COSIT, TASFUED
Journal of Science and Information Technology (JOSIT)

Identifying Significant Structural Factors associated with Knee Pain Severity in Patients with Osteoarthritis using Hybrid Bio-BERT Bi-LSTM CNN Model

O. L. Usman

Department of Computer Science, Tai Solarin Federal University of Education, Ijagun, Ogun State, Nigeria.
Corresponding Author: usmanol@tasued.edu.ng

Abstract

Knee osteoarthritis (OA) is a progressive, long-term joint disorder marked by pain, stiffness, and impaired mobility, significantly diminishing patients' quality of life. Accurately predicting the severity of knee pain is challenging due to the complex, multifactorial aetiology of OA and the heterogeneity in both structural changes and clinical symptoms among patients. This study introduces a hybrid deep learning model which integrates Bidirectional Encoder Representations from Transformers for Biomedical Text Mining (Bio-BERT), Bidirectional Long Short-Term Memory (Bi-LSTM), and Convolutional Neural Network (CNN) models to predict knee OA pain severity by combining clinical text and Magnetic Resonance Imaging (MRI) datasets. Using a multimodal Kaggle datasets which contains 163,064 clinical records and 2000 MRI scans, the model achieved a maximum test accuracy of 98.11%, with strong precision, recall, and F1-scores, particularly for severe pain cases. These results demonstrate its effectiveness in synthesizing textual and visual features, supporting potential clinical applications in diagnosis, early intervention, and treatment planning. Future work will focus on improving interpretability, validating external generalizability, enabling real-time Electronic Health Record (HER) integration, and ensuring ethical AI practices, including patient privacy and transparency.

Keywords: Knee Osteoarthritis, Bio-BERT, Bi-LSTM, Magnetic Resonance Imaging, Pain Severity

INTRODUCTION

According to the Osteoarthritis Research Society International (2020), osteoarthritis (OA) is among the most widespread degenerative joint disorders globally, with knee osteoarthritis representing the most prevalent clinical manifestation. The condition is marked by the progressive deterioration of articular cartilage and subchondral bone, leading to chronic pain, stiffness, and functional limitation. Although OA has been extensively studied, accurately predicting the severity of knee pain remains a persistent clinical challenge (Morals et al. 2021; Panwar et al. 2024). This difficulty arises from the complex and non-linear interactions among multiple structural and biological factors, including joint space narrowing, cartilage degeneration, osteophyte formation, and synovial inflammation. Furthermore, the relationship between radiographic findings and patient-reported pain is often inconsistent, complicating clinical assessment. A clearer understanding of the relative contribution of these structural abnormalities to pain severity is therefore essential, as it could improve diagnostic precision, support personalized treatment planning, and inform the development of targeted therapeutic interventions aimed at slowing disease progression and improving patient outcomes.

Johnson & Hunter (2019) reported that a range of methodologies have been employed to predict pain severity in patients with

Cite as:

Usman, O. L. (2025). Identifying Significant Structural Factors associated with Knee Pain Severity in Patients with Osteoarthritis using Hybrid Bio-BERT Bi-LSTM CNN Model. *Journal of Science and Information Technology (JOSIT)*, Vol. 19 No. 2, pp. 18-28.

osteoarthritis using clinical and structural data. Conventional statistical techniques, particularly logistic regression, have been widely adopted; however, these approaches often struggle to capture the complex, non-linear relationships that characterize OA progression. For example, logistic regression models have been applied to predict knee pain severity from radiographic features, but their relatively simple structure limits the ability to model interactions among multiple contributing factors. In contrast, more advanced machine learning techniques, including support vector machines and random forests, have demonstrated improved predictive performance by better capturing complex patterns within the data (Pedoia et al., 2021; Teoh et al. 2023). Random forest models have been applied to the analysis of imaging data, showing promise in identifying salient structural abnormalities. Nevertheless, such approaches typically concentrate on individual clinical or imaging features and often lack the capability to effectively integrate heterogeneous data modalities (Lazzarini et al., 2017; Lin et al. 2023; Usman et al. 2021).

