOA-CCTNet. A large number of parameter and hyperparameter configurations were explored to make the most out of performance, the ability to handle large volumes of data, and the reduction of training durations. On the one hand, the proposed model was more accurate in diagnosis compared to Swin Transformer, Vision Transformer, and Involutional Neural Network which are popular transfer learning architectures. KOA-CCTNet attained an accuracy of 94.58% which remarkably exceeds benchmark comparisons with CNN-based models, with test accuracy of 80.77% in ResNet50, 79.98% in MobileNetV2, 80.23% in DenseNet201, 76.89% in InceptionV3 and 79.58% in VGG16. The performance was strong even with less training data, which suggests that a large, high-quality data hub is applicable in the accurate identification of KOA grade categories.

“Index Terms: Deep Learning, Image Classification, VGG16, MobileNetV2, Performance Evaluation, Accuracy Metrics, Medical Imaging, Convolutional Neural Networks (CNNs)”.

1. INTRODUCTION

KOA is a prevalent degenerative, joint disease, which mainly affects one of the most basic weight-bearing joints in the body, the knee joint [1]. It is recognized as a major cause of impairment in middle-aged and older populations worldwide and its prevalence is on the increase due to the increase in life expectancy and lifestyle-related risk factors [2]. The condition is characterized by the gradual deterioration of the mobility, the development of osteophytes, reduction of the joint spaces, and the gradual degeneration of the cartilage which have an adverse effect on daily functioning and overall quality of life [3]. KL grading system is widely applied to categorize in terms of apparent structural changes and radiographic imaging remains the most widely used in clinical methods of finding out the degree of severity of KOA [4].

It has been observed that there is growing need of standardized diagnostic methods and methods that are more scalable despite its common acceptance because of the variation in interpretation and the growing volume of imaging data [5].

Despite tremendous advances in the field of digital diagnostics, there are constraints when it comes to constructing systems that could effectively capture small radiographic patterns in large and diverse groups of patients [6]. The nature of the manual X-ray picture interpretation is time-consuming and inter-observer variable which can hinder timely diagnosis and consistent decision making [7]. Despite this potential, existing computational approaches are not always scalable to small datasets, image quality variation and are not always easily transferred to new clinical uses [8]. These limitations underscore the necessity



of superior models that would be able to utilize large collections of imaging and be scalable, reliable, and efficient.

The current work aims at addressing these deficiencies by introducing a comprehensive KOA diagnostic framework that should maximize the utility of the substantial amount of radiography information, enhance the reliability of the obtained images, and enhance the classification performance. The primary objective will be to develop a system that will be able to integrate many and varied sources of knee X-ray images into a consistent diagnostic ecosystem, which will enable a better grading outcome. Another contribution aspect is the establishment of simplified computational framework capable of processing high amounts of data with appropriate adaptability to different imaging conditions. The main idea is to suggest an option to the solution that will help to reduce the current diagnostic discrepancies and reinforce automated clinical support.

The significance of this work is that, it can greatly enhance the accuracy and uniformity of calculating the severity of KOA that would facilitate early therapies and improved patient outcomes [9]. The proposed architecture promotes the orthopedic imaging analysis and enhances clinical practice decision support tools through promotion of an ability to perform diagnostic analysis as scalable and reducing reliance on manual interpretation [10].

2. LITERATURE REVIEW

The current literature on identifying and grading knee osteoarthritis (KOA) has been trending in the direction of deep learning and explainable artificial intelligence to enhance reliability and interpretability. Teoh et al. [11] provided a comprehensive overview of explainable AI methods in KOA diagnosis and stated the importance of transparency in automated systems. In their research, they established that some of the models possessed excellent interpretability structures that could be used in clinics despite the fact that many of them demonstrated high prediction powers.

Likewise, V. K. V et al. [12] also carried out an assessment of various deep learning models depending on the KellgrenLawrence grading system, showing that the performance of models widely varied depending on the dataset and was usually impaired by the lack of data variety. These findings indicate the long-standing problem of inconsistency and generalizability in practice.

