AMERICAN JOURNAL OF MANAGEMENT AND IOT MEDICAL COMPUTING
Peer Reviewed, Referred & Indexed Journal
E-ISSN: 3069-0110
Vol.5, No.2(2026)
www.ajmimc.com

DEIT-Based Feature Extraction with Ensemble Machine Learning for Accurate Bone Fracture Detection in X-Ray Images

Mohammad Sharib[1], B. Sharada[2*], Kurra Sanjay[1], Mohammad Arif Pasha[1], Gogula Akshaya[1], Chintala Shivaram[1]

[1]UG Student, [2]Assistant Professor, [1,2]Department of Computer Science and Engineering (AI&ML)

[1,2]Vaagdevi Engineering College, Bollikunta, Warangal, 506005, Telangana, India

*Correspondence: B. Sharada (sharadabura11@gmail.com)

Abstract

Bone fracture detection using X-ray imaging has traditionally depended on expert radiological assessment, which is time-intensive and subject to inter-observer variability. With the increasing need for rapid and accurate diagnosis in clinical settings, automated computer-aided diagnostic systems have gained significant attention. Although advanced imaging modalities such as CT scan and MRI provide high diagnostic accuracy, they are costly and less accessible, making them unsuitable for large-scale screening. Consequently, X-ray-based intelligent systems have emerged as a scalable and cost-effective alternative. In this study, a hybrid framework is proposed by integrating transformer-based feature extraction with ensemble machine learning techniques. A pre-trained Data-efficient Image Transformer (DeiT) is employed as a fixed feature extractor to convert raw X-ray images into high-dimensional feature representations. These features are then used to train multiple classifiers, including Ridge Classifier, Passive Aggressive Classifier, and Nearest Centroid Classifier, along with a proposed Fast Interpretable Greedy-tree Sums (FIGS) model for enhanced decision-making. To improve transparency, an Explainable Artificial Intelligence (XAI) module is incorporated as a preliminary validation layer, ensuring that the input is a bone X-ray and providing contextual outputs such as fracture location, type, and severity. This enhances trust and interpretability in clinical decision-making. Performance evaluation using accuracy, precision, recall, and F1-score demonstrates the effectiveness of the proposed approach. The system is further deployed via a Telegram Bot interface, enabling real-time analysis and offering a practical, accessible solution for assisting medical professionals in fracture diagnosis.

Keywords: bone fracture, deep learning, ensemble learning, explainable AI, image classification, X-ray imaging

1. Introduction

Shoulder fractures were commonly caused by conditions such as dislocations, contact sports injuries, and motor vehicle accidents. The shoulder structure primarily consisted of three bones: the humerus (upper arm bone), scapula (shoulder blade), and clavicle (collarbone), which together enabled a wide range of motion. The proximal end of the humerus formed a ball-like joint that articulated with the scapula at the glenoid cavity, making it more susceptible to fractures under high impact. The distribution of fracture types varied across age groups, where clavicle fractures were more frequent in children, while adults predominantly experienced fractures in the proximal humerus region. In this study, X-ray images from the MURA dataset were utilized, where annotated samples provided reliable ground truth for fracture identification. These anatomical insights supported the development of an automated classification system using transformer-based feature extraction and ensemble learning models for accurate fracture detection as shown in Figure 1.

The types of shoulder bone fractures are divided into three categories in general: clavicle fractures, which are the most common shoulder fracture, frequently the result of a fall, scapula fractures, which rarely occur, and resulting fractures, which occur as cracks in the upper part of the arm in individuals over 65 years of age. The images from X-ray devices are primarily used for imaging of the shoulder bone for diagnosis and treatment of such fractures, while MRI or CT devices may also be used when required [1]. Deep learning classification procedures of shoulder bone X-ray images were carried out in the study. The main contributions of this study are as follows: The most suitable model for the classification of shoulder bone X-ray images as a fracture or non-fracture is determined [2]. An approach that can be used in similar studies is developed via new ensemble learning models. The study can assist physicians who are not experts in the field in the classification stage, especially in cases of shoulder fractures, which are frequently encountered in the emergency departments of hospitals [3]. The method suggested in the study contributes to the literature with two different ensemble approaches. With the first model proposed in the study, a performance study was conducted with transfer learning for the MURA dataset, which is widely used in X-ray studies.

