# Enhancing Knee Osteoarthritis severity level classification using diffusion augmented images

Paleti Nikhil Chowdary[1], Gorantla V N S L Vishnu Vardhan[1], Menta Sai Akshay[1], Menta Sai Aashish[1], Vadlapudi Sai Aravind[1], Garapati Venkata Krishna Rayalu[1], and Aswathy P[1]

Amrita School of Artificial Intelligence, Coimbatore, Amrita Vishwa Vidyapeetham, India.
nikhil.paleti@outlook.com, vishnuvardhangorantla0308@gmail.com, akshaymenta24@gmail.com, aashishmenta249@gmail.com, aravindvadlapudi2003@gmail.com, gkrishnarayalu@gmail.com, p_aswathy@cb.amrita.edu

**Abstract.** This research paper explores the classification of knee osteoarthritis (OA) severity levels using advanced computer vision models and augmentation techniques. The study investigates the effectiveness of data preprocessing, including Contrast-Limited Adaptive Histogram Equalization (CLAHE), and data augmentation using diffusion models. Three experiments were conducted: training models on the original dataset, training models on the preprocessed dataset, and training models on the augmented dataset. The results show that data preprocessing and augmentation significantly improve the accuracy of the models. The EfficientNetB3 model achieved the highest accuracy of 84% on the augmented dataset. Additionally, attention visualization techniques, such as Grad-CAM, are utilized to provide detailed attention maps, enhancing the understanding and trustworthiness of the models. These findings highlight the potential of combining advanced models with augmented data and attention visualization for accurate knee OA severity classification.

**Keywords:** Deep Learning· Computer Vision· Knee Osteoarthritis (OA)· CLAHE· Data Augmentation · Diffusion Models.

## 1 Introduction

Osteoarthritis (OA) is a chronic degenerative condition marked by cartilage degradation that worsens over time and finally results in bone deterioration. Knee osteoarthritis (KOA), one of its many variations, primarily affects The medial, lateral, and patellofemoral joints—the three compartments of the knee joint. It usually develops gradually over a period of 10 to 15 years, causing interruptions in daily life [7, 12].


C. Kiran Mai et al. (eds.), *Proceedings of the Fourth International Conference on Advances in Computer Engineering and Communication Systems (ICACECS 2023)*, Atlantis Highlights in Computer Sciences 18,
https://doi.org/10.2991/978-94-6463-314-6_27

According to Wang et.al [13], in patients over 60 years of age, the prevalence of symptomatic KOA ranges from 10.0% to 16.0%, whereas that of radiographic KOA ranges from 35.0% to 50.0%. The prevalence of KOA has increased by twofold in men and treble in women in the USA during the past 20 years, affecting around 250 million individuals globally and placing a heavy burden on society.

KOA diagnosis involves symptom evaluation, arthroscopy, X-rays, and MRI. The Kellgren and Lawrence grading system presented by Kellgren JH in [6] is the standard for radiographic classification. for knee osteoarthritis. The osteoarthritis grading system includes uncertain Joint space narrowing and potential osteophytic lipping (Grade 1), osteophytes with clear signs of joint space narrowing, sclerosis, and potential bony deformity (Grade 2), multiple osteophytes with clear signs of joint space narrowing, sclerosis, and potential bony deformity (Grade 3), and prominent osteophytes with clear signs of joint space narrowing, severe sclerosis, and certain bony deformity (Grade 4).

This research addresses the gap in knee osteoarthritis (KOA) classification by introducing cutting-edge augmentation techniques such as diffusion models. It aims to classify X-ray images into 5 classes, highlights the impact of data preprocessing and augmentation, and enhance interpretability by utilizing gradcam images to explain model predictions. [14, 8]

Three experiments were conducted: The first used the original OAI dataset to train computer vision models. The second employed preprocessed OAI data using the CLAHE [9] method. The third introduced synthetic data through diffusion models for augmentation, then trained models on it. Comparative analysis of models followed, and model gradients were used to generate gradcam images, highlighting significant image areas for classification.

