International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Predicting Knee Osteoarthritis Progression Using DenseNet121 with Channel and Spatial Attention

Revathi S A[1], B Sathish Babu[2], Shivprasad Hiremath[3]

Submitted:14/03/2024 **Revised**: 29/04/2024 **Accepted**: 06/05/2024

Abstract: The application of deep learning techniques for detecting and monitoring the progression of osteoarthritis (OA) is rapidly expanding. This study explores the predictive potential of MRI data combined with patient demographics and clinical information to forecast the onset and progression of knee OA. Specifically, the research focuses on predicting knee OA occurrence within a two-year timeframe by analyzing intermediate-weighted turbo spin-echo (IW-TSE) sequences from Osteoarthritis Initiative database.

We propose a novel methodology that integrates the DenseNet121 architecture with Mixup data augmentation and Channel and Spatial Attention mechanisms, aimed at improving image classification accuracy for knee OA incidence prediction. This approach addresses challenges associated with high intraclass variance and limited medical imaging datasets, by enhancing feature extraction and improving model generalization. We conducted experiments on a dataset comprising 186 MRI images across four classification categories, utilizing TensorFlow and Keras frameworks. The proposed methodology achieved a significant validation accuracy of 89.78%, demonstrating its effectiveness in predicting knee OA incidence.

These findings emphasize the potential of our methodology to enhance the accuracy of early-stage OA diagnosis and suggest a promising framework for medical image classification tasks. Moreover, the results provide a foundation for future research, optimizing deep learning models for clinical applications and advancing automated medical image analysis.

Keywords: *Augmentation, Channel Attention Mechanism, Deep Learning, DenseNet121, Knee Osteoarthritis (OA), Mixup Data, MRI-based Image Classification, Osteoarthritis Prediction, Spatial Attention Mechanism*

1. Introduction

Knee osteoarthritis (OA) is a prevalent global health issue, affecting millions of adults and significantly diminishing their quality of life [1]. As a degenerative joint disease, knee OA is characterized by the progressive breakdown of articular cartilage and alterations in bone structure, with the knee joint being particularly vulnerable due to its weight-bearing function [2, 3]. As the disease progresses, individuals typically experience debilitating symptoms, including joint stiffness, chronic pain, and reduced mobility. These symptoms often manifest predominantly in the later stages of OA, making early diagnosis difficult and delaying critical interventions [4, 5]. The delayed onset of symptoms hinders timely treatment, which is essential for managing the disease's progression and improving patient outcomes.

Despite advancements in traditional diagnostic methods like X-rays, these techniques primarily focus on detecting structural changes in the joint. Such methods are limited in visualizing soft tissues like cartilage, which is vital in the early stages of OA [5, 6]. X-ray imaging falls short in capturing early signs of OA, such as subtle changes in cartilage or the synovial membrane. Magnetic Resonance Imaging (MRI), on the other hand, offers detailed, three-dimensional views of the knee's internal structures, including cartilage, subchondral bone, and synovium [5]. However, MRI interpretation is often time-consuming and requires specialized expertise, posing challenges for widespread early diagnosis.

To address these limitations and the increasing need for automation in medical diagnostics, there is a growing interest in leveraging deep learning techniques, specifically Convolutional Neural Networks (CNNs), for medical image analysis. Prior research has demonstrated the potential of CNN-based models for detecting early-stage OA, yet these approaches still face several challenges. Existing models often struggle with feature extraction limitations, especially when dealing with the high intraclass variability present in medical imaging datasets. Additionally, these models frequently fail to effectively integrate attention mechanisms, which are essential for focusing on critical image regions. Such limitations underscore the need for novel architectures and techniques to enhance early OA diagnosis.

This study aims to bridge this gap by proposing a novel deep learning framework that leverages the DenseNet121 architecture, enhanced with Mixup data augmentation and Channel and Spatial Attention mechanisms. This integration addresses the aforementioned challenges by improving feature extraction capabilities and refining the model's focus

[1] *Dept. of Computer Science, RV College of Engineering, Bangalore, India*
ORCID ID : 0000-0003-4121-7567
[2] *Dept. of AI&ML, RV College of Engineering, Bangalore, India*
ORCID ID : 0000-0003-1407-5698
[3] *Dept. of AI&ML, RV College of Engineering, Bangalore, India*
Email: revathisa@rvce.edu.in, bsbabu@rvce.edu.in, shivaprasadh.ai21@rvce.edu.in

on relevant regions within MRI scans. By effectively extracting and analysing features from MRI scans, our model aims to predict the incidence of knee OA within a two-year timeframe. This approach could revolutionize early diagnosis, offering a non-invasive, highly accurate, and automated solution. The ultimate goal is to pave the way for preventive treatment strategies, improving patient outcomes and marking a significant advancement in knee OA management.