In recent years, deep learning approaches have gained prominence in medical image analysis and natural language processing (NLP). In particular, convolutional neural networks (CNNs) have been effectively utilized to analyze MRI scans of osteoarthritis patients, enabling the prediction of structural changes such as cartilage thinning and joint space narrowing. Studies have shown that CNNs can successfully learn imaging patterns associated with OA progression (Guo et al. 2024). However, despite these advances, CNN-based models are typically restricted to imaging data and fail to leverage complementary clinical text information, such as patient-reported symptoms and medical histories, which can provide critical contextual insight into pain severity (Smith & Lee, 2021; Muhammad et al. 2021; Thomas et al. 2020; Nguyen et al. 2020; Panfilov et al. 2022).

To overcome these limitations, this study introduces a novel hybrid deep learning framework that integrates clinical text and structural imaging data through the combined use of Bio-BERT and Bi-LSTM architectures. Bio-BERT (Bidirectional Encoder Representations from Transformers) is well established for extracting rich semantic representations from biomedical text, whereas Bi-LSTM (Bidirectional Long Short-Term Memory) networks are effective in modelling sequential dependencies in longitudinal or structured data. By fusing these complementary architectures, the proposed system is designed to identify key structural determinants associated with knee pain severity in osteoarthritis patients. This integrative approach exploits the strengths of deep learning for multimodal data analysis, enabling a more comprehensive characterization of both symptomatic and structural contributors to pain. Consequently, it offers a robust framework for improving the accuracy of knee pain severity prediction, with potential implications for enhanced clinical decision-making and personalized treatment planning (Liu et al., 2022; Usman et al. 2025; Usman & Adeusi, 2025; Owoade et al. 2025).

Accordingly, this study addresses a critical gap in existing research by developing a hybrid Bio-BERT Bi-LSTM model capable of effectively combining clinical narratives and imaging information. The proposed method aims to deliver a more accurate and interpretable solution for predicting knee pain severity while elucidating the structural factors that significantly influence disease progression.

MATERIALS AND METHODS

Materials: Data Collection and Preprocessing

This study employed a multimodal data collection strategy to support the prediction of knee pain severity in osteoarthritis (OA) patients by integrating clinical text, structured patient information, and imaging-related features. The dataset, sourced from reputable institutions and publicly available repositories such as Kaggle database, comprised high-quality structured clinical records stored in CSV file format and included patients diagnosed with or at risk of knee OA. Key variables encompassed demographic factors (e.g., age and BMI), clinical indicators (e.g., pain frequency, swelling, crepitus, prior knee surgery, and functional limitations), and KOOS (Knee injury and Osteoarthritis Outcome Score) pain scores, which served as ground truth labels. Pain severity was categorized into three classes: mild, moderate, and severe, to enable supervised learning. The study utilized a dataset comprising 163,064 clinical text records and 2,000 MRI scans. Preprocessing procedures were undertaken to ensure effective

synchronization between the textual and imaging data. Specifically, clinical records and MRI scans belonging to the same patient were systematically paired. The textual data were processed using natural language processing (NLP) techniques, including noise and artifact removal, tokenization, and normalization. Subsequently, Bio-BERT was employed to generate contextual embeddings, enabling the extraction of structured clinical features such as pain scores and symptoms (e.g., stiffness and swelling). For the imaging modality, MRI scans were standardized through intensity normalization and resized to focus on the region of interest (ROI), particularly the knee joint. A convolutional neural network (CNN) was then applied to extract relevant features and facilitate the segmentation of cartilage and bone structures. Finally, feature alignment was conducted to integrate the embedded textual representations with the extracted image features, thereby preparing the data for subsequent multimodal analysis.

To ensure suitability for machine learning, categorical clinical variables were encoded numerically, while continuous features were standardized for uniform scaling. Structured clinical attributes were further transformed into natural language sentences to allow semantic processing by the Bio-BERT model. These text representations were tokenized and embedded using the Bio-BERT tokenizer with a fixed sequence length, enabling the extraction of contextual relationships among clinical features. Imaging data, including MRI scans, were subjected to normalization and resizing procedures to ensure uniformity across samples. Additional preprocessing steps, such as noise reduction and data standardization, were applied to enhance image quality and consistency. Collectively, these preprocessing operations ensure that the multimodal datasets is appropriately structured and optimized for effective model training and accurate prediction of knee pain severity in osteoarthritis (OA) patients. It is therefore, sufficient to state that the data collection and processing procedures adhered to strict ethical and quality control protocols, including informed consent and privacy protection as applicable to publicly available datasets. This comprehensive and well-curated dataset provided a robust foundation for training and evaluating the proposed hybrid Bio-BERT Bi-LSTM–CNN model for knee pain severity prediction.

