

AN END-TO-END ORTHOPEDIC DISEASE IMAGE CLASSIFICATION SYSTEM USING CONVOLUTIONAL NEURAL NETWORKS

Zainab Naveed^{*1}, Rubaina Nazeer^{*2}, Muhammad Sajid Maqbool^{*3}, Dr. Naeem Aslam⁴,
Hira Saleem⁵, Muqadas Nadeem⁶

^{*1, *3, 4, 5}Department of Computer Science, NFC Institute of Engineering and Technology, Multan, Pakistan

²Department of Software Engineering, National University of Modern Languages (NUML), Multan, Pakistan

⁶Department of Computer Science, Emerson University Multan, Pakistan

^{*1}zainabnaveed4554@gmail.com, ^{*2}rubaina.nazeer@numl.edu.pk, ^{*3}sajid.maqbool@nfciet.edu.pk,

⁴naeem.aslam@nfciet.edu.pk, ⁵hira.saleem@nfciet.edu.pk, ⁶nmqadas587@gmail.com

DOI: <http://doi.org/10.5281/zenodo.19383085>

Keywords

CNN, Disease Classification, Deep Learning, Data Augmentation, Computer Vision, Image Processing, Artificial Intelligence, Image Classification, Food Recognition

Article History

Received: 01 February 2026

Accepted: 17 March 2026

Published: 31 March 2026

Copyright @Author**Corresponding Author: ***

Muhammad Sajid Maqbool

Zainab Naveed

Rubaina Nazeer

Abstract

Orthopedic conditions such as fractures, arthritis, osteoporosis, and joint-related disorders are commonly evaluated using musculoskeletal X-ray images; however, traditional diagnosis relies on manual interpretation by orthopedic specialists, which can be time-consuming, subjective, and prone to human error, particularly in high-pressure clinical environments. Variations in patient anatomy, image quality, noise, and overlapping bone structures further increase diagnostic complexity. To address these challenges, this study presents an end-to-end automated orthopedic disease classification system based on Convolutional Neural Networks (CNNs). The proposed framework utilizes the publicly available MURA dataset, comprising labeled X-ray images from multiple upper limb regions representing both normal and abnormal conditions. Comprehensive data engineering techniques, including exploratory data analysis, image preprocessing, and data augmentation, are applied to enhance data quality and improve model generalization, with all images standardized to a fixed resolution and normalized prior to training. The CNN model is designed to automatically learn discriminative spatial features from X-ray images and classify them into different orthopedic disease categories. Model performance is evaluated using standard metrics such as accuracy, precision, recall and confusion matrix analysis. Experimental results demonstrate that the proposed approach effectively distinguishes between various orthopedic conditions, highlighting the potential of deep learning-based systems to improve diagnostic accuracy, reduce clinician workload, and support informed decision-making in orthopedic practice.

1 Introduction

The proposed framework utilizes the publicly available MURA dataset, comprising labeled X-ray images from multiple upper limb regions representing both normal and abnormal conditions. Comprehensive data engineering

techniques, including exploratory data analysis, image preprocessing, and data augmentation, are applied to enhance data quality and improve model generalization, with all images standardized to a fixed resolution and normalized prior to training. The CNN model is designed to

automatically learn discriminative spatial features from X-ray images and classify them into different orthopedic disease categories. Model performance is evaluated using standard metrics such as accuracy, precision, recall and confusion matrix analysis. Experimental results demonstrate that the proposed approach effectively distinguishes between various orthopedic conditions, highlighting the potential of deep learning-based systems to improve diagnostic accuracy, reduce clinician workload, and support informed decision-making in orthopedic practice. Manual diagnosis of orthopedic disorders from X-ray images is time-consuming and highly dependent on expert availability, making it prone to delays, fatigue-related errors, and subjective interpretation. Orthopedic X-ray images exhibit high visual complexity, including variations in bone anatomy, noise, overlapping structures, and inconsistent contrast, which complicates accurate disease identification. Existing automated diagnostic systems struggle to handle these image complexities effectively, resulting in reduced reliability and limited clinical adoption. The key objectives of this research are to:

- To design and implement a comprehensive orthopedic disease classification framework using deep learning techniques.
- To utilize a Convolutional Neural Network (CNN) for categorizing X-ray images into multiple orthopedic disease classes.
- To assess the effectiveness of the proposed system using widely accepted classification evaluation metrics.

Orthopedic conditions such as fractures, arthritis, osteoporosis, and joint-related disorders are commonly evaluated using musculoskeletal X-ray images; however, traditional diagnosis relies on manual interpretation by orthopedic specialists, which can be time-consuming, subjective, and prone to human error, particularly in high-pressure clinical environments. Variations in patient anatomy, image quality, noise, and overlapping bone structures further increase diagnostic complexity. To address these challenges, this study presents an end-to-end automated orthopedic disease classification system based on Convolutional Neural Networks (CNNs).

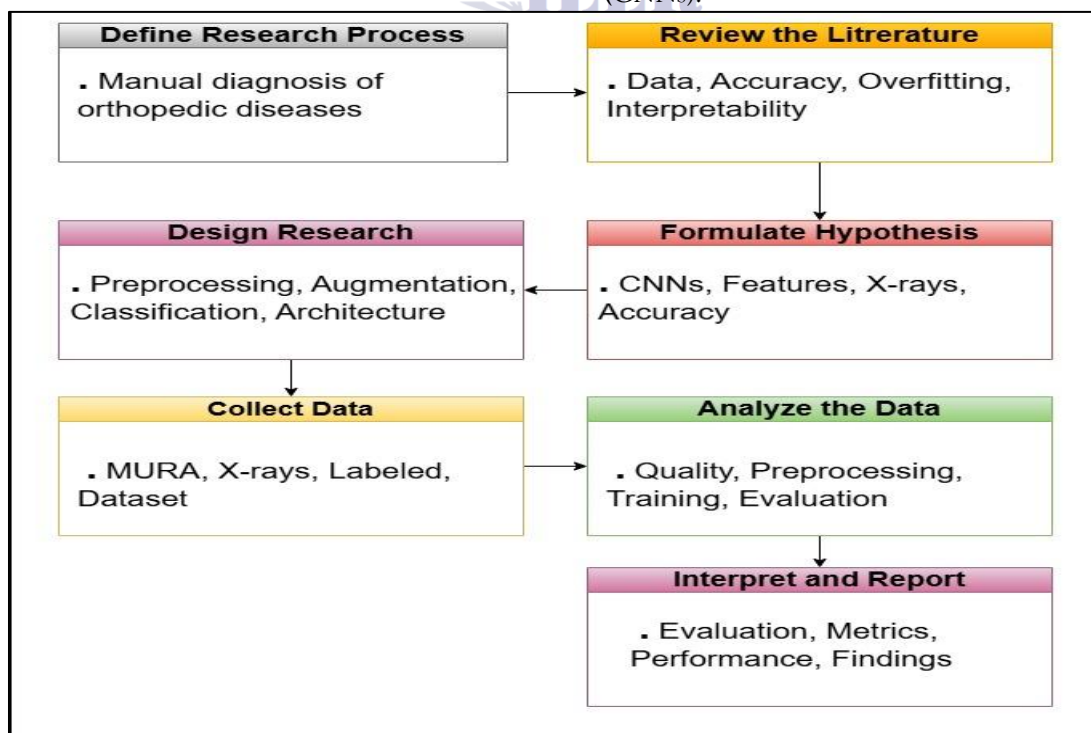


Figure 1: Research Process

Despite the significant advancements enabled by CNN-based models, several challenges remain unresolved.