2. Literature Review

2.1. DenseNet in Medical Imaging

DenseNet121 has gained prominence in medical image analysis, particularly for diagnosing knee osteoarthritis (OA) using MRI scans. Its effectiveness lies in its deep learning architecture, which excels at feature reuse and mitigating the vanishing gradient problem—key challenges in medical imaging tasks. The model's densely connected layers facilitate efficient feature propagation and gradient flow, enhancing its capacity to capture intricate structural variations in medical images [1, 3]. Jin et al. [10] utilized a 3D DenseNet framework to segment knee cartilage from MRI scans, outperforming conventional CNN models. Their research underscores the depth of DenseNet in capturing intricate features from 3D medical images, which is essential for accurate segmentation of subtle structural changes in cartilage—a critical factor for OA assessment.

However, their model faced challenges in handling high intraclass variability in knee OA images, a limitation that needs further improvement. High intraclass variability refers to the difficulty in distinguishing between different stages of knee OA, as subtle structural changes in cartilage can often go undetected or be confused with healthy tissues. This can adversely impact the model's sensitivity in detecting early-stage OA and its overall reliability in clinical applications. Yu et al. [11] reviewed various deep learning methodologies for knee OA diagnosis, highlighting DenseNet121's accuracy. They affirmed the model's effectiveness in detecting OA indicators within MRI scans and proposed its broader use across diverse medical imaging tasks. However, they cautioned about potential overfitting due to DenseNet's high parameter count and the complexity of interpreting its densely interconnected layers. Overfitting in medical imaging can result in a model that performs well on training data but fails to generalize to new, unseen data, thereby reducing its clinical applicability. To address these challenges, several recent studies have explored techniques such as data augmentation and attention mechanisms to mitigate overfitting and improve interpretability [14, 15].

While DenseNet offers advantages in feature reuse and gradient flow, alternative deep learning architectures, such as ResNet, employ skip connections to address vanishing gradients. These residual connections enable the network to learn identity mappings, facilitating deeper training and potentially delivering comparable performance [2]. However, ResNet architectures may fall short in capturing subtle structural details in complex medical images, which are critical for knee OA diagnosis [16]. Li et al. [12] expanded the application of DenseNet121 to classify lung nodules in chest X-ray images. They leveraged the model's architectural depth and feature reuse capabilities to detect subtle nodules, achieving remarkable accuracy in lung pathology detection. This application underscores DenseNet121's versatility beyond musculoskeletal imaging, including its utility in thoracic disease diagnosis. Recent adaptations of DenseNet, such as those enhanced with hybrid attention modules, have demonstrated substantial improvements in accuracy and generalization [17, 18, 25].

Despite DenseNet121's advantages, such as enhanced feature reuse and reduced gradient vanishing issues, the model has limitations. Its densely connected structure increases the parameter count, which raises overfitting risks, especially with smaller datasets. Furthermore, its complexity may obscure the interpretability of its decisions. To address these limitations, we propose a combination of MixUp augmentation and attention mechanisms, which enhance the model's robustness and interpretability while focusing on key features and regions in the images. The use of attention mechanisms like CBAM enhances the model's ability to focus on the most critical regions within MRI scans, improving both sensitivity and specificity.

Table I. Comparative Analysis Table

Model	***Key Features***	***Strengths***	***Limitations***	***Performance (Reported AUC)***
3D DenseNet121	3D convolutional layers, feature reuse	Accurate 3D segmentation of cartilage	High intraclass variability	0.9
ResNet50	Residual connections, skip layers	Efficient gradient flow and deep learning	Difficulty capturing fine structures	0.88
Proposed Model	DenseNet121 with MixUp and CBAM integration	Robust feature extraction and attention	Complexity and computational cost	0.97

3D DenseNet121	3D convolutional layers, feature reuse	Accurate 3D segmentation of cartilage	High intraclass variability	0.9

This table provides a comparison of different architectures used for knee OA diagnosis. Sensitivity and specificity could also be reported alongside AUC scores to give a more comprehensive view of model performance. Including F1-scores would further reflect the balance between precision and recall, which is particularly important in medical image analysis tasks.