Methods

Figure 1 illustrates the flowchart, offering a visual overview of the workflow of the proposed system and method employed in this study, while Figures 2(a) and 2(b) present the architectural representations of the proposed Bio-BERT Bi-LSTM–CNN model, respectively. The system architecture adopts a deep learning pipeline tailored for predicting knee pain severity in patients with OA. It integrates Bio-BERT for processing clinical text, Bi-LSTM for modelling sequential clinical information, and a CNN for extracting localized features from imaging data. The overall architecture is organized into the following components:

1. **Data Collection:**

i. **Clinical Data:** Comprises demographic information (e.g., age, BMI), medical history, and symptom-related variables. This structured dataset includes indicators such as prior surgeries, presence of swelling, and frequency of pain episodes.

ii. **Textual Data:** Structured clinical information is converted into natural language sentences to facilitate processing by the Bio-BERT model. For instance, a patient's record may be represented as a sentence such as, "Age 45 years, BMI 27, frequent pain 1, surgery 0…"

iii. **Imaging Data:** MRI and X-ray scans are utilized to evaluate structural alterations in the knee joint and to identify localized patterns associated with key indicators of knee pain severity, including variations in pain intensity and relevant risk factors.

2. **Preprocessing Layer:**

i. **Text Generation:** Structured clinical data is converted into descriptive textual format through a custom 'create_text' function.

ii. **Label Encoding:** Categorical variables, including "FREQUENT PAIN," "SURGERY," "RISK," and similar features, are transformed into numerical representations using LabelEncoder.

iii. **Tokenization:** The generated textual data is processed using the Bio-BERT tokenizer to prepare it for input into the model, with sequences padded to a uniform length of 128 tokens.

3. **Feature Extraction Layer:**

i. **Bio-BERT for Text Data:** Bio-BERT, a pre-trained transformer model, is applied to the tokenized clinical notes, producing embedding that encode rich semantic information from the clinical data.

ii. **Bi-LSTM for Sequential Data:** A Bidirectional Long Short-Term Memory (Bi-LSTM) layer is employed to model temporal dependencies and sequential relationships among clinical features, enabling the network to learn from patients' medical histories and disease progression.

4. **Feature Fusion Layer:** The features obtained from Bio-BERT and Bi-LSTM are concatenated into a unified representation, which is subsequently fed into the CNN layer for additional feature extraction.

5. **CNN Layer for Local Feature Extraction from Images:** A convolutional neural network (CNN) analyses the sequence of features produced by the Bi-LSTM, identifying localized patterns in imaging data that correspond to key indicators of knee pain severity, including variations in pain intensity and relevant risk factors.

6. **Output Layer:**

i. **Pain Severity Classification:** The output layer employs a softmax activation function to categorize pain severity into three levels: Severe, Moderate, or Mild, based on the integrated features derived from Bio-BERT Bi-LSTM, and CNN.

ii. **Softmax Output:** The softmax layer produces a probability distribution across the three classes, Severe, Moderate, and Mild.

7. **Training and Optimization:**

i. **Loss Function and Optimizer:** The model is compiled using the Adam optimizer in conjunction with the sparse categorical cross-entropy loss function. Training is conducted on the designated training dataset, and performance is assessed by monitoring accuracy and loss metrics across multiple epochs.

ii. **Training Data Input:** The training dataset, comprising textual features and their corresponding labels, is supplied to the model in batches, with model weights updated iteratively using the gradient descent algorithm.

8. **Performance Evaluation:** The model's performance is assessed using multiple evaluation metrics, including accuracy, confusion matrix, F1-score, precision, and recall. These measures provide a comprehensive assessment of the model's ability to correctly classify pain severity in the test dataset.