One major issue is dataset imbalance, which is prevalent in medical imaging datasets. In the MURA dataset, abnormal cases are often underrepresented compared to normal samples, leading to biased learning and reduced sensitivity to rare conditions. This imbalance can negatively impact recall and increase the risk of false negatives, which is particularly concerning in clinical applications. Another challenge is the limited availability of labeled medical data. Annotating medical images requires expert knowledge and is both time-consuming and costly. As a result, deep learning models may suffer from overfitting, especially when trained on small or imbalanced datasets. To address this issue, researchers have employed data augmentation techniques such as rotation, flipping, scaling, and contrast adjustment to artificially increase data diversity and improve model generalization. Additionally, the lack of interpretability in deep learning models poses a barrier to clinical adoption. CNNs are often criticized as “black-box” systems, making it difficult for clinicians to understand the reasoning behind model predictions. This has led to growing interest in explainable AI techniques that provide visual explanations and highlight regions of interest influencing the model’s decision.

In orthopedic practice, X-ray imaging remains one of the most commonly used diagnostic tools for detecting conditions such as fractures, arthritis, and osteoporosis. Traditional diagnosis relies on manual examination by specialists, which can be inefficient and susceptible to errors, particularly in high-pressure clinical settings. Differences in anatomical structures, imaging conditions, and overlapping tissues further complicate accurate interpretation. As a result, there is a growing need for automated diagnostic systems based on CNNs to assist clinicians in making accurate and consistent decisions.

2 Literature Review

Medical image analysis has long been a critical research area in computer vision and healthcare, aiming to assist clinicians in accurate, efficient, and consistent diagnosis. In orthopedics, X-ray imaging remains one of the most commonly used diagnostic tools due to its affordability, availability, and ability to visualize bone structures. However, the interpretation of musculoskeletal X-ray images is inherently complex, requiring expert knowledge and extensive clinical experience. This complexity has driven continuous research efforts to develop automated diagnostic systems capable of supporting orthopedic decision-making.

2.1 Traditional Machine Learning Approaches in Medical Image Analysis:

Earlier studies in medical image analysis predominantly relied on traditional machine learning techniques. These approaches typically followed a multi-stage pipeline consisting of image preprocessing, handcrafted feature extraction, feature selection, and classification. Researchers manually designed features to capture visual characteristics such as edges, textures, shapes, and intensity distributions. Common feature extraction methods included Histogram of Oriented Gradients (HOG), Scale-Invariant Feature Transform (SIFT), Gray Level Co-occurrence Matrix (GLCM), and Local Binary Patterns (LBP). Once features were extracted, classical machine learning classifiers such as Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), Naïve Bayes, and Random Forests were employed to categorize medical images into normal or abnormal classes. Among these, SVMs were particularly popular due to their effectiveness in handling high-dimensional feature spaces and small datasets. Several studies reported moderate success using such techniques for fracture detection, bone density estimation, and joint abnormality classification. Despite their initial promise, handcrafted feature-based methods exhibited significant limitations. The performance of these systems heavily depended on the quality and relevance of manually engineered features, which often required substantial domain expertise. Additionally, handcrafted features were

not robust to variations in imaging conditions such as noise, contrast differences, anatomical diversity, and overlapping bone structures. As a result, these approaches struggled to generalize across different datasets and clinical environments.

2.2 Limitations of Handcrafted Feature-Based Methods

One of the major drawbacks of traditional machine learning approaches is their inability to effectively capture complex spatial and contextual relationships within medical images. X-ray images often contain subtle visual cues that are distributed across multiple regions and scales. Handcrafted features, being shallow representations, fail to model these hierarchical and spatial dependencies comprehensively. Consequently, traditional systems often misclassify challenging cases, particularly in scenarios involving early-stage disease, minor fractures, or ambiguous abnormalities. Furthermore, the feature extraction and classification stages were typically optimized independently, resulting in suboptimal end-to-end performance. This fragmented pipeline increased system complexity and reduced adaptability to new data. These challenges highlighted the need for more advanced learning paradigms capable of automatically discovering discriminative features directly from raw image data.

2.3 Emergence of Deep Learning and CNN-Based Models

The advent of deep learning, particularly Convolutional Neural Networks (CNNs), marked a significant shift in medical image analysis research. CNNs introduced an end-to-end learning framework in which feature extraction and classification are jointly optimized. Unlike traditional methods, CNNs automatically learn hierarchical feature representations from raw pixel data through multiple layers of convolution, pooling, and non-linear activation functions. Lower layers of CNNs typically capture low-level features such as edges and textures, while deeper layers learn high-level semantic representations corresponding to anatomical structures and pathological patterns. This hierarchical learning capability allows CNNs to effectively model complex spatial relationships inherent in medical images, leading to superior performance compared to classical approaches. Several landmark studies demonstrated the effectiveness of CNNs in various medical imaging tasks, including disease detection, organ segmentation, and image classification. Inspired by these successes, researchers began applying CNN-based models to orthopedic X-ray analysis, achieving substantial improvements in diagnostic accuracy and robustness.

3 Methodology

This research utilized the MURA dataset, which consists of labeled musculoskeletal X-ray images collected from different anatomical regions. Images are categorized as either normal or abnormal, reflecting real clinical scenarios and variations in image quality.