Thus, three models that give the best classification results have been combined into a single model, and the classification performance has been increased [4]. The developed model can be used to classify many medical X-ray images. With the second model, it is determined which model finds which class best by looking at the success of finding the classes on the models in the dataset. It ensures that the model decides the prediction of that class [5]. Thus, regardless of the dataset studied, a similar decision system can be designed with models that find the classes in each dataset, and a higher performance can be achieved compared to a single model. The proposed ensemble model approaches can be applied and generalized by performing similar preprocessing steps in other X-ray biomedical datasets. In addition, the proposed method can be easily used in studies with transfer learning.

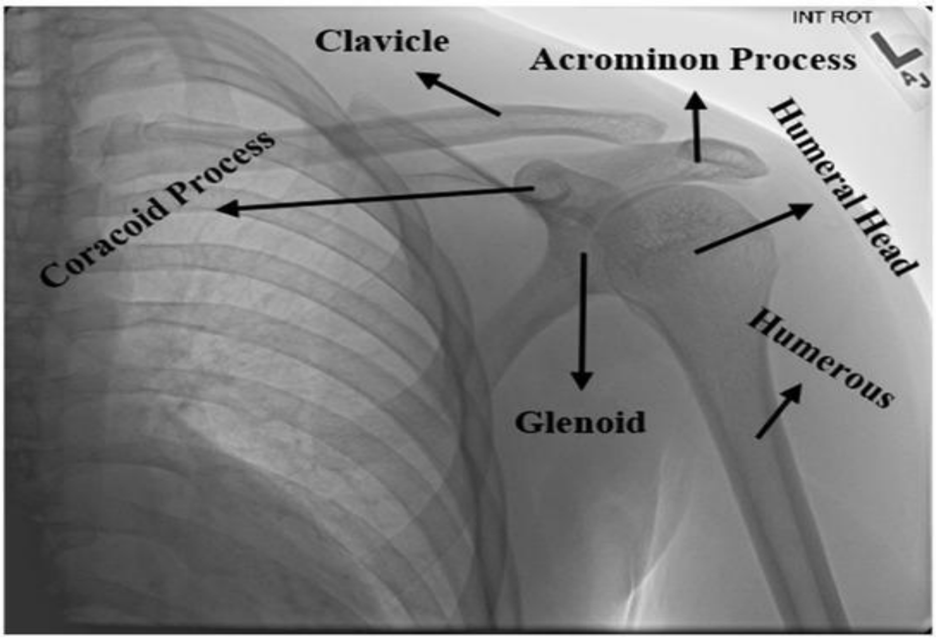

Figure 1: The anatomy of the shoulder bone.

2. Literature Survey

Uysal, et al. [6] considered pre-trained models in the detection of shoulder fractures were evaluated on the musculoskeletal radiographs (MURA) dataset, and two ensemble learning models (EL1 and EL2) were developed. The pre-trained models used are ResNet, ResNeXt, DenseNet, VGG, Inception, MobileNet, and their spinal fully connected (Spinal FC) versions. In the EL1 and EL2 models developed using pre-trained models with the best performance, test accuracy was 0.8455, 0.8472, Cohen’s kappa was 0.6907, 0.6942 and the area that was related with fracture class under the receiver operating characteristic (ROC) curve (AUC) was 0.8862, 0.8695. Sumon, et al. [7] assessed the diagnostic efficacy of the artificial intelligence (AI) model before and after optimization and compared its performance in detecting fractures or not. The training and evaluation dataset consists of fractured and non-fractured X-rays from various anatomical locations, including the hips, knees, lumbar region, lower limb, and upper limb.

Su, et al. [8] summarized based on key aspects such as the bones involved, research objectives, dataset sizes, methods employed, results obtained, and concluding remarks. This process distills the diverse approaches into a generalized processing framework or workflow. Tanzi, et al. [9] aimed to analyze and evaluate a selection of papers, chosen according to their representative approach, where the authors applied different deep learning techniques to classify bone fractures, to select the strengths of each of them and try to delineate a generalized strategy. Tariq, et al. [10] aimed to increase the existing state-of-the-art convolutional neural networks (CNNs)' performance by using various ensemble techniques. In this approach, different CNNs are used to classify the images; rather than choosing the best one, a stacking ensemble provides a more reliable and robust classifier. The ensemble model outperforms the results of individual CNNs by an average of 10%.