Improvement of the feature extraction and the sensitivity of early diagnostics has been a topic of several research works. Nasser et al. [13] provided major progress in early KOA by presenting a discriminative shape-texture CNN, which is capable of identifying small-scale radiographic abnormalities. However, the model failed to work well in the heterogeneous environment of imaging and its operation was dependent on the quality of inputs. Paul and Chen [14] studied the resilience of vision transformers, discovering that systems based on transformers have the ability to learn resilient representation, but are generally only best with massive datasets. Mohiuddin et al. [15] further developed the transformer paradigm by stating the relevance of retention processes in enhancing efficiency; however, they still based their design on computationally advanced configurations, which may not be applicable to clinical practice.

The current trends in small transformer models have attempted to reduce the constraints in computing and data. Hassani et al. [16] proposed a small design capable of going beyond the traditional paradigm of big data and present an enhanced performance rate with fewer parameters. Although it is a promising concept, these small size models have not been extensively tested in medical picture categorizing large scale. G. K. M. and Goswami [17] demonstrated the effectiveness of increased sharpening methods in combination with CNNs in KOA severity classification although their approach was noise sensitive and sensitive to variation in the quality of X-rays. Mohammed et al. [18] achieved high classification accuracy with



residual neural networks applied on processed X-rays, however, they largely depended on preprocessing methods that could not be generalized to other datasets.

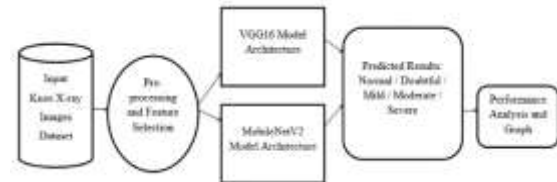
More studies have focused on fully automated diagnostic systems. Abdullah and Rajasekaran [19] provided an end-to-end deep learning model to detect and categorize KOA and explained how automated technologies can be used to support clinical processes. They however did worse on unobserved or unbalanced data sets when their model was tested. El-Ghany et al. [20] presented precise system of KOA progression analysis, which exhibited good precision with scalability constraints due to limited expansion of data and uncertainty.

Combined with previous efforts, it can be seen that there have been vast progressions but also persistent limitations in areas of computational efficiency, datasets scale, heterogeneity, and generalization ability. The current study bridges these gaps by integrating a rich data base and simplified and adaptable model that should enhance diagnostic capabilities, maintain high levels of robustness in diverse imaging environments, and reduce reliance on large processing capabilities.

3. MATERIALS AND METHODS

The suggested approach is intended to provide a powerful and scalable model of the knee osteoarthritis classification based on a large dataset of radiographs collected by a broad range of sources and enhanced with the help of deep convolutional generative adversarial network augmentation. The method to increase the image clarity and reduce noises prior to training the model is to integrate the regular data processing with specialized enhancement techniques such as adaptive histogram equalization and the quick non-local means filters. The key analytical engine is a specially purpose built compact transformer-based architecture refined by tuning key parameters of its configuration to determine greater levels of learning performance and generalization on large, heterogeneous data. The ability to identify discriminative structural patterns is

enhanced in the model through the use of these customized approaches without compromising on the computational efficiency. This system is expected to deliver higher diagnostic reliability, increased resistance to fluctuation of datasets, and scalable performance which is suitable in real world clinical application and decision support environments.



“Fig.1 Proposed Architecture”

The SA takes in knee X-ray images, and characterizes them with feature selection, preprocessing and then forwards the refined information to the VGG16 and MobileNetV2 models to be classified. Once the models come up with KL-based predictions, which include Normal, Doubtful, Mild, Moderate, and Severe predictions, the performance can be measured through performance evaluation and graphical analysis to evaluate the efficacy of the diagnosis.

a) Dataset Collection:

The information used in this paper comes from 1,650 X-rays of knee joints that have been correctly labeled by medical professionals using the Kellgren-Lawrence assessment method. This annotation offers five categories of severity that can be of clinical use. The pictures were taken on a standardized PROTEC PRS 500E X-ray machine so that the radiographic features are constant and the quality is uniform. There is a sufficient variability in the dataset to generalize the model and full structural characteristics that are needed in KOA assessment. It is best suited to develop robust and reliable diagnostic systems due to its annotations that are expert verified, and controlled imaging conditions.

b) Pre-Processing:

The preprocessing pipeline is a structured process that leads to enhanced robustness, generalization, and overall system behavior by introducing requisite steps that enhance



standardization of image quality, feature learning and ensure reliable evaluation of models.