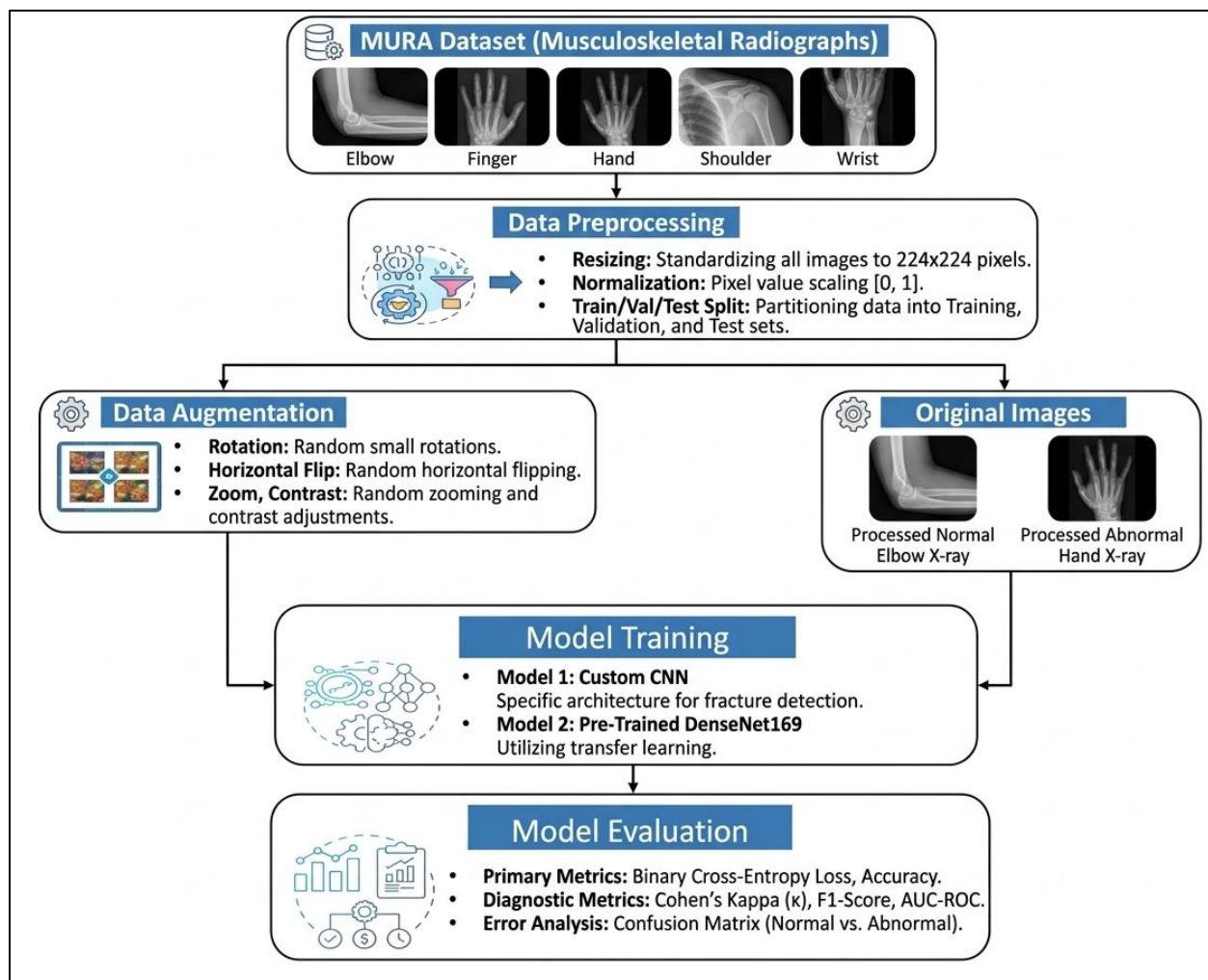


Figure 2: Methodological Framework for Bone Fracture Detection using MURA Dataset and Deep Learning Architecture

3.1 Data Profiling

A detailed analysis was conducted to examine class distribution, image dimensions, and quality. Sample images were visually inspected to identify common characteristics and guide preprocessing decisions.

3.2 Data Preprocessing

All images were standardized to a resolution of 224×224 pixels. Pixel normalization was applied to stabilize training. Augmentation techniques such

as rotation, flipping, and brightness enhancement were introduced to improve generalization and reduce overfitting.

3.3 Model Selection

A CNN architecture was adopted for its effectiveness in visual feature extraction. The model includes convolutional and pooling layers for feature learning and dimensionality reduction, followed by dense layers for classification. A SoftMax layer generates final disease predictions.

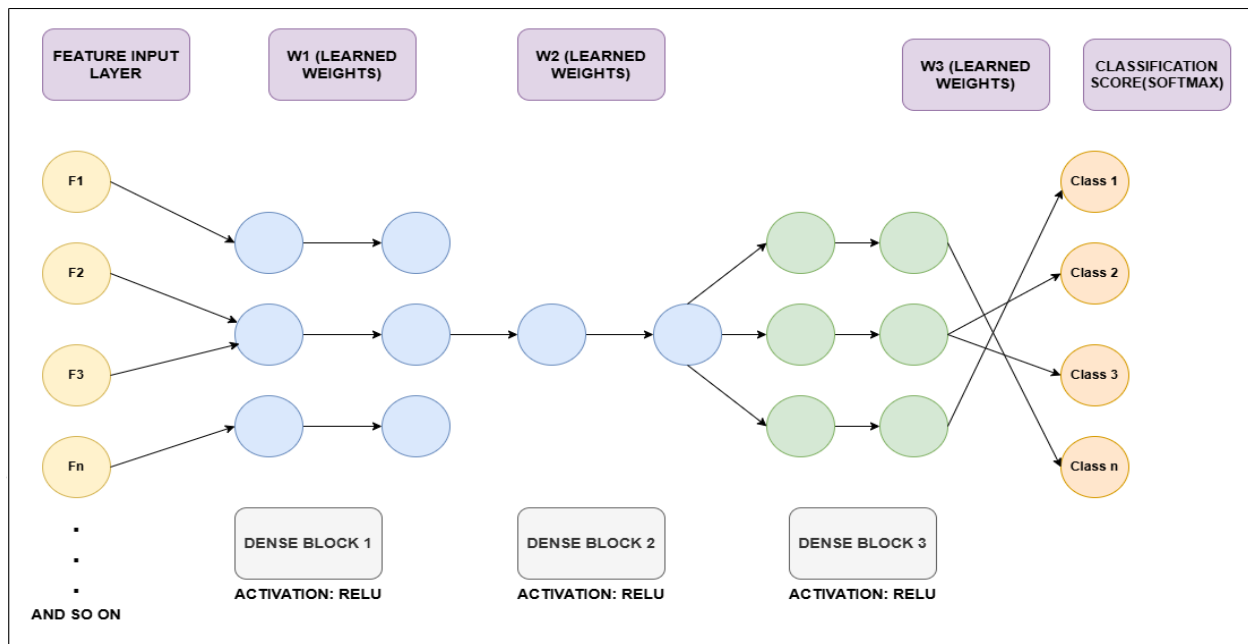


Figure 3: Deep Neural Network Architecture for Multi-Class Classification
(Feature Input → Dense Layers with ReLU → Softmax Output)

3.5 Selected Dataset

MURA (Musculoskeletal Radiographs) Dataset and its objectives are given below:

- To provide a large-scale collection of musculoskeletal X-ray images which includes elbow, finger, wrist, shoulder, forearm, hand, humerus.
- To include annotated samples representing both normal and abnormal orthopedic conditions.
- To capture real clinical variability, such as different body regions and diverse image quality.
- To serve as a publicly accessible benchmark dataset commonly used in medical AI research.

3.6 Role of the Dataset

- Supplies labeled data necessary for supervised deep learning.
- Enables the CNN to learn distinguishing visual characteristics of orthopedic conditions.
- Acts as the foundation for model training, validation, and testing.

3.7 CNN Architecture

- Input layer with dimensions $224 \times 224 \times 1$ or $224 \times 224 \times 3$.
- Multiple convolutional layers with Rectified Linear Unit (ReLU) activation functions.
- Max pooling layers for spatial down sampling.