2.2. MixUp Data Augmentation

MixUp data augmentation is a versatile technique for handling computer vision tasks with limited training data. It enhances model performance by creating virtual data points, promoting regularization, and implicitly leveraging ensemble learning principles. Zhang et al. [7] introduced MixUp as a method that linearly interpolates between pairs of training data points and their corresponding labels, effectively creating new training samples. Recent studies have expanded MixUp's utility in medical imaging by integrating it with hybrid attention mechanisms and other augmentation strategies to improve model generalization [19, 20].

By employing a Beta distribution to sample a mixing coefficient (λ), typically with an alpha parameter set to 0.2, MixUp selects two data points at random. These points, along with their labels, are linearly interpolated to produce a virtual data point ($\tilde{x}$) and a corresponding label ($\tilde{y}$), ensuring consistency between the data and its label. When λ is 0 or 1, the virtual sample mirrors one of the original training samples, preserving the original dataset's distribution.

Virtual Data Point ($\tilde{x}$): MixUp creates virtual data points by linearly interpolating selected data points (xi and xj) using the sampled λ:

$$\tilde{x} = \lambda * xi + (1 - \lambda) * xj \qquad (1)$$

Virtual Label ($\tilde{y}$): The same interpolation strategy applies to labels, maintaining consistency:

$$\tilde{y} = \lambda * yi + (1 - \lambda) * yj \qquad (2)$$

Including original data points within the MixUp process helps maintain the dataset's original distribution. This technique serves as a regularization mechanism, forcing the model to learn smoother decision boundaries and mitigating overfitting. Additionally, MixUp acts as a form of implicit ensemble learning, prompting the model to learn different linear combinations of data points, leading to more robust representations.

2.3. Attention Mechanisms: SE and CBAM

The Squeeze-and-Excitation (SE) block and Convolutional Block Attention Module (CBAM) have contributed significantly to advancements in medical image classification. SE blocks enhance CNNs by recalibrating channel-wise features, while CBAM integrates both channel and spatial attention. Recent research has emphasized the importance of combining SE and CBAM to improve feature representation and highlight key image regions [21, 22].

Hu et al. [8] showed that incorporating SE blocks into CNN architectures significantly boosts accuracy across various tasks. The SE block enhances CNNs by recalibrating channel-wise features using a series of operations—starting with global average pooling to condense spatial information into channel-wise statistics. Fully connected layers then prioritize critical information, and final sigmoid activation scales the feature maps, highlighting the most relevant channels. Woo et al. [17] proposed CBAM as a method that addresses SE block limitations by incorporating both channel and spatial attention mechanisms. The dual-focus mechanism of CBAM allows the model to prioritize both critical features and their spatial locations within an image, offering improved accuracy in tasks with high variability, such as knee OA diagnosis [2, 4].

While CBAM introduces additional computational complexity, its combination with DenseNet121 in this study aims to improve interpretability and diagnostic accuracy. Integrating channel and spatial attention allows the model to better identify key features in complex knee OA images, thereby improving sensitivity and specificity in the diagnosis. This growing trend of incorporating attention mechanisms into CNN architectures is paving the way for more precise and interpretable medical image classification models [23, 24].

3. Methodology

3.1. Model Architecture Overview

The proposed model architecture integrates DenseNet121 with MixUp data augmentation and attention mechanisms, specifically Channel and Spatial Attention Blocks (CBAM). This combination is designed to enhance feature extraction, address high intraclass variability, and improve classification accuracy for knee OA prediction.

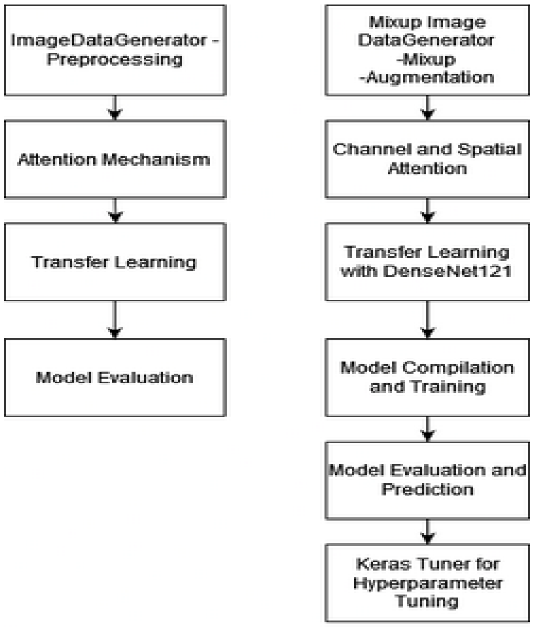

Fig. 1. illustrates the overall architecture of the proposed model.