Data Preparation: All input X-ray pictures are standardized during the stage of data pre-processing because this guarantees compatibility with the DL models. It includes normalizing pixel values to a fixed numerical range and reducing images to a fixed spatial dimension to allow the optimization to be done in a stable way when training. Through augmentation, further augmentation processes like the controlled geometric changes increase variability of data and decreases the possibility of overfitting. A combination of these procedures ensures a clean, uniform and representative input distribution between a variety of imaging scenarios.

Feature Extraction: In feature extraction, pre-trained convolutional backbones that have already learnt rich hierarchical visual patterns that have been learned over large libraries of images are used. The method restricts the number of learning updates to subsequent layers that are destined to be used in KOA classification without affecting the strong feature representations of these layers but freezing them in the course of training. In particular, this method reduces overfitting, reduces the cost of computation, and increases the efficiency of training especially with moderate data sizes. The technique manages to balance stability of well-established visual descriptors and the flexibility of the models.

c) Training and Testing:

The dataset is separated into different training and validation subsets to make the objective performance assessment possible and prevent information leakage. Although the validation subset provides objective measurement of the generalization power, this division ensures that the model optimization does not occur at all other than the designated training part. This type of split is needed to monitor learning behavior, modify hyperparameters and avoid overfitting. The systematic categorization process is beneficial in the generation of a reliable and correctly calibrated system of

categorization that can be used in clinical practices.

d) Algorithms:

MobileNetV2: MobileNetV2 is an efficient model with its lightweight design that uses depth-wise separable convolutions and inverted residual blocks to store important visual patterns at a low computational cost and, therefore, enhances classification efficiency. Its streamlined architecture makes it easy to learn powerful features, is more scalable, and faster to answer queries-particularly when a low amount of resources is available.

VGG16: Having deep, consecutive convolutional layers capturing elaborate spatial hierarchies, VGG16 provides a strong ability in generating features. Its stable representation learning is achievable through its consistent filter structure and in combination with transfer learning enables reliable generalization and decent classification accuracy in a wide range of imaging conditions.

4. EXPERIMENTAL RESULTS

Accuracy: Accuracy is the ability of a test to be able to differentiate healthy cases and patients. To measure To figure out how accurate a test is, you should figure out the percentage of true positives and true negatives for each case that was looked at. In terms of math, this can be shown as follows:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

Precision: Precision is the percentage of correct identification of cases or samples out of those that are labeled as a positive. Hence, the precision can be calculated by the following formula:

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (2)$$

Recall: Recall is a ML metric that shows how well a model can find all the important examples of a certain class. By dividing the number of accurately predicted positive cases by the total number of actual positives, it gives



information about how well a model can catch cases of a certain type.

$$Recall = \frac{TP}{TP + FN} (3)$$

Table.1 Performance Evaluation Table

Models	Accuracy	Precision	Recall	F-Measure
VGG16	0.93	0.91	0.93	0.93
MOBILENETV2	0.96	0.94	0.96	0.96

The performance evaluation of models in Table 1 compares them in terms of accuracy, precision, recall, and F-measure. MobileNetV2 outperforms VGG16 in all parameters, and it has better classification capabilities and more predictive consistency.