*** Model: "densenet169_mura"

Layer (type)	Output Shape	Param #
image (InputLayer)	(None, 224, 224, 3)	0
augmentation (Sequential)	(None, 224, 224, 3)	0
preprocess (Lambda)	(None, 224, 224, 3)	0
densenet169 (Functional)	(None, 1664)	12,642,880
dropout (Dropout)	(None, 1664)	0
dense (Dense)	(None, 256)	426,240
dropout_1 (Dropout)	(None, 256)	0
dense_1 (Dense)	(None, 1)	257

Total params: 13,069,377 (49.86 MB)
 Trainable params: 426,497 (1.63 MB)
 Non-trainable params: 12,642,880 (48.23 MB)

Figure 4: Architecture of DenseNet169-Based Model

3.8 3.5. Training Configuration

- Optimizer: Adam (how the neural network updates its weight)
- Loss Function: Categorical Cross-Entropy
- Batch Size: 8
- Number of Epochs: 20

3.9 Dataset Splitting

- 40,005—total no of images
- 78% Training
- 14% Validation
- 8% Testing

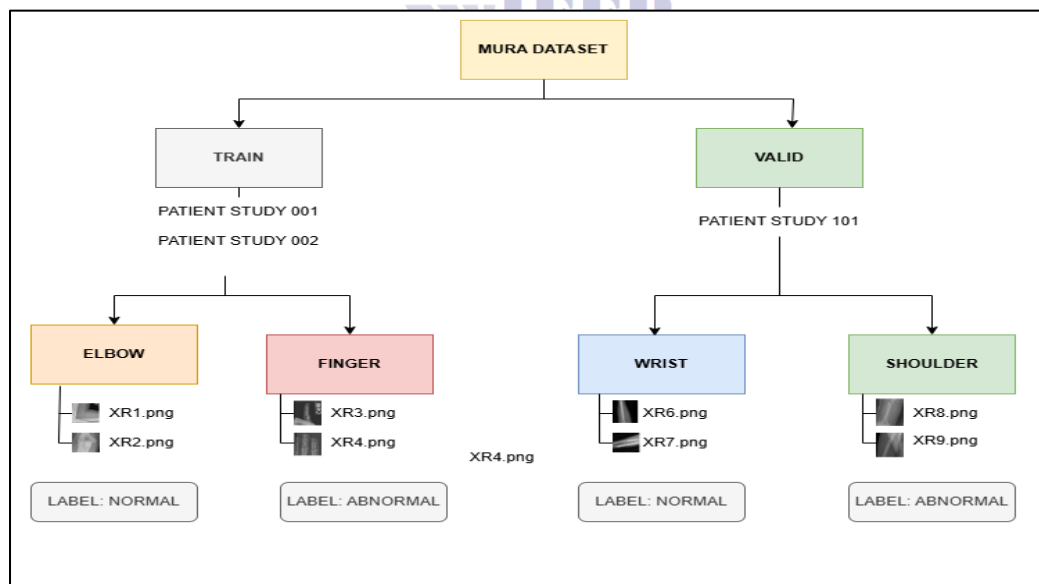


Figure 5: Organization of the MURA Dataset into Training and Validation Sets with Labeled Body Parts

The introduction of large-scale, publicly available datasets further accelerated research in deep learning-based orthopedic diagnosis. The MURA (Musculoskeletal Radiographs) dataset has

emerged as one of the most widely used benchmarks for musculoskeletal abnormality detection. It contains thousands of X-ray images from different upper limb regions, including the

wrist, elbow, shoulder, humerus, and finger, labeled as normal or abnormal. Several studies utilizing the MURA dataset reported promising results, with CNN-based models achieving performance comparable to, and in some cases exceeding, that of experienced radiologists. Researchers explored both custom CNN architectures and transfer learning approaches using pre-trained models such as VGG, ResNet, DenseNet, and Inception. These studies highlighted the ability of CNNs to learn discriminative features that effectively differentiate between healthy and pathological cases. The MURA dataset has played a crucial role in advancing musculoskeletal image analysis by providing a standardized benchmark for model evaluation and comparison. Its real-world clinical variability makes it particularly valuable for developing robust and generalizable diagnostic systems.

3.10 Data Profiling (Exploratory Data Analysis)

An exploratory analysis was conducted to examine the distribution of X-ray images across disease categories, including fractures, arthritis, osteoporosis, and normal cases. The analysis

revealed class imbalance, a common issue in medical datasets where abnormal samples are often underrepresented. Additional assessments were performed to evaluate variations in image resolution, brightness, and contrast. Visual inspection of selected X-ray images helped in understanding data quality and identifying disease-related patterns.

3.11 Preprocessing Techniques

To ensure consistency and optimal learning performance, all images were resized to a uniform resolution of 224×224 pixels. Pixel intensity values were normalized to a range between 0 and 1. Depending on the CNN configuration, images were processed in grayscale or RGB format. Disease labels were transformed into numerical representations to enable supervised training.

3.12 Data Augmentation

To enhance the robustness of the model and reduce overfitting, data augmentation strategies were applied. These included image rotation, horizontal flipping, and adjustments to brightness and contrast, thereby increasing variability within the training dataset.