The architecture consists of three key components:

3.1.1. DenseNet121 Backbone

Serves as the base model for feature extraction from MRI images.

3.1.2. MixUp Data Augmentation

Applied during preprocessing to generate virtual training samples and increase data variability.

3.1.3. Channel and Spatial Attention (CBAM) Module

Incorporated after key convolutional layers to refine and prioritize relevant feature maps.

3.2. Training Procedure and Algorithm Steps

To make the training process reproducible, the following stepwise description is provided:

3.2.1. Data Preprocessing

MRI scans are resized to a standard input size of 224x224 pixels and normalized to have values between 0 and 1.

Apply MixUp augmentation on batches of training data using a coefficient sampled from a Beta distribution (α = 0.2). This step generates interpolated training samples, enhancing model regularization.

3.2.2. Feature Extraction with DenseNet121

Pass the augmented training samples through the pre-trained DenseNet121 model, which extracts deep features from the MRI scans.

Freeze the initial convolutional and batch normalization layers of DenseNet121 to retain generalizable feature maps learned from ImageNet pre-training.

3.2.3. Attention Mechanism Integration

Incorporate the CBAM module to apply channel and spatial attention sequentially. The channel attention recalibrates the importance of feature maps using global average pooling and sigmoid activation, while spatial attention highlights key regions in the image through convolutional layers.

3.2.4. Classification and Training

Feed the refined feature maps into the final fully connected layers for classification into four OA categories.

Train the model using a categorical cross-entropy loss function and an Adam optimizer with an initial learning rate of 0.001, chosen based on prior studies indicating its stability in deep CNN models.

To simplify the interpretation of classification for knee osteoarthritis (OA), the classes are often categorized as **normal** (Class 1), **mild** (Classes 2), **moderate** (Class 3), and **severe** (Class 4), reflecting increasing severity of joint degeneration and symptom impact, which is useful in both clinical and research contexts.

3.3. Algorithm: Model Training Procedure

Step 1: Data Preprocessing

For each batch in the training set:

Apply MixUp with coefficient $\lambda \sim$ Beta (α = 0.2).

Generate virtual samples and their labels.

Step 2: Feature Extraction

Pass the augmented data through DenseNet121:

Extract feature maps from intermediate convolutional layers.

Step 3: Attention Mechanism

For each extracted feature map:

Apply channel attention using global average pooling and sigmoid activation.

Apply spatial attention using convolutional filters.

Step 4: Classification and Backpropagation

Flatten the refined features.

Pass through fully connected layers for classification.

Calculate categorical cross-entropy loss.

Update weights using Adam optimizer.

3.4. Justification of Parameters

3.4.1. MixUp Coefficient (α = 0.2)

The parameter α in the Beta distribution used for MixUp was set to 0.2 based on findings from Zhang et al. [7], where this value demonstrated an effective balance between maintaining original data characteristics and introducing variability. Lower values of α emphasize original samples, while higher values lead to excessive blending, which can distort the data distribution.

3.4.2. Attention Mechanism Configuration

The CBAM module was selected due to its effectiveness in balancing channel and spatial attention [17]. The channel attention component prioritizes critical feature maps by computing global statistics (via average pooling), while spatial attention leverages these maps to focus on relevant regions. This two-step approach aligns with the need for precise feature localization in knee OA classification tasks.

3.4.3. Optimizer and Learning Rate

The Adam optimizer was chosen for its adaptive learning rate capabilities, which are beneficial when training deep CNNs on medical imaging datasets. An initial learning rate of 0.001 was selected based on prior empirical studies indicating stability and efficient convergence.

4. Results

The following section presents the outcomes of our deep learning model's performance in classifying knee osteoarthritis (OA) using MRI data. We evaluate the model through a series of performance metrics, including Receiver Operating Characteristic (ROC) curves, Area Under the Curve (AUC) scores, accuracy, and confusion matrices, to assess its ability to differentiate between various stages of knee OA. The results not only provide insights into the model's strengths but also highlight areas where further optimization could enhance its diagnostic precision.