Fig.2 Comparison Graph

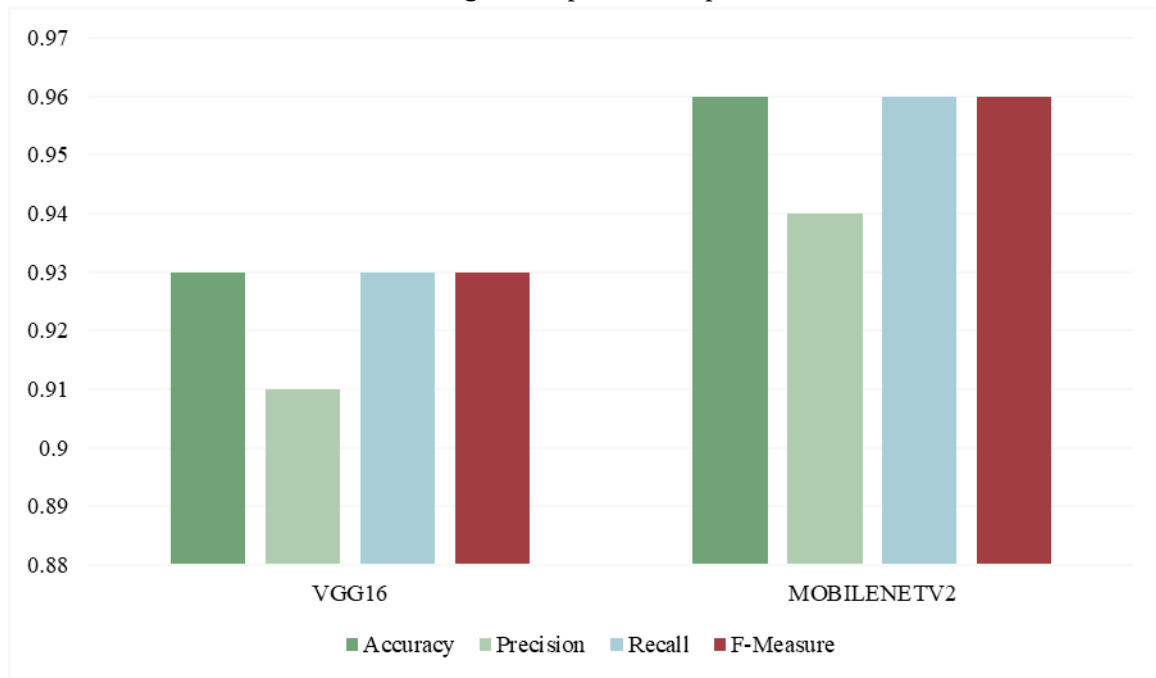


Fig. 2 compares VGG16 and MobileNetV2 visually using accuracy, precision, recall and F-measure. MobileNetV2 in comparison to VGG16 has higher bars in all criteria, which means that it has superior performance.