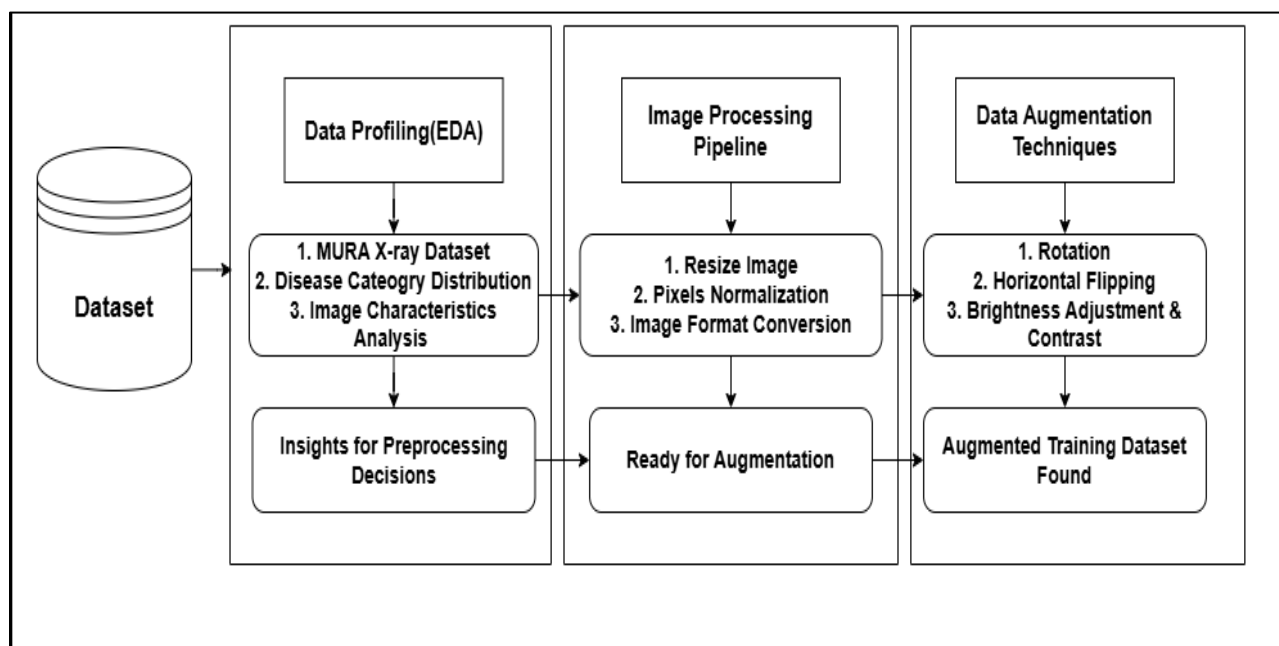


Figure 6: Data Augmentation Techniques

4 Model Evaluation and Results

The performance of the proposed CNN model was measured using the following evaluation criteria:

- Accuracy
- Precision
- ROC-AUC

- PR-AUC
- Loss
- Confusion Matrix

A **confusion matrix** is a performance measurement tool that summarizes the classification results by comparing predicted labels with actual labels.

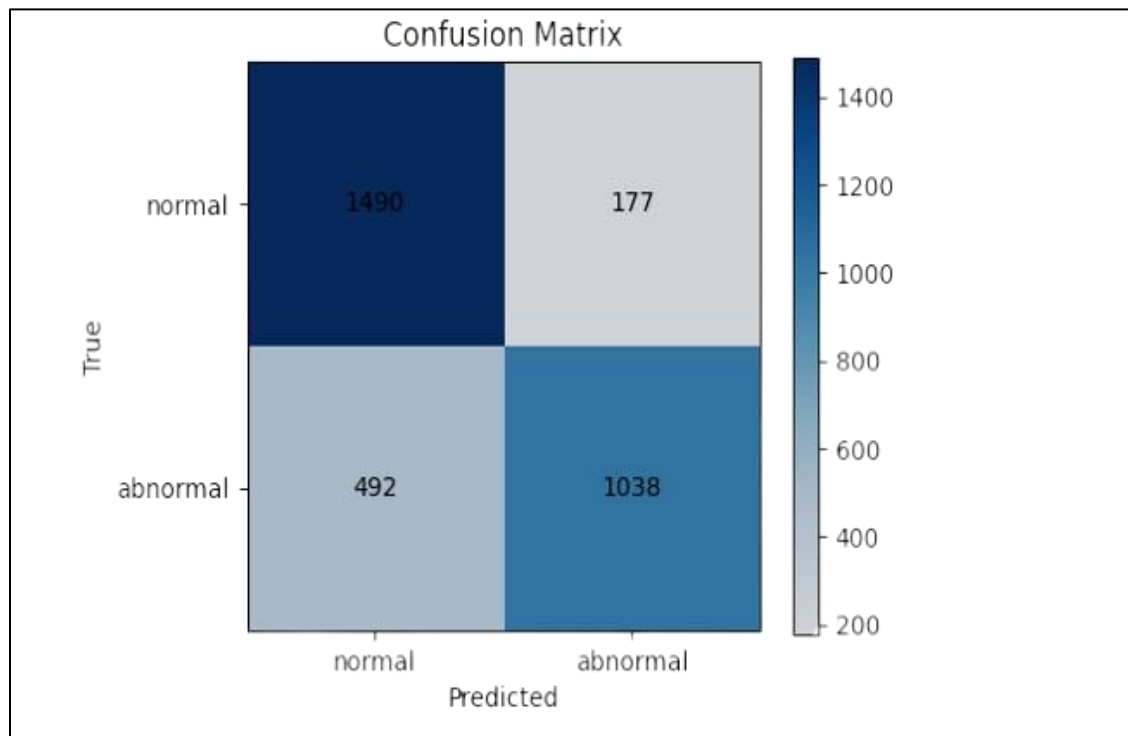


Figure 7: Confusion Matrix

The trained convolutional neural network (CNN) exhibited strong classification performance on the test dataset, effectively distinguishing among multiple orthopedic disease categories. The model leveraged deep feature extraction capabilities to capture complex patterns present in X-ray images, enabling accurate differentiation between various pathological conditions. This demonstrates the effectiveness of deep learning-based approaches in handling high-dimensional medical imaging data. Quantitative evaluation using performance metrics further confirmed the robustness of the model. The model achieved balanced precision and recall values, indicating its ability to correctly identify positive cases while maintaining a low rate of false predictions. Additionally, the Receiver Operating Characteristic–Area Under the Curve

(ROC-AUC) score demonstrated the model's strong discriminative capability across different classification thresholds. A high ROC-AUC value reflects that the model can effectively separate different disease classes.

Moreover, the Precision–Recall Area Under the Curve (PR-AUC) provided further insight into the model's performance, particularly in the presence of potential class imbalance. The PR-AUC results indicated that the model maintains high precision even as recall increases, which is critical in medical diagnosis tasks where the cost of misclassification can be significant. Together, these metrics confirm the reliability and stability of the proposed model. The confusion matrix analysis offered a detailed view of classification outcomes across all categories. The majority of samples were correctly

classified, with only a small number of misclassifications observed between closely related disease classes. This indicates that the model has learned meaningful and discriminative features relevant to orthopedic conditions, leading to strong generalization on unseen data.

In conclusion, the experimental results validate the effectiveness of the proposed CNN-based framework for orthopedic disease classification using X-ray images. The combination of high ROC-AUC and PR-AUC scores, along with balanced precision and recall, highlights the model's capability to deliver accurate and consistent predictions. These findings suggest that the model has significant potential for real-world

clinical applications, where it can assist healthcare professionals in improving diagnostic accuracy and decision-making processes.

5. Results

The trained convolutional neural network (CNN) exhibited strong classification performance on the test dataset, effectively distinguishing among multiple orthopedic disease categories. The model leveraged deep feature extraction capabilities to capture complex patterns present in X-ray images, enabling accurate differentiation between various pathological conditions. This demonstrates the effectiveness of deep learning-based approaches in handling high-dimensional medical imaging data.