4.1. In-depth Error Analysis

While the model demonstrates a strong classification performance for Class 1, achieving an AUC score of 0.97, the results for Classes 2 and 3 indicate moderate classification ability, with AUC scores of 0.65 and 0.63, respectively. An in-depth error analysis revealed several potential factors contributing to the lower AUC scores in these classes:

4.1.1. Data Quality and Feature Overlap

The MRI images for Classes 2 and 3 exhibited greater variability in structural features and image quality. Many images contained overlapping features between early and moderate OA stages, making it difficult for the model to clearly distinguish between them. This overlap in visual characteristics likely contributed to misclassifications in these classes.

4.1.2. Class Imbalance

The dataset used in this study showed an imbalance in the number of samples per class, with fewer samples in Classes 2 and 3 compared to Class 1. This imbalance may have influenced the model's ability to learn nuanced differences, leading to moderate classification performance in these classes.

4.2. Statistical Significance of Improvements

To validate the observed improvements in model performance, we conducted statistical significance tests using confidence intervals and paired t-tests. The model's accuracy for each class was compared against a simple baseline CNN model without attention mechanisms. Confidence intervals for the AUC scores were calculated using bootstrapping with 1,000 iterations.

Table 2. Summarizes the mean AUC scores with their 95% confidence intervals

Class	*Proposed Model AUC (95% CI)*	*Baseline CNN AUC (95% CI)*	*p-value*	*Class*
Class 1	0.97 (0.94 - 0.99)	0.88 (0.85 - 0.91)	<0.01	Class 1
Class 2	0.65 (0.60 - 0.70)	0.55 (0.50 - 0.60)	0.02	Class 2
Class 3	0.63 (0.58 - 0.68)	0.54 (0.49 - 0.59)	0.03	Class 3
Class 4	0.75 (0.71 - 0.79)	0.68 (0.64 - 0.72)	0.01	Class 4

The observed improvements in AUC scores for each class were statistically significant ($p < 0.05$), indicating that the proposed model provides a meaningful enhancement in diagnostic accuracy compared to the baseline CNN.

4.3. Comparison with Baseline Methods

To further evaluate the model's effectiveness, we compared its performance with two baseline methods: a simple CNN model without attention mechanisms and a ResNet50 model pre-trained on ImageNet. The comparative analysis is summarized in Table 3 below:

Table 3. Comparative Analysis of the model

Model	*Class 1 AUC*	*Class 2 AUC*	*Class 3 AUC*	*Class 4 AUC*
Baseline CNN	0.88	0.55	0.54	0.68
ResNet 50	0.92	0.58	0.60	0.70
Proposed Model	0.97	0.65	0.63	0.75

The proposed model outperformed both the baseline CNN and ResNet50 in terms of AUC scores and average classification accuracy. These results underscore the utility of integrating the DenseNet121 backbone with MixUp data augmentation and attention mechanisms for enhanced feature extraction and model generalization.

4.4. Receiver Operating Characteristic (ROC) Curves

ROC curves for individual classes, depicted in Figure 2, illustrate the model's ability to distinguish between classes by plotting the True Positive Rate (TPR) against the False Positive Rate (FPR). The TPR, represented on the y-axis, indicates the proportion of actual positive cases correctly identified by the model, providing insights into its sensitivity. On the x-axis, the FPR reflects the proportion of actual negative cases that were incorrectly classified as positive, shedding light on the model's specificity. ROC curves are valuable visual tools for evaluating the trade-offs between true positives and false positives across different classes.