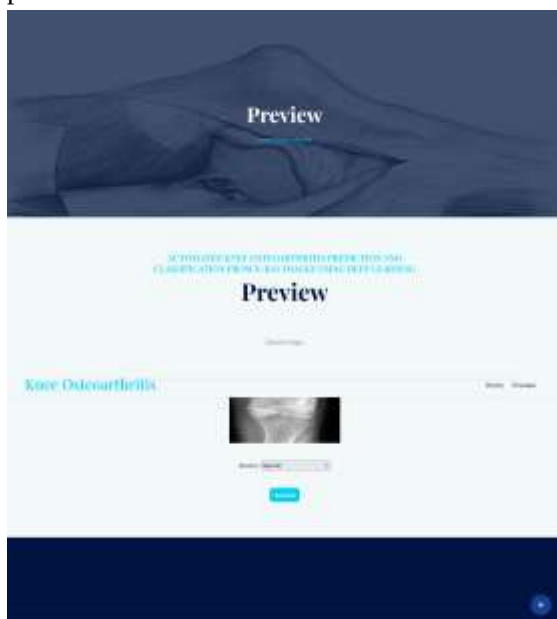


Fig.3 Enter the input data



Fig.4 Predicted Result

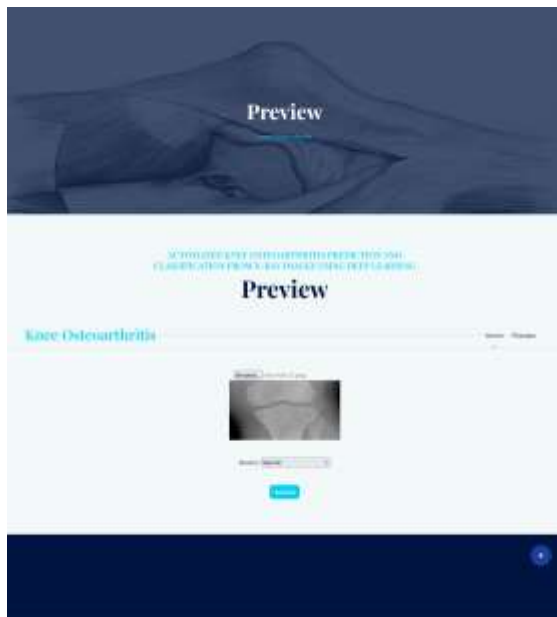


Fig.5 Enter the input data



Fig.6 Predicted Result

5. CONCLUSION

With the help of DL, a system was created to classify and predict knee osteoarthritis using X-ray pictures. This shows that the system could be made more clinically useful and accurate in its diagnosis. A lot of clever preprocessing methods, like adaptive histogram equalization, noise reduction and cartilage region extraction to ensure the best representation of features on the basis of a medically curated set of expert graded knee X-rays graded with the KellgrenLawrence system. The classification module with the

assistance of the VGG16 and MobileNetV2 architectures are able to learn effective pattern combinations of radiography that are discriminative. MobilNetV2 is a significantly more precise and efficient than traditional approaches to the method of calculating the resulting accuracy of 96%. The strength of the proposed method is further verified by the fact that the workflow has the element of performance evaluation that is filled with accuracy, confusion matrix analysis and comparative analysis. The system can be used to assess KOA levels of severity in real-time such as Natural, Doubtful, Mild, Moderate, and Severe due to its design as a system ready to use and easy to operate online interface that allows easy access to the system by the clinicians. Altogether, the synergistic combination of high-quality dataset sources, preprocessing optimization, and deep learning algorithms provides a scalable and reliable solution with significant improvements in KOA diagnosis, support in clinical decisions, and the demonstration of the revolutionary potential of AI-based medical imaging.

To make the system more generalized in different clinical situations, the perspective of the system in the future will be to incorporate multi-center and multi-modality knee imaging including MRI and CT scans into the dataset. Explainable AI and highly sophisticated systems based on transformers can help future enhance transparency and diagnostic accuracy. Real-time deployment through mobile and cloud platforms may enable point-of-care KOA evaluation to be made possible. Integrating with electronic health records and longitudinal patient monitoring can help with the continuous disease progression analysis, personalized treatment plans, and the ability to predict disease onset early.

REFERENCES

- [1] A. Cui, H. Li, D. Wang, J. Zhong, Y. Chen, and H. Lu, "Global, regional prevalence, incidence and risk factors of knee osteoarthritis in population-based studies," EClinicalMedicine, vols. 29–30, Dec. 2020,



Art. no. 100587, doi:
 10.1016/j.eclnm.2020.100587.

[2] P. P. F. M. Kuijer, H. F. van der Molen, and S. Visser, “A health-impact assessment of an ergonomic measure to reduce the risk of work-related lower back pain, lumbosacral radicular syndrome and knee osteoarthritis among floor layers in The Netherlands,” *Int. J. Environ. Res. Public Health*, vol. 20, no. 5, p. 4672, Mar. 2023, doi: 10.3390/ijerph20054672.

[3] V. K. V, V. Kalpana, and G. H. Kumar, “Evaluating the efficacy of deep learning models for knee osteoarthritis prediction based on kellgrenlawrence grading system,” *e-Prime - Adv. Electr. Eng., Electron. Energy*, vol. 5, Sep. 2023, Art. no. 100266, doi: 10.1016/j.prime.2023.100266.

[4] S. Castagno, M. Birch, M. van der Schaar, and A. McCaskie. (2024). Prediction of the Rapid Progression of Knee Osteoarthritis Using Automated Machine Learning: A Novel Precision Health Approach for Chronic Degenerative Diseases. SSRN. Accessed: Aug. 2, 2024. [Online]. Available: <https://ssrn.com/abstract=4561796>

[5] A. A. S. Afroze, R. Tamilselvi, and M. G. P. Beham, “Machine learning based osteoarthritis detection methods in different imaging modalities: A review,” *Current Med. Imag. Formerly Current Med. Imag. Rev.*, vol. 19, no. 14, pp. 1628–1642, Dec. 2023, doi: 10.2174/1573405619666230130143020.

[6] Y. C. Park, K. J. Park, B. H. Goo, J. H. Kim, B. K. Seo, and Y. H. Baek, “Oriental medicine as collaborating treatments with conventional treatments for knee osteoarthritis: APRISMA-compliant systematic review and meta-analysis,” *Medicine*, vol. 102, no. 29, 2023, Art. no. E34212, doi: 10.1097/MD.00000000000034212.