The subsequent section will provide a summarization of the relevant literature on knee osteoarthritis, followed by an explanation of the methodology employed in this study. Subsequently, the results and discussion section will present the findings derived from our experiments and delve into their analysis. Finally, the concluding section will present our overall conclusion.

## 2 Related Works

In [1], Fabi Prezja et al. used data augmentation with a WGAN to generate synthetic X-ray DeepFake images for knee osteoarthritis (KOA). Surgeons achieved 65.28% accuracy and radiologists achieved 59.40% accuracy in classifying real or fake images. In [11], Joseph Humberto Cueva et al. developed a semi-automatic CADx model using Deep Siamese CNNs and ResNet-34. Their model achieved a 61% accuracy for KOA detection and classification. In [10], Kevin A Thomas et al. used traditional augmentation techniques and a custom CNN model with ImageNet weights to automate KOA severity classification. They achieved 71% accuracy and an f1-score of 0.70. In [3], Abdelbasset Brahima et al. created a decision support tool for early KOA detection using X-ray imaging and machine learning. They achieved an accuracy of 82.90% with a random forest classifier.

These studies highlight various methods for processing KOA datasets, including traditional techniques and advanced approaches like GANs. However, there is limited research on using diffusion models for data augmentation and integrating them with advanced traditional processing techniques. This research aims to investigate the effectiveness of diffusion models for data augmentation and the impact of advanced image-processing techniques on KOA classification.

# 3 Methodology

We first preprocess the dataset using Contrast-Limited Adaptive Histogram Equalization (CLAHE) technique and prepared a preprocessed dataset which is used to train Diffusion models. Figure 1 displays the preprocessed Images for each class. Images are then generated from the diffusion models using a Denoising Diffusion Implicit Models (DDIM) scheduler and a final dataset is prepared by combining all the images. then various experiments are run as stated before and the current section describes all the aspects in detail.

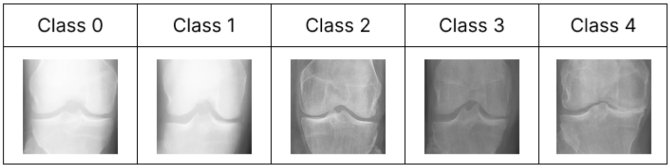


**Fig. 1.** Preprocessed Images

## 3.1 Dataset

The OAI dataset encompasses a comprehensive collection of X-ray images focusing on knee joint detection and knee KL grading. Knee Osteoarthritis Severity Grading Dataset [2] is organized from the OAI dataset which encompasses 9786 knee X-ray images of size 224x224x1, categorized into five distinct groups, known as KL-grades, which provide an indication of the severity of knee osteoarthritis. Detailed information about the dataset is presented in Table 1.

In this study, data preprocessing was performed using CLAHE [9] to enhance the quality and usability of the dataset. The block size was set to 8x8 pixels and the clip limit to 0.03 for preprocessing.

## 3.2 Data Augmentation

In order to address a dataset's limited amount of image samples and reduce the possibility of bias toward particular classes [4], data augmentation is essential. Diffusion models are one effective method for data augmentation. Diffusion

**Table 1.** OAI Dataset

| Severity (Class) | Train (75%) | Test (15%) | Valid (10%) | Total |
|---|---|---|---|---|
| 0 | 2,890 | 639 | 328 | 3,857 |
| 1 | 1,321 | 296 | 153 | 1,770 |
| 2 | 1,919 | 447 | 212 | 2,578 |
| 3 | 957 | 223 | 106 | 1,286 |
| 4 | 217 | 51 | 27 | 295 |

models develop an invertible generative process that enables the creation of fresh data samples based on the discovered distribution. We can add artificial images to our dataset that accurately reflect the statistical characteristics of the original data by utilizing diffusion models.

As the preferred diffusion model for data augmentation in our study, we used denoising diffusion implicit models (DDIMs). The pre-processed images were initially scaled to (64 x 64 x 1) dimensions before DDIMs were trained on them. Then, 200 photos for each class except class 0 were produced using the trained DDIM. Figure 2 shows a sample-generated image from each class

Diffusion models for data augmentation, like DDIMs, offer a potent method to increase the dataset's diversity and enhance the generalization skills of the trained models. Additionally, it is important to note that a DDIM scheduler, which chooses the level of diffusion at each training, was in charge of managing and directing the augmentation process step. This scheduler aids in regulating the trade-off between the diversity of the generated samples and their resemblance to the original data, so achieving a balance that improves model performance and reduces potential bias.