SIMULATION RESULTS AND DISCUSSION

The proposed Bio-BERT Bi-LSTM–CNN model was trained and evaluated under three different train–validation–test partitioning schemes (80:10:10, 75:10:15, and 70:15:15) to assess its robustness and generalization across varying data distributions. All experiments were conducted in Python using the TensorFlow and Keras frameworks, with training performed in a GPU-enabled environment to accelerate computation. The experimental setup is consistent with established medical imaging and clinical text analysis pipelines and reflects the resource considerations emphasized in energy-efficient modelling research. Figure 3 depicts the Python-based one-dimensional representation of the proposed hybrid Bio-BERT Bi-LSTM–CNN architecture, while Figure 4 presents screenshots of normalized sample images corresponding to severe, moderate, and mild osteoarthritis pain severity cases, respectively.

For each of the three experimental runs, model training was conducted for a maximum of 20 epochs. The input dimensions were 365 × 224 × 224 × 1 for the training samples, 78 × 224 × 224 × 1 for the validation samples, and 53 × 224 × 224 × 1 for the testing samples. Table 1 summarizes the model's performance across the different data split configurations according to the severity levels of knee osteoarthritis (OA) cases.

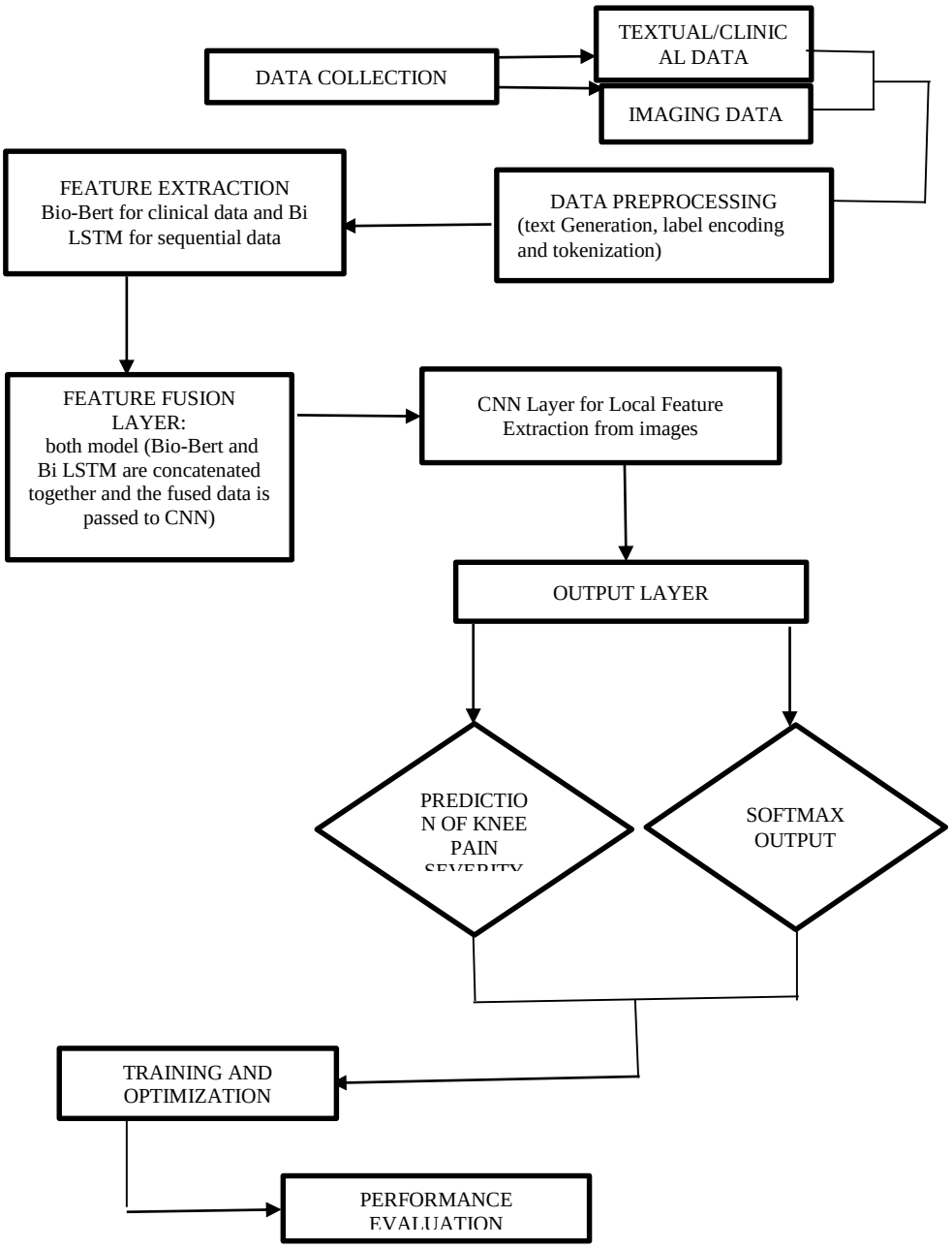
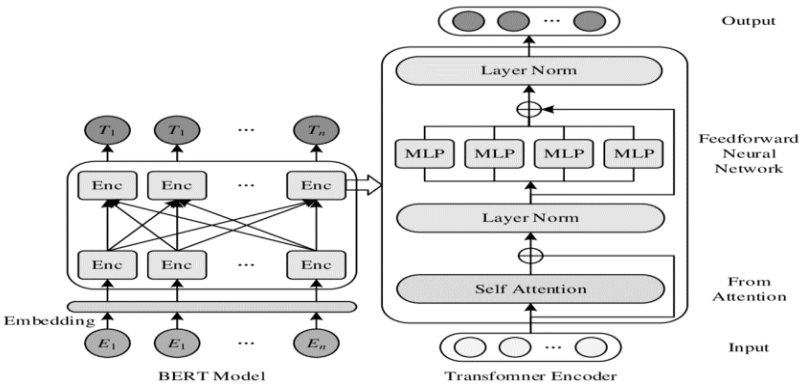
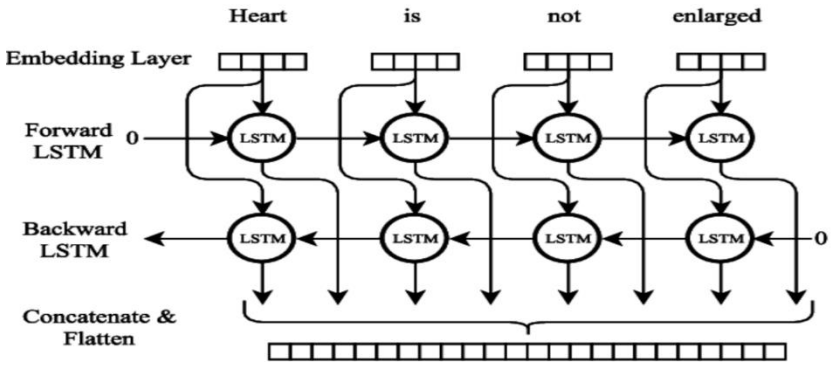