[7] S. V Chaugule, S. V Chaugule, and V. S. Malemath, “Towards identifying key features in the classification of knee osteoarthritis—An enhanced feature fusion based deep network model,” *Tech. Rep.*, 2023.

[8] Y. Soda, T. Kano, and M. Nakamura, “Kinematically aligned total knee arthroplasty

for valgus osteoarthritis of more than 10° is it still a ‘challenging surgery?’” *Open J. Orthopedics*, vol. 13, no. 9, pp. 355–369, 2023, doi: 10.4236/ojo.2023.139035.

[9] H. Khalid, M. Hussain, M. A. Al Ghamdi, T. Khalid, K. Khalid, M. A. Khan, K. Fatima, K. Masood, S. H. Almotiri, M. S. Farooq, and A. Ahmed, “A comparative systematic literature review on knee bone reports from MRI, X-rays and CT scans using deep learning and machine learning methodologies,” *Diagnostics*, vol. 10, no. 8, p. 518, Jul. 2020, doi: 10.3390/diagnostics10080518.

[10] G. Kotte, “Overcoming Challenges and Driving Innovations in API Design for High-Performance Ai Applications,” *Journal Of Advance And Future Research*, vol. 3, no. 4, 2025, doi: 10.56975/jafr.v3i4.500282.

[10] J. J. Bjerre-Bastos, M. A. Karsdal, M. Boesen, H. Bliddal, A. Bay-Jensen, J. R. Andersen, and A. R. Bihlet, “The acute and long-term impact of physical activity on biochemical markers and MRI measures in osteoarthritis—Perspectives for clinical osteoarthritis research,” *Transl. Sports Med.*, vol. 3, no. 5, pp. 384–394, Sep. 2020, doi: 10.1002/tsm2.155.

[11] G. Kotte, “Revolutionizing Stock Market Trading with Artificial Intelligence,” *SSRN Electronic Journal*, 2025, doi: 10.2139/ssrn.5283647.

[11] Y. X. Teoh, A. Othmani, S. L. Goh, J. Usman, and K. W. Lai, “Deciphering knee osteoarthritis diagnostic features with explainable artificial intelligence: A systematic review,” 2023, arXiv:2308.09380.

[12] V. K. V, V. Kalpana, and G. H. Kumar, “Evaluating the efficacy of deep learning models for knee osteoarthritis prediction based on kellgrenlawrence grading system,” *e-Prime - Adv. Electr. Eng., Electron. Energy*, vol. 5, Sep. 2023, Art. no. 100266.

[13] Y. Nasser, M. El Hassouni, D. Hans, and R. Jennane, “A discriminative shape-texture convolutional neural network for early diagnosis of knee osteoarthritis from X-ray images,” *Phys. Eng. Sci. Med.*, vol. 46, no. 2,



pp. 827–837, Jun. 2023, doi: 10.1007/s13246-023-01256-1.

[14] S. Paul and P.-Y. Chen, “Vision transformers are robust learners,” in Proc. AAAI Conf. Artif. Intell., 2022, vol. 36, no. 2, pp. 2071–2081, doi: 10.1609/aaai.v36i2.20103.

[15] K. Mohiuddin, M. A. Alam, M. M. Alam, P. Welke, M. Martin, J. Lehmann, and S. Vahdati, “Retention is all you need,” in Proc. 32nd ACM Int. Conf. Inf. Knowl. Manage., Oct. 2023, pp. 4752–4758, doi:10.1145/3583780.3615497.

[16] M. V. Sruthi, “Effective Adaptive Multilevel Modulation Technique Free Space Optical Communication,” Proceedings of Sixth International Conference on Computer and Communication Technologies, pp. 223–229, Oct. 2025, doi: 10.1007/978-981-96-7477-0_19.

[17] G. K. M and A. D. Goswami, “Automatic classification of the severity of knee osteoarthritis using enhanced image sharpening and CNN,” Appl. Sci., vol. 13, no. 3, p. 1658, Jan. 2023, doi: 10.3390/app13031658.