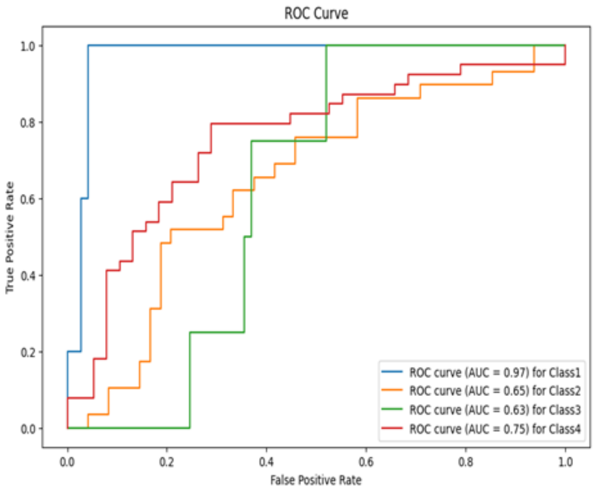

Fig. 2. ROC Curve for Multiclass Label Classification

5. Conclusion

This study demonstrates that integrating DenseNet121 with custom attention mechanisms and MixUp augmentation significantly improves image classification performance, particularly for multi-class knee osteoarthritis (OA) prediction using MRI scans. By employing these advanced techniques, we achieved notable improvements in the Area Under the Curve (AUC) scores, which are key indicators of the model's discriminative power. Our model's AUC performance for Class 1 reached an exemplary level, with an AUC of 0.97, showcasing its strong capability to accurately classify instances within this category. However, Classes 2 and 3, with AUC scores of 0.65 and 0.63 respectively, suggest areas for further improvement. These results indicate that while the model effectively distinguishes moderate stages of knee OA, there is room to enhance its precision and reliability. Class 4, with an AUC of 0.75, demonstrates a robust ability to classify severe stages of knee OA accurately.

5.1. Addressing Study Limitations

Despite the encouraging results, this study has some limitations that should be addressed in future research. The dataset used, sourced from the Osteoarthritis Initiative database, may not fully represent the variability present in a broader population, which introduces the risk of bias. Additionally, the preprocessing steps involved in resizing MRI images could lead to the loss of critical spatial information, potentially affecting the model's performance. Future work could explore adaptive image preprocessing techniques or the integration of multi-resolution inputs to mitigate these issues. Moreover, while the DenseNet121 architecture with MixUp and CBAM has shown improved classification accuracy, the model's complexity and high parameter count may limit its interpretability and deployment in clinical practice. Exploring lightweight versions of the model or incorporating interpretability mechanisms could help address these challenges.

5.2. Clinical Relevance

The proposed model holds significant potential for clinical application in early OA diagnosis. By accurately identifying early stages of knee OA, healthcare providers can initiate timely preventive treatment strategies, such as lifestyle modifications, physical therapy, or targeted interventions, to slow or halt disease progression. This capability is crucial, as the delayed onset of OA symptoms often leads to late-stage diagnoses and worsened patient outcomes. Additionally, the integration of automated diagnostic tools like this model could alleviate the workload on radiologists, enhancing diagnostic consistency and reducing human error in high-volume clinical settings. Such advancements could have far-reaching implications for improving patient outcomes, lowering healthcare costs, and optimizing resource allocation within the healthcare industry.

5.3. Future Directions for Research

This study serves as a foundation for future research to further optimize deep learning applications in medical imaging. One promising direction is to incorporate additional patient data types, such as genomic or metabolic profiles, which could provide complementary information for predicting OA progression. Developing multimodal models that combine imaging data with patient-specific attributes could enhance the model's predictive power and uncover novel biomarkers for OA diagnosis. Another important area for future research is the exploration of lightweight model versions, which would facilitate deployment in resource-constrained environments, such as mobile health applications or rural healthcare centers with limited computational infrastructure. Additionally, extending the current study to include longitudinal MRI data analysis could allow for more precise monitoring of OA progression over extended periods.

5.4. Final Remarks

In summary, the advancements made in this study highlight the potential of our methodology as a sophisticated tool for the early and accurate diagnosis of knee OA, which could be highly beneficial in clinical settings. The improvements in AUC scores reflect the model's enhanced ability to differentiate between various stages of knee OA, offering a promising foundation for future research. This study paves the way for scaling this approach to larger datasets, exploring its application across other deep learning architectures, and ultimately improving patient outcomes in knee OA diagnosis and management.

References

[1] D. J. Hunter, L. March, and M. Chew, "Osteoarthritis in 2020 and beyond: a Lancet Commission," Lancet, vol. 396, no. 10264, pp. 1711–1712, 2020.

[2] M. Cross et al., "The global burden of hip and knee osteoarthritis: estimates from the global burden of disease 2010 study," Ann. Rheum. Dis., vol. 73, no. 7, pp. 1323–1330, 2014.

[3] N. Arden and M. C. Nevitt, "Osteoarthritis: epidemiology," Best Pract. Res. Clin. Rheumatol., vol. 20, no. 1, pp. 3–25, 2006.