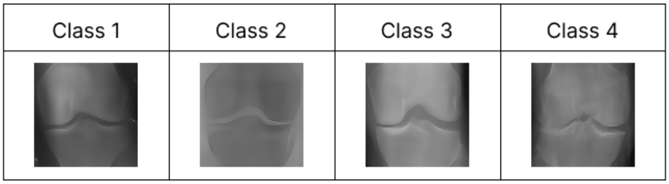


**Fig. 2.** Generated Images for each Class

The original dataset consists of knee osteoarthritis images with dimensions of 224 x 224. To improve visualization and analysis, we apply the CLAHE technique for image preprocessing [5]. However, considering computational constraints, we resize the preprocessed images to 64 x 64 for training the diffusion model.

To augment the dataset, we employ a trained diffusion model to generate additional images. For each class, except the control class, we generate 100 augmented images. Since the desired image size for classification is 224 x 224, we upscale the augmented images (originally 64 x 64) using the HAT Transformer and chaiNNer GUI. This upscale process increases the dimensions to 256 x 256, and we subsequently resize the images back to 224 x 224 using the Lancos interpolation method, which is clearly explained in Figure 3

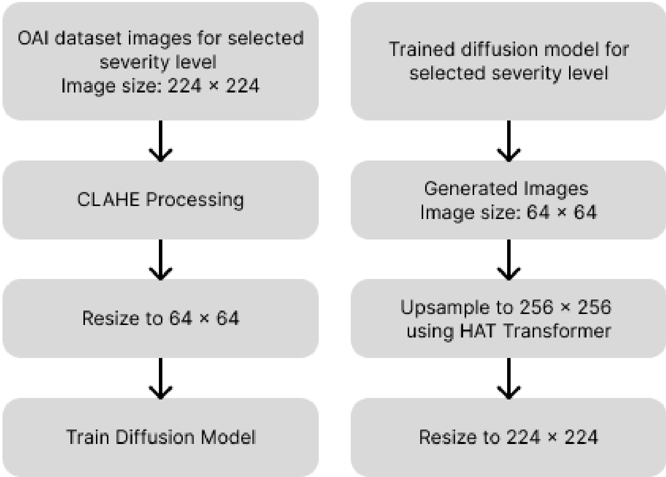


**Fig. 3.** Diffusion Methodology

### 3.3 Training

In this work, we examined feature extraction and fine-tuning as two methods for training vision models. Both methods made use of an augmented dataset, but we chose to concentrate on the benefits of fine-tuning rather than feature extraction. Figure 4 illustrates the visual differences in feature extraction and fine-tuning processes.

We make sure that the pre-trained weights, which capture important knowledge from a different task or domain, are kept in the fine-tuning procedure by initially freezing the basic model and training the models for a specific number of epochs. As a result, the model can take advantage of the underlying model's extensive representations and feature extraction skills.

We unfroze a portion of the base model layers after the initial training phase, usually the final 15 layers, and we continued training for several epochs. This action permits the pre-trained weights to be adjusted by the model to the particular task at hand, fine-tuning them to the subtleties and specifics of the augmented dataset. The model can learn task-specific attributes necessary for optimum performance on the target task by updating these unfrozen layers.

After preparing the datasets, The model was optimized using Lora with a projection rank of 16 after being loaded into memory with 8-bit precision. The optimized model was then trained for 10 epochs with an 8-batch size and a $10^{-3}$ learning rate.