Figure 1. Proposed system architecture and method.



(a)



(b)

Figure 2. Architectural model representation: (a) Bio-BERT model (b) Bi-LSTM model.

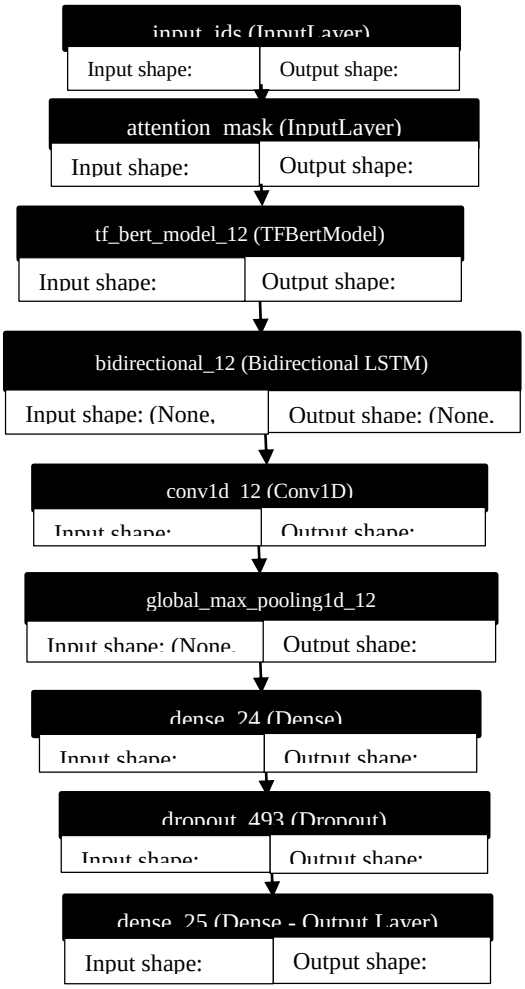


Figure 3. Proposed 1D hybrid Bio-BERT + Bi-LSTM + Conv1D model.