Kandel, et al. [11] proposed a global-local feature fusion convolutional neural network, including a global pathway to capture the global contextual information and a local pathway to extract the fine-grained information from local patches. Hui, et al. [12] presented the appropriate method to classify musculoskeletal images by transfer learning and by training from scratch. They applied six state-of-the-art architectures and compared their performance with transfer learning and with a network trained from scratch. Kandel, et al. [13] reviewed provides a focused analysis of CNN evolution and architectures as applied to medical image analysis, highlighting their application and performance in different medical fields, including oncology, neurology, cardiology, pulmonology, ophthalmology, dermatology, and orthopaedics.

Mienye, et al. [14] developed a tool to streamline mammogram classification that maintains high reliability across different data sources. They used images from the DDSM data set and a proprietary data set, YERAL, which comprises 943 mammograms from Mexican patients. Berrones-Reyes, et al. [15] discussed the Wide ResNet-40, DenseNet-121, and EfficientNet-B7 are chosen, fine-tuned, and used as base models, and a Bayesian-based probabilistic ensemble learning method is proposed for fracture detection in cervical spine CT images. The proposed method considers the prediction's uncertainty of the base models and combines the predictions obtained from them, to improve the overall performance significantly. Pandey, et al. [16] reviewed an overview of the use of DL in bone imaging to help radiologists to detect various abnormalities, particularly fractures. They also discussed the challenges and problems faced in the DL-based method, and the future of DL in bone imaging.

Meena, et al. [17] addressed the multi-label classification challenge of chest X-ray images using the Chest X-ray14 dataset. They propose a novel online ensemble technique that differs from previous penalty-based methods by focusing on combining individual model losses with the overall ensemble loss. Katona, et al. [18] addressed the multi-label classification challenge of chest X-ray images using the Chest X-ray14 dataset. They propose a novel online ensemble technique that differs from previous penalty-based methods by focusing on combining individual model losses with the overall ensemble loss.

Rekha Gangula et al. [19] proposed a methodology for early Alzheimer's disease detection using Brain Magnetic Resonance Imaging (MRI). The system performed image preprocessing and feature extraction. The classification model improved early-stage disease identification. Rekha Gangula et al.

[20] proposed a diabetes prediction model using Logistic Regression. The framework analyzed clinical datasets for risk prediction. The model improved prediction accuracy and disease prognosis. Pavankumar Nagapuri et al. [21] proposed a deep learning hybrid model for lung disease detection using Chest X-Ray (CXR) images. The model combined feature extraction and classification stages. The approach improved diagnostic accuracy in medical imaging. Rekha Gangula et al. [22] proposed an ensemble machine learning approach for dengue disease prediction. The framework utilized multiple classifiers for improved performance. The model enhanced prediction accuracy and reliability.

3. Proposed Methodology

The proposed system automatically detected and classified bone fractures from X-ray images using an intelligent machine learning framework. Deep features were extracted using the DeiT model to capture critical structural patterns in the images. These features were used to train multiple classifiers, where the FIGS model was selected based on superior performance. An XAI-based module validated the input image and provided fracture-related insights before final prediction. The system was deployed through a Telegram Bot interface to enable real-time fracture analysis and visualization of results as shown in Figure 2.