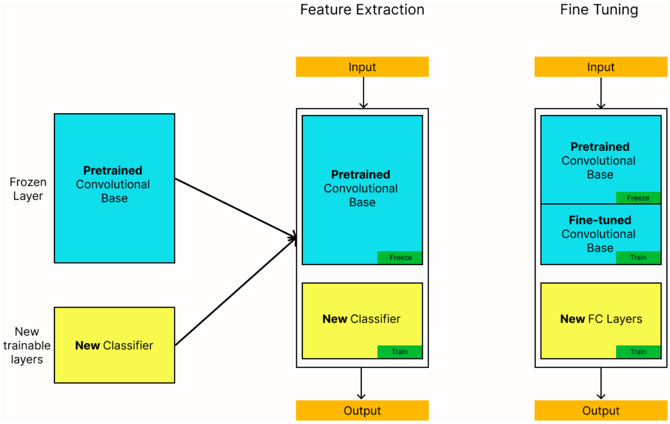


**Fig. 4.** Feature Extraction Vs Fine-tuning

## 4 Results and Discussion

The results obtained from the experiments are summarized in Table 2. The table presents the accuracy scores of different models when trained on three different datasets: original, preprocessed, and augmented. These results provide insights into the performance of the models and the impact of data preprocessing and augmentation techniques on the classification of knee osteoarthritis (OA) severity levels.

**Table 2.** Comparison of various models on 3 types of dataset

| Model | Original | Preprocessed | Augmented |
|---|---|---|---|
| Xception | 52% | 76% | 79% |
| VGG16 | 78% | 77% | 82% |
| ConvNeXtTiny | 72% | 80% | 81% |
| EfficientNet B3 | 68% | 76% | 84% |
| Densenet 201 | 60% | 76% | 79% |
| Vision Transformer | 69% | 71% | 72% |
| Swin Transformer V2 | 72% | 73% | 73% |

Overall, the results indicate that data preprocessing and augmentation techniques have a positive impact on the performance of the models for classifying knee OA severity levels. Data augmentation, in particular, consistently improved the accuracy of the models across different architectures. The EfficientNetB3 model achieved the highest accuracy of 84% on the augmented dataset, showcasing the potential of combining advanced models with augmented data for accurate classification of knee OA severity.

### 4.1 Visualizing GradCam

This study emphasizes the need to visualize attention in CNNs and ViTs. To understand how these models focus on different image regions and make decisions, attention visualization methods like Grad-CAM, CAM, and self-attention heatmaps are explored.

Grad-CAM, which can offer precise attention maps for intermediate layers as shown in Figure 5, is our favored technology among these ones. Grad-CAM can be used with different CNN and ViT designs because it is model-agnostic. Grad-CAM has the unique ability to overlay attention maps on the original image, providing a clear visual representation of the model's attentional focus.

Using the EfficientNetB3 model, our study classifies osteoarthritis, focusing on joint spacing as a vital feature. This aligns with osteoarthritis severity indicators, deepening our understanding of the model's decision-making. By employing Grad-CAM and similar techniques, we aim to enhance CNNs and ViTs transparency, reliability, and credibility, thereby improving model quality, fairness, and the ability to identify biases or weaknesses.

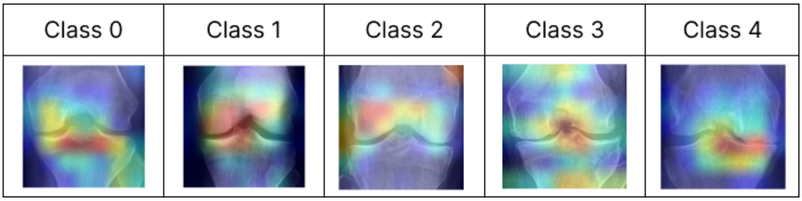


**Fig. 5.** AttentionMap Visualization Using Gradcam

## 5 Conclusion and Future Scope

In Conclusion, CNN models are effective in correctly classifying OA severity levels, as shown by our study on the classification of knee osteoarthritis (OA) severity levels. We have identified optimal techniques for preprocessing and data augmentation, including diffusion-based augmentation and explainable AI (Grad-CAM), while comparing various CNN and transformer designs. Efficientnet-B3 outperformed other models, displaying greater accuracy in OA severity classification. This automated analysis of knee X-ray images holds promise for advancing OA diagnosis, treatment, informed decision-making, and patient care. The results underscore deep learning's potential to address knee OA assessment challenges, paving the way for specialized treatments and improved patient outcomes.