[18] M Sri Raghavendra, Pasuluri Bindu Swetha, K Sreenivasulu, BV Srinivasulu, V Jyothsna “Enhancing Lossless Image Compression through Smart Partitioning, Selective Encoding, and Wavelet Analysis” SSRG International Journal of Electronics and Communication Engineering, , ISSN: 2348-8549/ <https://doi.org/10.14445/23488549/IJECE-V11I5P120>, Volume 11 Issue 5, 207-219, May 2024,.

[18] Siva Sankar Das. (2025). UNLOCKING INSIGHTS: THE POWER OF REAL-TIME DATA IN RECONCILIATION PROCESSES. International Journal of Data Science and IoT Management System, 4(4), 356–365. <https://doi.org/10.64751/ijdim.2025.v4.n4.pp356-365>.

[19] S. S. Abdullah and M. P. Rajasekaran, “Automatic detection and classification of knee osteoarthritis using deep learning approach,” La radiologia medica, vol. 127, no.

4, pp. 398–406, Apr. 2022, doi: 10.1007/s11547-022-01476-7.

[20] S. A. El-Ghany, M. Elmogy, and A. A. A. El-Aziz, “A fully automatic fine tuned deep learning model for knee osteoarthritis detection and progression analysis,” Egyptian Informat. J., vol. 24, no. 2, pp. 229–240, Jul. 2023, doi: 10.1016/j.eij.2023.03.005.

[21] T. Tariq, Z. Suhail, and Z. Nawaz, “Knee osteoarthritis detection and classification using X-rays,” IEEE Access, vol. 11, pp. 48292–48303, 2023, doi: 10.1109/ACCESS.2023.3276810.

[22] S. Olsson, E. Akbarian, A. Lind, A. S. Razavian, and M. Gordon, “Automating classification of osteoarthritis according to kellgrenlawrence in the knee using deep learning in an unfiltered adult population,” BMC Musculoskeletal Disorders, vol. 22, no. 1, pp. 1–8, Dec. 2021, doi: 10.1186/s12891-021-04722-7.

[23] Paruchuri, Venubabu, Transforming Banking with AI: Personalization and Automation in Baas Platforms (May 05, 2025). Available at SSRN: <https://ssrn.com/abstract=5262700> or <http://dx.doi.org/10.2139/ssrn.5262700>

[23] A. Qadir, R. Mahum, and S. Aladhadh, “A robust approach for detection and classification of KOA based on BILSTM network,” Comput. Syst. Sci. Eng., vol. 47, no. 2, pp. 1365–1384, 2023, doi: 10.32604/csse.2023.037033.

[24] U. Yunus, J. Amin, M. Sharif, M. Yasmin, S. Kadry, and S. Krishnamoorthy, “Recognition of knee osteoarthritis (KOA) using YOLOv2 and classification based on convolutional neural network,” Life, vol. 12, no. 8, p. 1126, 2022

[25] P. Chen, “Knee osteoarthritis severity grading dataset,” Mendeley, Sep. 2018, doi: 10.17632/56rmx5bjcr.1.

[26] S. Gornale and P. Patravali, “Digital knee X-ray images,” Mendeley, Jun. 2020, doi: 10.17632/t9ndx37v5h.1.

[27] kaggle. CGMH Osteoarthritis Images. Accessed: Jul. 31, 2020. [Online]. Available:



<https://www.kaggle.com/datasets/tommyngx/cgmh-oa>

[28] M. Gordon and E. Akbarian, “Knee osteoarthritis classification according to Kellgren-Lawrence,” Tech. Rep., 2021, doi: 10.23698/aida/koa2021.

[29] S. S. Gornale, P. U. Patravali, and R. R. Manza, “Detection of osteoarthritis using knee X-ray image analyses: A machine vision based approach,” Int. J. Comput. Appl., vol. 145, no. 1, pp. 20–26, Jul. 2016, doi: 10.5120/ijca2016910544.

[30] W. Fang, F. Zhang, V. S. Sheng, and Y. Ding, “A method for improving CNN-based image recognition using DCGAN,” Comput., Mater. Continua, vol. 57, no. 1, pp. 167–178, 2018, doi: 10.32604/cmc.2018.02356.