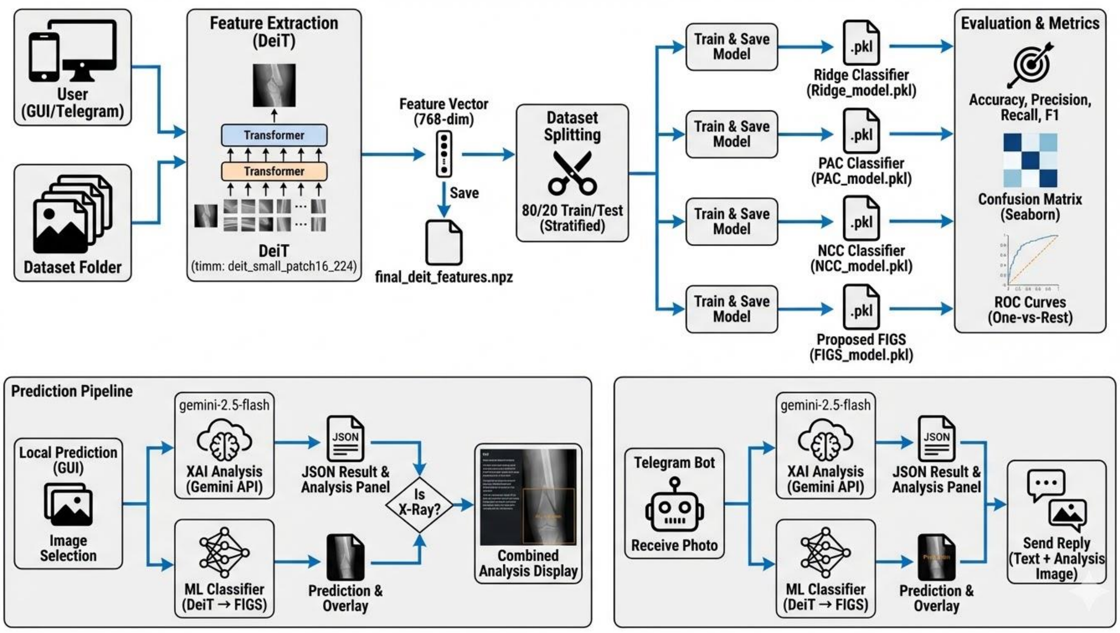

Figure 2: Proposed System architecture of Bone fracture classification

Step 1. Data Acquisition and Loading: The process begins with the user selecting the directory containing the bone X-ray images. The system reads these directory names to identify the available classification categories. This step effectively loads the raw image data and defines the scope of the classification problem.

Step 2. DeiT Feature Extraction: The images are then processed using the pre-trained DeiT , which acts as a powerful feature extractor. Each image is transformed into a fixed-size, high-dimensional feature vector by passing it through the transformer's layers. This critical step reduces the complexity of working directly with high-resolution image data, resulting in the numerical matrices X (features) and Y (labels) required for the subsequent machine learning models.

Step 3. Dataset Preparation and Splitting: The extracted feature vectors (X) and their labels (Y) are prepared for model training. The data is partitioned into training and testing sets using an 80/20 split, employing stratification to ensure that the distribution of fracture classes remains consistent in both sets. This rigorous split guarantees that the models are trained on a representative portion of the data and evaluated fairly on completely unseen data.

Step 4. Model Training and Evaluation (Existing Classifiers): Baseline machine learning classifiers including the RC, PAC, and NCC are trained on the DeiT-extracted features. Their performance is immediately assessed using a comprehensive suite of metrics like Accuracy, Precision, Recall, and F1-score. These results, along with visual aids like the Confusion Matrix and ROC Curves, establish a necessary performance benchmark.

Step 5. Training the Proposed Classifier: The central classification model, the Proposed FIGS classifier, is trained on the same prepared training data. FIGS, designed for speed and clarity, learns the patterns in the DeiT features to classify fractures. Once trained, this final model is permanently saved using joblib for deployment, serving as the system's primary diagnostic tool.

Step 6. XAI-Powered Pre-Screening and Analysis: In the deployment phase (for new inputs), the system first uses a Generative AI for XAI analysis. This module performs robust image type detection (ensuring it's a bone X-ray) and, if confirmed, provides a detailed medical assessment, identifying the affected bone, fracture type, and severity in a structured JSON output. This step confirms the image's validity and provides essential diagnostic context.

Step 7. Final Prediction and Visualization: Following XAI confirmation, the image is passed through the DeiT feature extractor, and the resulting features are fed into the trained Proposed FIGS model for the final automated fracture classification. The system then displays a comprehensive visual result, combining the original X-ray image, the detailed XAI report panel, and the FIGS model's prediction in a unified interface, ensuring both accuracy and interpretability.

Step 8. Telegram Bot Integration: For maximum accessibility, the entire prediction pipeline is wrapped into a Telegram Bot. Users interact by sending an image, and the bot automatically triggers the XAI screening, the DeiT feature extraction, and the FIGS classification. It then replies instantly with a text summary of the findings and the visual analysis panels, enabling real-world, mobile usage.

4. Results and Discussion

Figure 3 displays the prediction result comparison between the original X-ray image and the model’s classification output generated using the transformer-based fracture detection system. The left panel shows the input bone X-ray image, while the right panel presents the Transformer Prediction, where the model has identified the fracture type and displayed the label “Predicted: Comminuted fracture” on the image. This output demonstrates how the trained model analyzes bone structure patterns from the X-ray image and accurately classifies the fracture type, helping in automated and efficient fracture diagnosis.