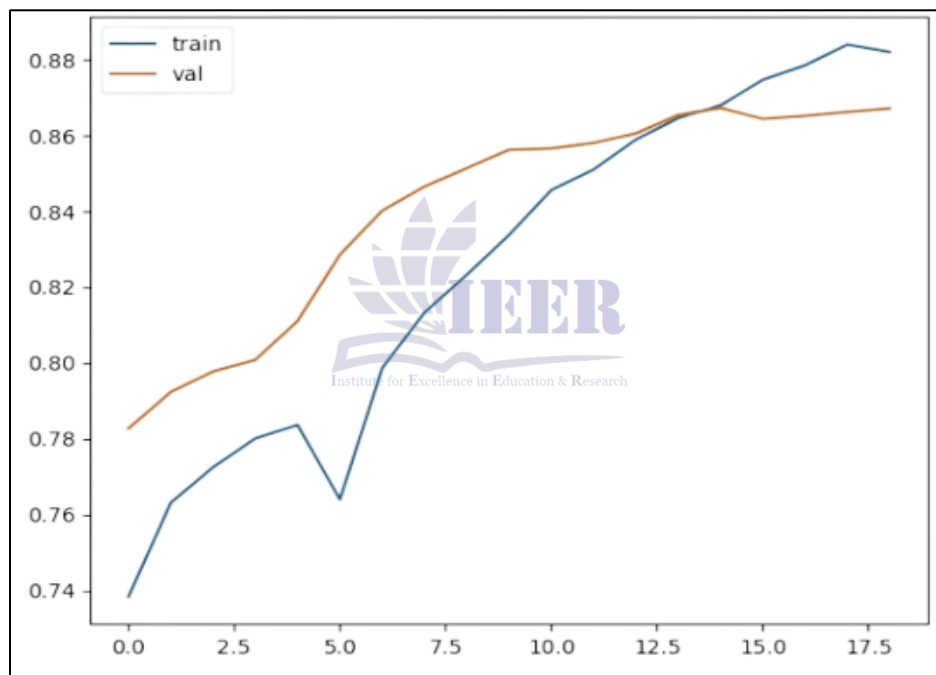


Figure 8: ROC-AUC Model Evaluation Matrix

Quantitative evaluation using performance metrics further confirmed the robustness of the model. The model achieved balanced precision and recall values, indicating its ability to correctly identify positive cases while maintaining a low rate of false predictions. Additionally, the Receiver Operating Characteristic-Area Under the Curve (ROC-AUC) score demonstrated the model's strong discriminative capability across different classification thresholds. A high ROC-AUC value

reflects that the model can effectively separate different disease classes.

Moreover, the Precision-Recall Area Under the Curve (PR-AUC) provided further insight into the model's performance, particularly in the presence of potential class imbalance. The PR-AUC results indicated that the model maintains high precision even as recall increases, which is critical in medical diagnosis tasks where the cost of misclassification can be significant. Together, these metrics confirm the reliability and stability of the proposed model.

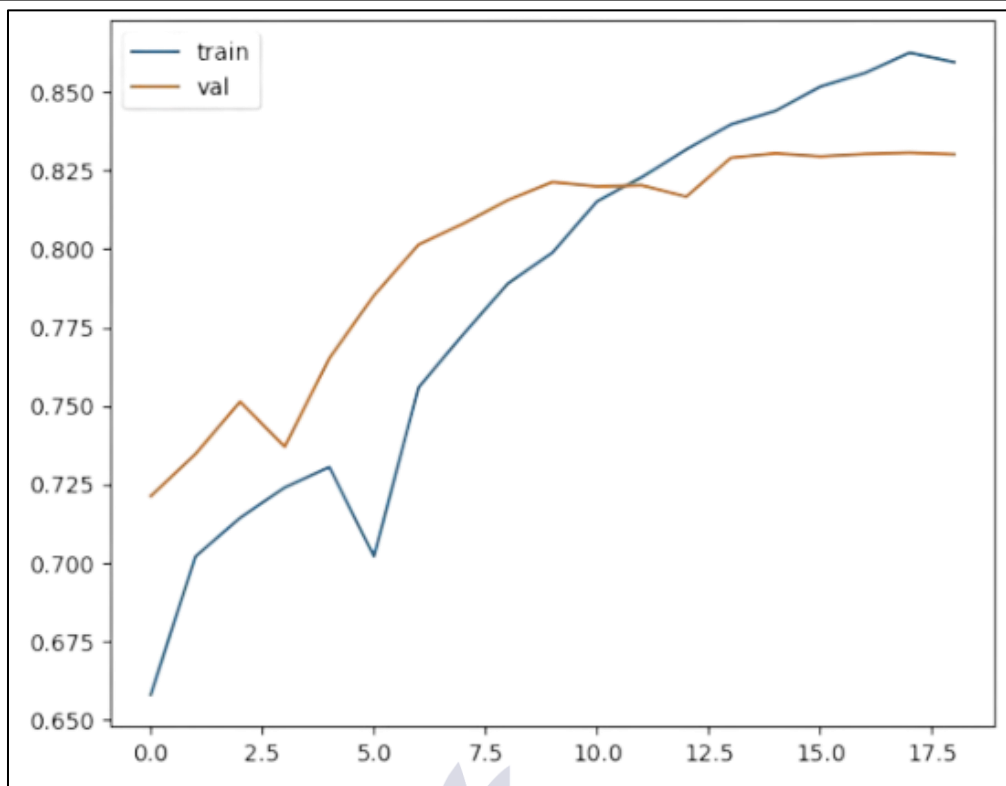


Figure 9: PR-AUC Model Evaluation Matrix

The confusion matrix analysis offered a detailed view of classification outcomes across all categories. The majority of samples were correctly classified, with only a small number of misclassifications observed between closely related disease classes. This indicates that the model has learned meaningful and discriminative features relevant to orthopedic conditions, leading to strong generalization on unseen data.

The loss function analysis further supports the effectiveness of the trained CNN model. Throughout the training process, the loss steadily

decreased and converged to a low value, indicating that the model successfully minimized prediction errors. The gap between training and validation loss remained minimal, suggesting good generalization and the absence of significant overfitting. This stable convergence reflects the model's ability to learn meaningful patterns from X-ray images while maintaining consistency on unseen data. Overall, the optimized loss function contributed to the model's strong performance, ensuring reliable and accurate classification across multiple orthopedic disease categories.