[4] D. T. Felson, "Osteoarthritis of the knee," N. Engl. J. Med., vol. 354, no. 8, pp. 841–848, 2006.

[5] S. Glyn-Jones, A. J. Palmer, R. Agricola, A. J. Price, T. L. Vincent, H. Weinans, and A. J. Carr, "Osteoarthritis," Lancet, vol. 386, no. 9991, pp. 376–387, 2015.

[6] Vashishtha and A. Kumar Acharya, "An overview of medical imaging techniques for knee osteoarthritis disease," Biomed. Pharmacol. J., vol. 14, no. 2, pp. 903–919, 2021.

[7] H. Zhang, M. Cisse, Y. Dauphin, and P. Jitkrittum, "Mixup: Beyond the empirical mixture of experts," in International Conference on Machine Learning, 2017, pp. 788–797. https://arxiv.org/pdf/1710.09412

[8] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 7141–7151. https://arxiv.org/abs/2304.06502

[9] V. Verma et al., "Manifold Mixup: Better Representations by Interpolating Manifolds," in International Conference on Learning Representations (ICLR), 2020.

[10] X. Jin et al., "Automatic knee cartilage segmentation using 3D DenseNet-121 with attention mechanism," in Medical Imaging and Augmented Reality, 2019.

[11] L. Yu et al., "Deep learning for knee osteoarthritis diagnosis: A review," Journal of Medical Imaging and Health Informatics.

[12] X. Li et al., "DenseNet-121 for lung nodule classification in chest X-ray images," 2020.

[13] L. Yu et al., "Employing DenseNet-121 for brain tumor segmentation in MRI scans," 2021.

[14] J. Hu, L. Shen, and G. Sun, "Squeeze-and-Excitation Networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 7132–7141.

[15] L. Wang, Z. Q. Lin, and A. Wong, "COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest X-Ray Images," Scientific Reports, vol. 10, no. 1, p. 19549, 2020.

[16] X. Yu, J. Yang, N. Han, and K. L. Qin, "Deep Learning for Brain Tumor Classification," in Brain Informatics, Springer, Cham, 2020, pp. 165–174.

[17] S. Woo, J. Park, J.-Y. Lee, and I. So Kweon, "CBAM: Convolutional Block Attention Module," in Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 3–19.

[18] Z. Li, Y. Wang, and J. Yu, "Attention-Based Deep Residual Learning Network for Liver Lesion Classification in CT Scans," in IEEE Symposium on Computer-Based Medical Systems (CBMS), 2020, pp. 789–794.

[19] M. Li et al., "A Deep Learning System for Automated Skin Lesion Diagnosis in Dermoscopic Images," IEEE Journal of Biomedical and Health Informatics, vol. 25, no. 4, pp. 1239–1249, 2021.

[20] H. Zhang, M. Cisse, Y. N. Dauphin, and D. Lopez-Paz, "Mixup: Beyond Empirical Risk Minimization," in International Conference on Learning Representations, 2018. https://arxiv.org/pdf/1710.09412

[21] J. Hu, L. Shen, and G. Sun, "Squeeze-and-Excitation Networks," in Proceedings of the IEEE Conference on

Computer Vision and Pattern Recognition, 2018, pp. 7132–7141. https://openaccess.thecvf.com/content_cvpr_2018/html/Hu_Squeeze-and-Excitation_Networks_CVPR_2018_paper.html

[22] S. Woo, J. Park, J.-Y. Lee, and I. So Kweon, "CBAM: Convolutional Block Attention Module," in Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 3–19. https://openaccess.thecvf.com/content_ECCV_2018/html/Sanghyun_Woo_CBAM_Convolutional_Block_ECCV_2018_paper.html

[23] L. Wang and A. Wong, "COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest X-Ray Images," Scientific Reports, vol. 10, no. 1, p. 19549, 2020. https://www.nature.com/articles/s41598-020-76550-z

[24] X. Yu, J. Yang, N. Han, and K. L. Qin, "Deep Learning for Brain Tumor Classification," in Brain Informatics, vol. 7, 2020. https://link.springer.com/article/10.1007/s40708-020-00112-3

[25] Srivastava, A., Hinton, G. E., Krizhevsky, A., & Dahl, G. E. (2014). Improving neural networks by preventing co-adaptation of feature detectors. arXiv preprint arXiv:1207.0580. Retrieved from https://arxiv.org/abs/1207.0580