Figure 4 shows a comparison between the original foot X-ray image and the model’s classification output generated by the transformer-based fracture detection system. The left panel displays the input X-ray image, while the right panel shows the Transformer Prediction, where the system has analyzed the bone structure and identified the condition as “Fracture Dislocation.” The predicted label is displayed on the image, demonstrating how the trained model automatically detects and classifies bone fracture types from medical X-ray images for accurate and efficient diagnosis.

AMERICAN JOURNAL OF MANAGEMENT AND IOT MEDICAL COMPUTING
Peer Reviewed, Referred & Indexed Journal
E-ISSN: 3069-0110
Vol.5, No.2(2026)
www.ajmimc.com

Figure 3: Prediction obtained using FIGS model

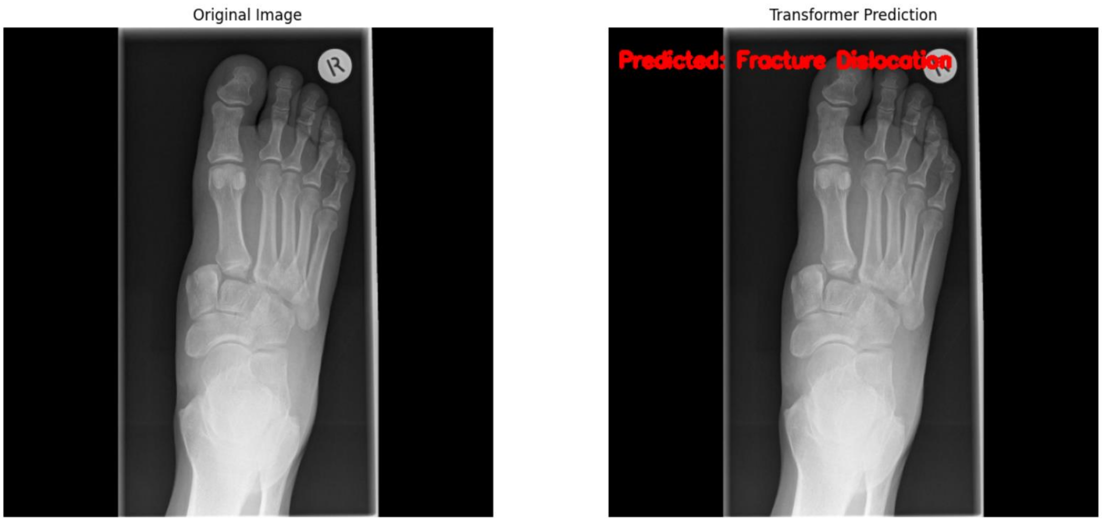

Figure 4: Prediction obtained using FIGS model

Figure 5 shows the Telegram that demonstrates real-time image analysis by processing images sent by users and providing automated responses. In the first case, when a handwritten document image is uploaded, the system analyzes the content and correctly identifies that the image is not an X-ray, classifying it as a handwritten document and preventing unnecessary fracture prediction. In the second case, when a foot X-ray image is submitted, the system successfully detects that it is a valid X-ray, identifies the bone region (5th metatarsal and cuboid), confirms the presence of a fracture, determines the fracture type as avulsion, estimates the severity as moderate, and finally provides the classifier prediction “Avulsion fracture.” These results demonstrate the effectiveness of the Telegram bot in performing real-time X-ray verification and automated fracture classification.

Table 1 presents the performance comparison of four classification algorithms such as RC, PAC, NCC, and the proposed FIGS model based on evaluation metrics including Accuracy, Precision, Recall, and F-score. The RC achieves an accuracy of 71.87%, with precision, recall, and F-score values of 71.66%, 71.87%, and 71.23%, indicating moderate classification capability in identifying different fracture categories. The PAC shows slightly improved performance with an accuracy of 73.5%, a higher precision of 81.33%, recall of 73.49%, and F-score of 74.38%, demonstrating better ability in correctly identifying fracture instances compared to the RC model. In contrast, the NCC performs poorly, achieving only 18.5% accuracy, with precision, recall, and F-score values of 26.67%, 18.49%, and

14.67%, suggesting that this distance-based method is not well suited for the complex feature patterns extracted from bone X-ray images. The proposed FIGS model significantly outperforms all existing classifiers, achieving 96.41% accuracy, 96.43% precision, 96.41% recall, and 96.42% F-score, indicating highly reliable and consistent classification performance. These results clearly demonstrate that the FIGS model effectively captures the discriminative patterns in transformer-extracted features, making it the most accurate and robust approach for multi-class bone fracture classification among the evaluated methods.