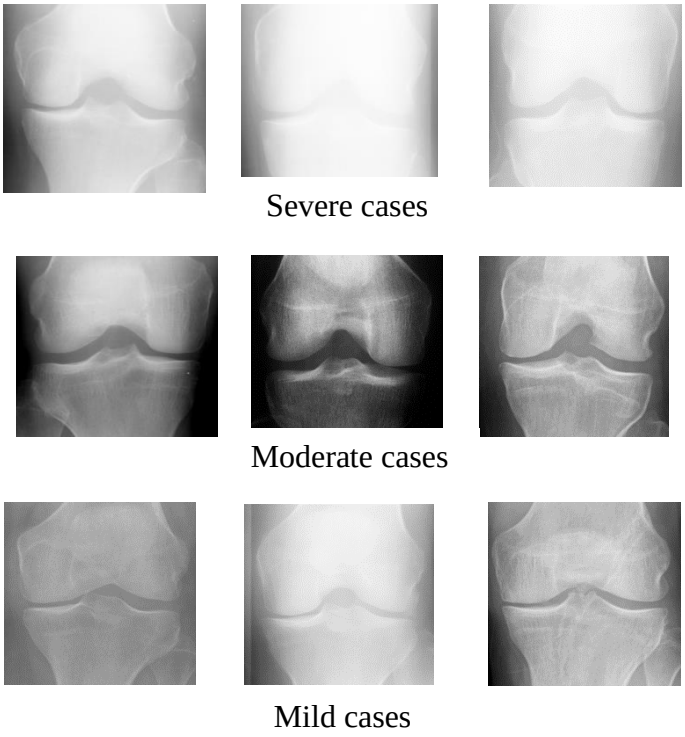


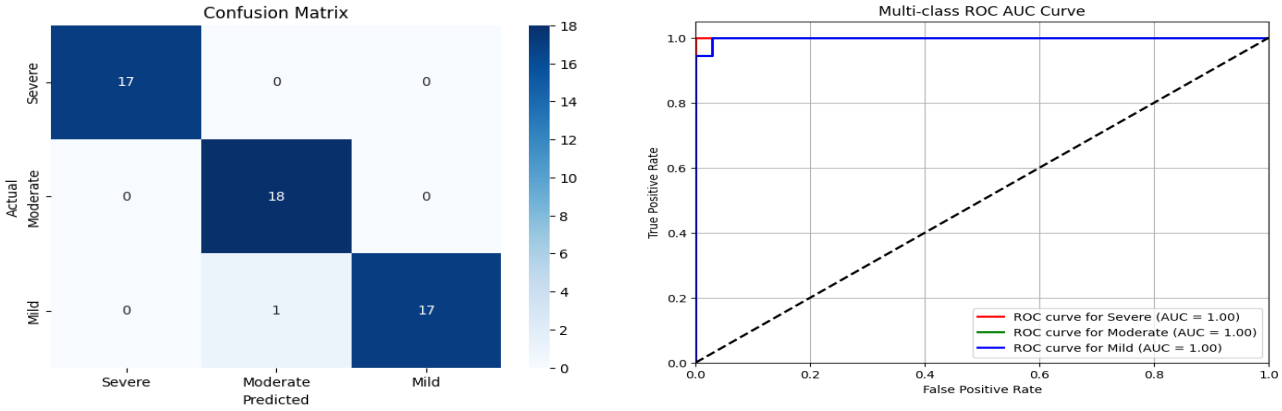
Figure 4. Screenshot of sample dataset for both mild, moderate, severe categories.

Table 1. Summary of classification performance of Bio-BERT Bi-LSTM–CNN model.