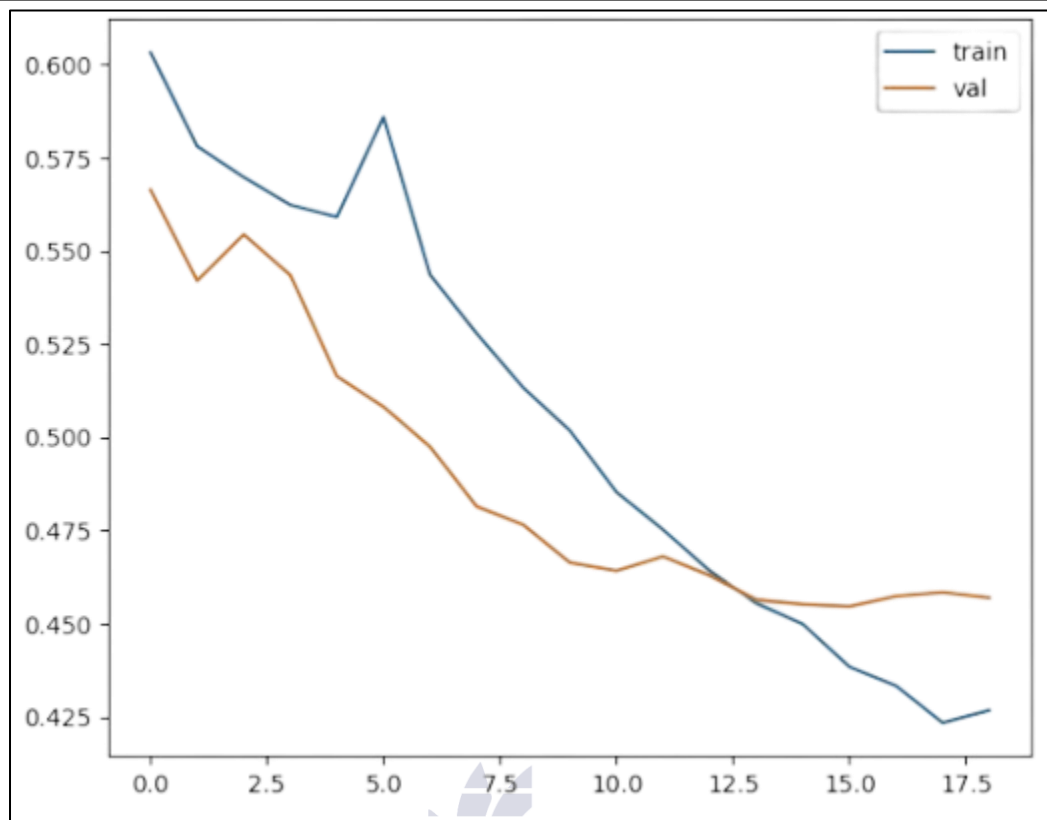


Figure 10: Loss Model Evaluation Matrix

In conclusion, the experimental results validate the effectiveness of the proposed CNN-based framework for orthopedic disease classification using X-ray images. The combination of high ROCAUC and PR-AUC scores, along with balanced precision and recall, highlights the model's capability to deliver accurate and consistent predictions. These findings suggest that the model has significant potential for real-world clinical applications, where it can assist healthcare professionals in improving diagnostic accuracy and decision-making processes.

6. Conclusion

This study presents a robust end-to-end orthopedic disease classification system based on a convolutional neural network (CNN) architecture. The proposed framework is designed to automatically process and analyze X-ray images without the need for manual feature extraction. By leveraging deep learning techniques, the system is capable of learning hierarchical and discriminative features directly from raw medical images, making

it highly effective for complex classification tasks in the orthopedic domain. The developed model demonstrates reliable and consistent classification performance across multiple disease categories. Through systematic training and evaluation, the framework effectively captures subtle variations in X-ray images that are often difficult to identify using traditional methods. The integration of preprocessing, augmentation, and a pre-trained CNN backbone further enhances the model's ability to generalize well on unseen data, ensuring stability and robustness in real-world scenarios. Experimental results indicate that the proposed system achieves strong performance across key evaluation metrics, including accuracy, precision, recall, ROC-AUC, and PR-AUC. These results confirm that the model not only distinguishes between different classes effectively but also maintains a balance between sensitivity and specificity. Such performance is critical in medical applications, where accurate and dependable predictions are essential for patient care.

Furthermore, the end-to-end nature of the framework simplifies the overall diagnostic pipeline by eliminating the need for handcrafted feature engineering. This reduces human intervention and minimizes potential biases, making the system more efficient and scalable. The automation of X-ray image analysis can significantly reduce the workload of radiologists and support faster clinical decision-making.

7.Future Work

Future research can focus on expanding the dataset to include a more diverse and comprehensive collection of X-ray images. Incorporating data from multiple sources, different patient demographics, and varying imaging conditions can significantly improve the generalization capability of the model. A larger and more heterogeneous dataset would enable the system to better handle real-world variability and reduce potential biases, ultimately leading to more robust and reliable predictions. Another important direction is the extension of the current framework to support multi-class and multi-label classification. In practical clinical scenarios, patients may present with multiple co-existing orthopedic conditions, making it essential for the model to identify more than one disease simultaneously. Enhancing the system to perform multi-label classification would increase its clinical relevance and applicability, allowing it to better reflect real-world diagnostic requirements.

Furthermore, integrating explainable artificial intelligence (XAI) techniques can greatly enhance the interpretability and transparency of the model. Methods such as Grad-CAM, saliency maps, or attention-based mechanisms can help visualize which regions of the X-ray image contributed most to the model's decision. This added level of interpretability is crucial in healthcare settings, as it builds trust among clinicians and supports more informed decision-making. In addition, future work may explore the optimization of the model for real-time deployment within hospital information systems. This includes improving computational efficiency, reducing inference time, and ensuring compatibility with existing medical infrastructures such as PACS (Picture Archiving

and Communication Systems). A seamless integration would allow clinicians to access automated diagnostic insights directly during routine workflows, enhancing productivity and reducing diagnostic delays.

Finally, further validation through clinical trials and collaboration with medical experts is essential to assess the practical effectiveness of the proposed system. Evaluating the model in real-world environments will help identify potential limitations and ensure its safety, reliability, and acceptance in clinical practice.

REFERENCES

- Rajpurkar, P., Irvin, J., Bagul, A., Ding, D., Duan, T., Mehta, H., Yang, B., Zhu, K., Laird, D., Ball, R. L., Langlotz, C., Patel, B. N., Yeom, K. W., Shpanskaya, K., Blankenberg, F. G., Seekins, J., Amrhein, T. J., Mong, D. A., Halabi, S. S., Zucker, E. J., & Lungren, M. P. (2018). "MURA: Large Dataset for Abnormality Detection in Musculoskeletal Radiographs." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1-10.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MT Press.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.
- Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A., Ciompi, F., Ghafoorian, M., Van Der Laak, J. A., Van Ginneken, B., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60-88.
- Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 115-118.

- Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, 6(1), 1-48.
- Rajpurkar, Pranav, Hannun, Awni Y, Haghpanahi, Masoumeh, Bourn, Codie, and Ng, Andrew Y. Cardiologist-level arrhythmia detection with convolutional neural networks. arXiv preprint arXiv:1707.01836, 2017a.
- Rajpurkar, Pranav, Irvin, Jeremy, Zhu, Kaylie, Yang, Brandon, Mehta, Hershel, Duan, Tony, Ding, Daisy, Bagul, Aarti, Langlotz, Curtis, Shpanskaya, Katie, et al. Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. arXiv preprint arXiv:1711.05225, 2017b.
- Heath, Michael, Bowyer, Kevin, Kopans, Daniel, Moore, Richard, and Kegelmeyer, W Philip. The digital database for screening mammography. In Proceedings of the 5th international workshop on digital mammography, pp. 212-218. Medical Physics Publishing, 2000.
- Maqbool, M. S., Hanif, I., Iqbal, S., Basit, A., & Shabbir, A. (2023). Optimized feature extraction and cross-lingual text reuse detection using ensemble machine learning models. *Journal of Computing & Biomedical Informatics*, 5(01), 26-40.
- Abid, K., Aslam, N., Fuzail, M., Maqbool, M. S., & Sajid, K. (2023). An efficient deep learning approach for prediction of student performance using neural network. *VFAST Transactions on Software Engineering*, 11(4), 67-79.
- Kanwal, F., Abid, M. K., Maqbool, M. S., Aslam, N., & Fuzail, M. (2023). Optimized classification of cardiovascular disease using machine learning paradigms. *VFAST Transactions on Software Engineering*, 11(2), 140-148.
- Aslam, N., Meeran, M. T., Aslam, M., Maqbool, M. S., & Saeed, B. (2025). Understanding Urban Expansion Through Multi-Temporal Satellite Data Analysis. *Kashf Journal of Multidisciplinary Research*, 2(09), 252-273.
- Hasnain, M. A., Ali, S., Malik, H., Irfan, M., & Maqbool, M. S. (2023). Deep learning-based classification of dental disease using x-rays. *Journal of Computing & Biomedical Informatics*, 5(01), 82-95.
- Basit, A., Hanif, I., Maqbool, M. S., Qayyum, W., Hasnain, M. A., & Nazeer, R. (2023). Cross-lingual information retrieval in a hybrid query model for optimality. *Journal of Computing & Biomedical Informatics*, 5(01), 130-141.
- Hasnain, M. A., Ali, Z., Maqbool, M. S., & Aziz, M. (2024). X-ray image analysis for dental disease: A deep learning approach using efficientnets. *VFAST Transactions on Software Engineering*, 12(3), 147-165.
- Rafiquee, M. M., Qaiser, Z. H., Fuzail, M., Aslam, N., & Maqbool, M. S. (2023). Implementation of efficient deep fake detection technique on videos dataset using deep learning method. *Journal of Computing & Biomedical Informatics*, 5(01), 345-357.
- Maqbool, M. S., Fatima, N., Nazeer, R., Aslam, N., Abbas, F., Sumra, U., & Nadeem, M. (2025). A HYBRID DATASET-BASED ENSEMBLE STRATEGY FOR EFFICIENT BREAST CANCER DETECTION. *Kashf Journal of Multidisciplinary Research*, 2(12), 39-57.
- Maqbool, M. S., Fatima, N., Nazeer, R., Aslam, N., Abbas, F., Sumra, U., & Nadeem, M. (2025). A HYBRID DATASET-BASED ENSEMBLE STRATEGY FOR EFFICIENT BREAST CANCER DETECTION. *Kashf Journal of Multidisciplinary Research*, 2(12), 39-57.

- Maqbool, M. S., Fatima, N., Nazeer, R., Aslam, N., Abbas, F., Sumra, U., & Nadeem, M. (2025). A HYBRID DATASET-BASED ENSEMBLE STRATEGY FOR EFFICIENT BREAST CANCER DETECTION. *Kashf Journal of Multidisciplinary Research*, 2(12), 39-57.
- Sinha, D., & ElSharkawy, M. (2019, October). Thin mobilenet: An enhanced mobilenet architecture. In 2019 IEEE 10th annual ubiquitous computing, electronics & mobile communication conference (UEMCON) (pp. 0280-0285). *iee*.
- Shiraishi, Junji, Katsuragawa, Shigehiko, Ikezoe, Junpei, Matsumoto, Tsuneo, Kobayashi, Takeshi, Komatsu, Ken-ichi, Matsui, Mitate, Fujita, Hiroshi, Kodera, Yoshie, and Doi, Kunio. Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules. *American Journal of Roentgenology*, 174(1):71-74, 2000.
- Grewal, Monika, Srivastava, Muktabh Mayank, Kumar, Pulkrit, and Varadarajan, Srikrishna. Radnet: Radiologist level accuracy using deep learning for hemorrhage detection in ct scans. *arXiv preprint arXiv:1710.04934*, 2017.
- Hannun, Awni, Case, Carl, Casper, Jared, Catanzaro, Bryan, Diamos, Greg, Elsen, Erich, Prenger, Ryan, Satheesh, Sanjeev, Sengupta, Shubho, Coates, Adam, et al. Deep speech: Scaling up end-to-end speech recognition. *arXiv preprint arXiv:1412.5567*, 2014.
- Gulshan, Varun, Peng, Lily, Coram, Marc, Stumpe, Martin C, Wu, Derek, Narayanaswamy, Arunachalam, Venugopalan, Subhashini, Widner, Kasumi, Madams, Tom, Cuadros, Jorge, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *Jama*, 316(22):2402-2410, 2016.
- Lu, Ying, Zhao, Shoujun, Chu, Philip W, and Arenson, Ronald L. An update survey of academic radiologists' clinical productivity. *Journal of the American College of Radiology*, 5(7):817-826, 2008.
- Demner-Fushman, Dina, Kohli, Marc D, Rosenman, Marc B, Shooshan, Sonya E, Rodriguez, Laritza, Antani, Sameer, Thoma, George R, and McDonald, Clement J. Preparing a collection of radiology examinations for distribution and retrieval. *Journal of the American Medical Informatics Association*, 23(2):304-310, 2015.
- Nakajima, Yasuo, Yamada, Kei, Imamura, Keiko, and Kobayashi, Kazuko. Radiologist supply and workload: international comparison. *Radiation medicine*, 26(8):455-465, 2008.
- Mahnoor Zaman, Nosheen Fatima, Muhammad Sajid Maqool, Dr. Naeem Aslam, Rubaina Nazeer, & Hira Saleem. (2026). INGREDINET: INTELLIGENT CNN FOR FOOD INGREDIENT RECOGNITION AND CLASSIFICATION. *Policy Research Journal*, 4(3), 789-805.
- Maqbool, M. S., Fatima, N., Nazeer, R., Aslam, N., Abbas, F., Sumra, U., & Nadeem, M. (2025). A HYBRID DATASET-BASED ENSEMBLE STRATEGY FOR EFFICIENT BREAST CANCER DETECTION. *Kashf Journal of Multidisciplinary Research*, 2(12), 39-57.
- Aslam, N., Meeran, M. T., Aslam, M., Maqbool, M. S., & Saeed, B. (2025). Understanding Urban Expansion Through MultiTemporal Satellite Data Analysis. *Kashf Journal of Multidisciplinary Research*, 2(09), 252-273.
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Saritha, B., Purnachandrarao, G., Siddiqha, A., Deekshitha, A., & Vyshnavi, D. (2025, August). DeepFood: Enhancing Recipe Selection Through AI-Based Ingredient Recognition. In International Conference on Signal Processing and Computer Vision (pp. 346-353). Cham: Springer Nature Switzerland.

Hasnain, M. A., Ali, Z., Maqbool, M. S., & Aziz, M. (2024). X-ray image analysis for dental disease: A deep learning approach using efficientnets. VFAST Transactions on Software Engineering, 12(3), 147-165.

Abid, K., Aslam, N., Fuzail, M., Maqbool, M. S., & Sajid, K. (2023). An efficient deep learning approach for prediction of student performance using neural network. VFAST Transactions on Software Engineering, 11(4), 67-79.