Figure 5: Telegram bot activated for real time prediction

Table 1: Performance comparison for the RC, PAC, NCC and Proposed FIGS Model

Algorithms Name	Accuracy	Precision	Recall	F-score
RC	71.87%	71.66%	71.87%	71.23%
PAC	73.5%	81.33%	73.49%	74.38%
NCC	18.5%	26.67%	18.49%	14.67%
FIGS model	96.41%	96.43%	96.41%	96.42%

5. Conclusion

The system was successfully developed to automatically detect and classify different types of bone fractures from X-ray images using advanced artificial intelligence techniques. The system utilizes the DeiT model to extract deep visual features from X-ray images, which are then used to train multiple machine learning classifiers including RC, PAC, NCC, and the proposed FIGS model. Performance evaluation using metrics such as accuracy, precision, recall, and F-score demonstrated that the proposed FIGS model significantly outperforms the existing models, achieving an accuracy of 96.41%, making it highly effective for multi-class fracture classification. The system also integrates an XAI-based image verification module to determine whether the uploaded image is a valid bone X-ray before performing fracture analysis, improving reliability and preventing incorrect predictions on non-medical images. In addition, the integration of a Telegram bot interface enables real-time remote analysis, allowing users to send images directly to the bot and receive automated fracture detection results along with fracture type, bone location, and severity information. The experimental results, visualization outputs, and real-time testing confirm that the proposed system provides an accurate, efficient, and user-friendly solution for automated bone fracture detection, which can assist medical professionals in preliminary diagnosis and reduce the time required for manual fracture analysis.

References

[1] Shoulder Fracture 2014. Available online: https://www.assh.org/handcare/condition/shoulder-fracture (accessed on 1 October 2025).

[2] Rajpurkar, P.; Irvin, J.; Bagul, A.; Ding, D.; Duan, T.; Mehta, H.; Yang, B.; Zhu, K.; Laird, D.; Ball, R.L.; et al. MURA Dataset: Towards Radiologist-Level Abnormality Detection in Musculoskeletal Radiographs. In Proceedings of the 1st Conference on Medical Imaging with Deep Learning, Amsterdam, The Netherlands, 4–6 June 2018.

[3] Guan, B.; Zhang, G.; Yao, J.; Wang, X.; Wang, M. Arm fracture detection in X-rays based on improved deep convolutional neural network. Comput. Electr. Eng. 2020, 81, 1–11.

[4] Galal, A.; Hisham, F.; Mohamed, M.; Hassan, S.; Ghanim, T.; Nabil, A. Automatic Recognition of Elbow Musculoskeletal Disorders using Cloud Application. In Proceedings of the 2019 8th International Conference on Software and Information Engineering, Cairo, Egypt, 9–12 April 2019.

[5] Liang, S.; Gu, Y. Towards Robust and Accurate Detection of Abnormalities in Musculoskeletal Radiographs with a Multi-Network Model. Sensors 2020, 20, 3153.

[6] Uysal, F.; Hardalaç, F.; Peker, O.; Tolunay, T.; Tokgöz, N. Classification of Shoulder X-ray Images with Deep Learning Ensemble Models. Appl. Sci. 2021, 11, 2723. https://doi.org/10.3390/app11062723.

[7] Sumon, R.I.; Ahammad, M.; Mozumder, M.A.I.; Hasibuzzaman, M.; Akter, S.; Kim, H.-C.; Al-Onaizan, M.H.A.; Muthanna, M.S.A.; Hassan, D.S.M. Automatic Fracture Detection Convolutional Neural Network with Multiple Attention Blocks Using Multi-Region X-Ray Data. Life 2025, 15, 1135. https://doi.org/10.3390/life15071135.

[8] Purmani, S. S. R. (2024). Aligning IT investment decisions with overall business strategy from an enterprise program management perspective, focusing on the integration of IT leadership in strategic decision-making processes. International Journal of Communication Networks and Information Security, 16(5), 1213–1219.