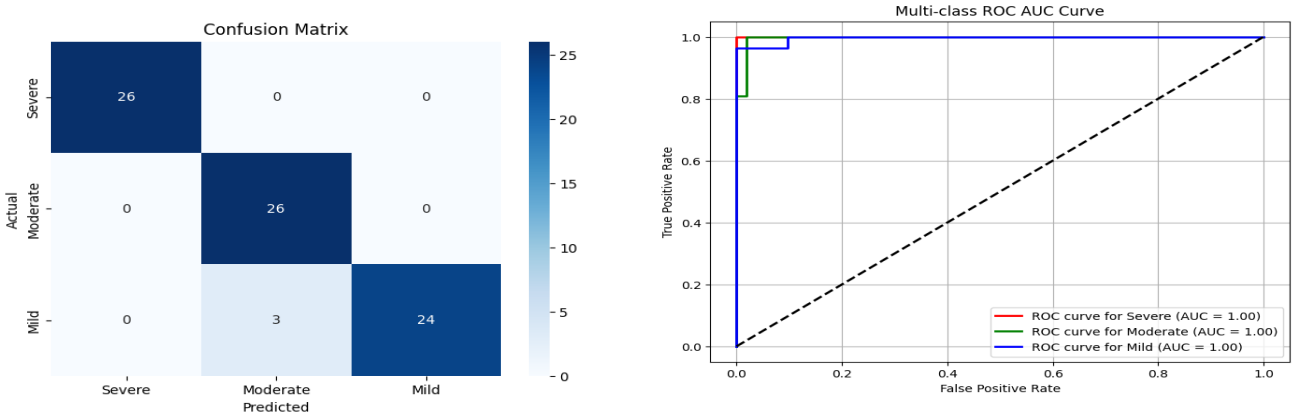
Data Partition Ratio	OA Severity Level	Precision	Recall	F1-Score	Weighted Average Accuracy
80:10:10	Severe	1.00	1.00	1.00	0.98
	Moderate	0.95	1.00	0.97	
	Mild	1.00	0.94	0.97	
75:15:10	Severe	1.00	1.00	1.00	0.96
	Moderate	0.90	1.00	0.95	
	Mild	1.00	0.89	0.94	
70:15:15	Severe	1.00	1.00	1.00	0.97
	Moderate	0.93	1.00	0.96	
	Mild	1.00	0.93	0.96	

As shown in Table 1, the 80:10:10 train–validation–test split yielded the best overall performance, achieving a test weighted average accuracy of 98.0%. This was followed by the 70:15:15 split with a weighted accuracy of 97.0%, while the 75:15:10 configuration attained a weighted accuracy of 96.0%. The confusion matrices and ROC–AUC curves presented in Figures 5(a)–(c) further demonstrate the stability and consistency of the proposed Bio-BERT Bi-LSTM–CNN model across all simulation runs, regardless of OA pain severity level, using the multimodal datasets. Collectively, these findings highlight the model's strong ability to balance sensitivity and precision while sustaining reliable classification performance. Overall, the ROC–AUC values ranged between 0.99 and 1.00, indicating excellent discriminative capability.

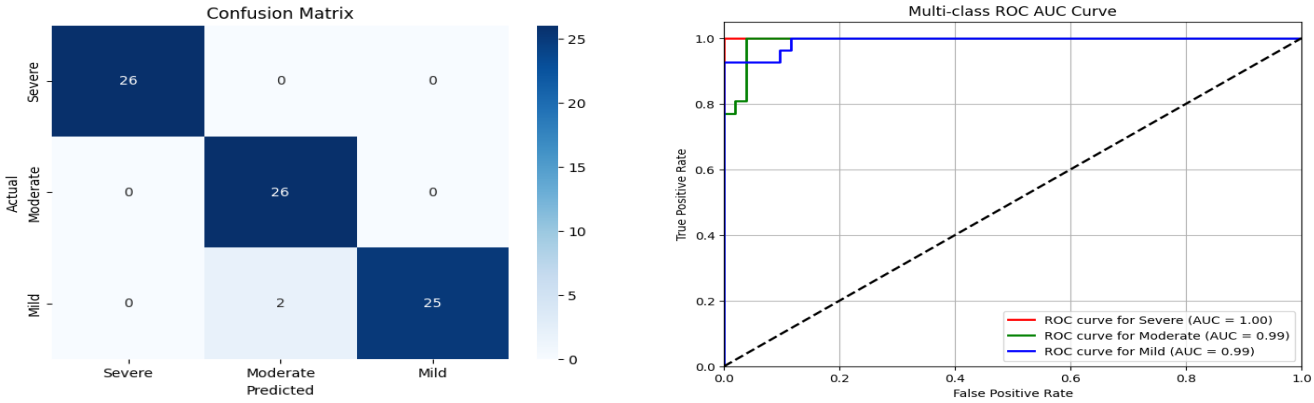
.



