Title: Hybrid Deep Learning Framework for Knee Osteoporosis Diagnosis Using CNNS and Fine-Tuned Multimodal Large Language Models

Author names: Md Rana Hossain; Tushar Kumar Ghosh; Ponchanon Datta Rone; Farhan Tanvir; Alaya Parven Alo; Md. Sadekur Rahman

Abstract: Early and accurate diagnosis of osteoporosis of the knee is of utmost significance for prompt treatment and planning. In this paper, we propose a hybrid deep learning system where Convolutional Neural Networks (CNNs) are combined with finetuned multimodal Large Language Models (LLMs) for enhanced diagnosis. Initially, we experimented with different leading CNN architectures, including InceptionV3, DenseNet121, MobileNetV2, and ResNet50. Even though such models were moderately successful, with about 80% accuracy, their performance was not yet sufficient for reliable use in clinics. To surpass such a deficiency, we blended an elementary multimodal approach that also exhibited some better results but not sufficient to achieve the desired level. Hence, we embarked on a high-performance multimodal LLM-based architecture, which effectively captured complex patterns from both imaging as well as associated clinical metadata. This integration significantly enhanced the diagnostic accuracy to 93%, which was superior to standard CNN-based and untuned multimodal methods. Our results show that utilizing the representational ability of fine-tuned multimodal LLMs in conjunction with CNNs offers a promising pathway towards improving medical image analysis, particularly for osteoporosis detection.

Key words: Protocols , HTTP , Pixel , Fuses , Artificial intelligence

https://ieeexplore.ieee.org/document/11451868/keywords#keywords

DOI: 10.1109/ICSADL67539.2026.11451868