[9] Tanzi, L.; Vezzetti, E.; Moreno, R.; Moos, S. X-Ray Bone Fracture Classification Using Deep Learning: A Baseline for Designing a Reliable Approach. Appl. Sci. 2020, 10, 1507. https://doi.org/10.3390/app10041507.

[10] Tariq, M.; Choi, K. YOLO11-Driven Deep Learning Approach for Enhanced Detection and Visualization of Wrist Fractures in X-Ray Images. Mathematics 2025, 13, 1419. https://doi.org/10.3390/math13091419.

[11] Kandel, I.; Castelli, M.; Popovič, A. Comparing Stacking Ensemble Techniques to Improve Musculoskeletal Fracture Image Classification. J. Imaging 2021, 7, 100. https://doi.org/10.3390/jimaging7060100.

[12] Hui, Q.; Wang, C.; Weng, J.; Chen, M.; Kong, D. A Global-Local Feature Fusion Convolutional Neural Network for Bone Age Assessment of Hand X-ray Images. Appl. Sci. 2022, 12, 7218. https://doi.org/10.3390/app12147218.

[13] Kandel, I.; Castelli, M.; Popovič, A. Musculoskeletal Images Classification for Detection of Fractures Using Transfer Learning. J. Imaging 2020, 6, 127. https://doi.org/10.3390/jimaging6110127.

[14] Mienye, I.D.; Swart, T.G.; Obaido, G.; Jordan, M.; Ilono, P. Deep Convolutional Neural Networks in Medical Image Analysis: A Review. Information 2025, 16, 195. https://doi.org/10.3390/info16030195.

[15] Berrones-Reyes, M.C.; Salazar-Aguilar, M.A.; Castillo-Olea, C. Use of Ensemble Learning to Improve Performance of Known Convolutional Neural Networks for Mammography Classification. Appl. Sci. 2023, 13, 9639. https://doi.org/10.3390/app13179639.

[16] Pandey, A.K.; Senapati, K.; Argyros, I.K.; Pateel, G.P. Improving Vertebral Fracture Detection in C-Spine CT Images Using Bayesian Probability-Based Ensemble Learning. Algorithms 2025, 18, 181. https://doi.org/10.3390/a18040181.

[17] Meena, T.; Roy, S. Bone Fracture Detection Using Deep Supervised Learning from Radiological Images: A Paradigm Shift. Diagnostics 2022, 12, 2420. https://doi.org/10.3390/diagnostics12102420.

[18] Katona, T.; Tóth, G.; Petró, M.; Harangi, B. Advanced Multi-Label Image Classification Techniques Using Ensemble Methods. Mach. Learn. Knowl. Extr. 2024, 6, 1281-1297. https://doi.org/10.3390/make6020060.

[19] Rekha Gangula,A.Manjula, , R Bhukya, R Saturi, N Gayatri, R Rajesh "Innovative Approaches for Early Alzheimer's Disease Detection through Novel Analysis of Brain MRI Images" Vol. 1 No. 1 (2024): Special Issue: International Conference of Advance Research and Innovation,https://doi.org/10.5750/ijme.v1i1.1402.

[20] Rekha Gangula, Chinnakka Sudha, K Sreeveda, Rajitha Bonagiri, C Bhagyashree, Swapna Saturi," "Prediction and Prognosis of Diabetes Using Logistic Regression," 2022 IEEE North Karnataka Subsection Flagship International Conference (NKCon), Vijaypur, India, 2022, pp. 1-7, doi: 10.1109/NKCon56289.2022.10126692.

[21] Pavankumar Nagapuri, Swapna Saturi, K Vinay Kumar, G Rekha, Tejaswi Kalva, Amith Reddy Maada," "Deep learning hybrid for X-ray lung disease detection," 2022 Second International Conference on Advanced Technologies in Intelligent Control, Environment, Computing & Communication Engineering (ICATIECE), Bangalore, India, 2022, pp. 1-7, doi: 10.1109/ICATIECE56365.2022.10047575.

[22] Rekha Gangula, Lingala Thirupathi, Rajashekar Parupati, K. Sreeveda, Saritha Gattoju,Ensemble machine learning based prediction of dengue disease with performance and accuracy elevation patterns,Materials Today: Proceedings,Volume 80, Part 3,2023, Pages3458-3463, ISSN2214-7853, https://doi.org/10.1016/j.matpr.2021.07.270