Future work includes investigating various pre-processing techniques and evaluating their impact on knee osteoarthritis categorization. Finding the best strategy for enhancing model accuracy and generalization will also depend on researching and evaluating the efficacy of various augmentation strategies.

## References

1. Brahim, A., Jennane, R., RIAD, R., Janvier, T., Khedher, L., Toumi, H., Lespessailles, E.: 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 **73** (04 2019). https://doi.org/10.1016/j.compmedimag.2019.01.007
2. Chen, P.: Knee osteoarthritis severity grading dataset. Mendeley Data **1** (2018)
3. Cueva, J., Castillo, D.P., Espinos Morato, H., Duran, D., Díaz, P., Lakshminarayanan, V.: Detection and classification of knee osteoarthritis. Diagnostics **12**, 2362 (09 2022)
4. E., N.K., M., K., P., P., R., A., S., V.: Tomato leaf disease detection using convolutional neural network with data augmentation. In: 2020 5th International Conference on Communication and Electronics Systems (ICCES). pp. 1125–1132 (2020). https://doi.org/10.1109/ICCES48766.2020.9138030
5. Harichandana, M., Vishvanathan, S., Sajithvariyar, V., Sivanpillai, R.: Comparison of image enhancement techniques for rapid processing of post flood images. vol. XLIV-M-2-2020, pp. 45–50 (11 2020). https://doi.org/10.5194/isprs-archives-XLIV-M-2-2020-45-2020
6. Kellgren, J.H., Lawrence, J.S.: Radiological assessment of osteoarthrosis. Annals of the Rheumatic Diseases **16**(4), 494–502 (1957). https://doi.org/10.1136/ard.16.4.494, https://ard.bmj.com/content/16/4/494
7. Lespasio, M.: Knee osteoarthritis: A primer. The Permanente Journal **21** (12 2017). https://doi.org/10.7812/TPP/16-183
8. Pooja, A., Mamtha, R., Sowmya, V., Soman, K.P.: X-ray image classification based on tumor using gurls and libsvm. In: 2016 International Conference on Communication and Signal Processing (ICCSP). pp. 0521–0524 (2016). https://doi.org/10.1109/ICCSP.2016.7754192
9. Reza, A.: Realization of the contrast limited adaptive histogram equalization (clahe) for real-time image enhancement. VLSI Signal Processing **38**, 35–44 (08 2004). https://doi.org/10.1023/B:VLSI.0000028532.53893.82
10. Song, J., Meng, C., Ermon, S.: Denoising diffusion implicit models. CoRR **abs/2010.02502** (2020), https://arxiv.org/abs/2010.02502
11. Thomas, K., Kidziński, u., Halilaj, E., Fleming, S., Venkataraman, G., Oei, E., Gold, G., Delp, S.: Automated classification of radiographic knee osteoarthritis severity using deep neural networks. Radiology: Artificial Intelligence **2**, e190065 (03 2020). https://doi.org/10.1148/ryai.2020190065
12. Veronese, N., Honvo, G., Bruyère, O., Rizzoli, R., Barbagallo, M., Maggi, S., Smith, L., Sabico, S., Al-Daghri, N., Cooper, C., Pegreffi, F., Reginster, J.Y.: Knee osteoarthritis and adverse health outcomes: an umbrella review of meta-analyses of observational studies. Aging Clinical and Experimental Research **35** (11 2022). https://doi.org/10.1007/s40520-022-02289-4
13. Wang, L., Lu, H., Chen, H., Jin, S., Wang, M., Shang, S.: Development of a model for predicting the 4-year risk of symptomatic knee osteoarthritis in china: a longitudinal cohort study. Arthritis Research & Therapy **23**, 1–13 (2021)
14. Yadav, P., Menon, N., Ravi, V., Vishvanathan, S.: Lung-gans: Unsupervised representation learning for lung disease classification using chest ct and x-ray images. IEEE Transactions on Engineering Management **70**(8), 2774–2786 (2023). https://doi.org/10.1109/TEM.2021.3103334