(a) 80:10:10 split ratio



(b) 75:15:10 split ratio



(c) 70:15:15 split ratio

Figure 5. Generated confusion matrices and ROC-AUC curves for different dataset splits.

DISCUSSION OF FINDINGS

This study aimed to develop a hybrid deep learning framework that integrates Bio-BERT embeddings with Bi-LSTM and CNN architectures to classify knee pain severity using structured clinical text from the Osteoarthritis Initiative (OAI) dataset. Model performance was evaluated under three train–validation–test split configurations (80:10:10, 75:10:15, and 70:15:15) to examine the effect of data partitioning on predictive accuracy. An initial analysis revealed substantial class imbalance, with a predominance of mild pain cases, which adversely affected the model's ability to accurately identify moderate and severe pain levels. To mitigate this limitation, data balancing techniques, including oversampling of minority classes, were applied prior to training. This intervention enhanced the model's capacity to learn discriminative features across all severity categories, resulting in more balanced, robust, and clinically meaningful classification performance.

Training the model with an 80:10:10 data split produced the strongest overall performance, attributed to the larger proportion of training data, which enabled robust learning of clinical patterns associated with knee pain severity. The model demonstrated consistently high validation and test performance, with minimal overfitting and strong precision and recall across all severity classes, including notably high recall for severe pain. Learning curves showed stable convergence, and confusion matrix analysis indicated very few misclassifications. Overall, this configuration yielded the most balanced and reliable results, achieving an accuracy of 0.98 and confirming the importance of a sufficiently large training set for effective deep learning performance.

Using the 75:10:15 split, the model maintained strong performance, though slightly below that of the 80:10:10 configuration, with a modest decline in accuracy and recall, most notably for the severe class, due to reduced training data. Nevertheless, it generalized well, achieving high precision and recall for the mild and moderate classes, and demonstrated stable validation loss without evidence of overfitting, resulting in an accuracy of 0.96. The 70:15:15 configuration, despite having the smallest training set, delivered competitive performance with an accuracy of 0.97. The larger validation and test sets enabled more reliable performance assessment, while the hybrid architecture remained effective across all classes, with only a slight reduction in severe-class recall. Overall, these results highlight the robustness and adaptability of the proposed model, which sustained high classification performance even under reduced training data conditions.

Overall, the findings highlight the critical roles of sufficient training data and effective class balancing in developing reliable predictive models. The initial class imbalance biased the model toward mild cases, but the application of oversampling techniques produced a more representative dataset, enabling improved discrimination across mild, moderate, and severe pain categories. The hybrid integration of Bio-BERT, Bi-LSTM, and CNN architectures leveraged complementary strengths in contextual, sequential, and spatial feature learning, resulting in robust and accurate knee pain severity classification. Among the evaluated data partitioning strategies, the 80:10:10 split demonstrated superior performance, yielding the highest accuracy, precision, and recall across all classes, particularly for severe pain detection. This configuration exhibited stable validation behaviour, minimal overfitting, and well-converged learning curves, underscoring its suitability as the most reliable and clinically relevant setup for the proposed model.

CONCLUSION

This study demonstrates that a hybrid deep learning framework integrating Bio-BERT, Bi-LSTM, and CNN architectures is an effective and innovative approach for predicting knee pain severity in osteoarthritis (OA) patients. By jointly leveraging unstructured clinical text and structured imaging data, the model overcomes key limitations of traditional statistical and conventional machine learning methods. Bio-BERT captures rich semantic information from clinical narratives, Bi-LSTM models temporal dependencies in patient data, and CNNs extract discriminative structural features from MRI images, resulting in a robust and clinically meaningful predictive system.

The model achieved consistently high performance across multiple data partitioning strategies, with strong accuracy, precision, recall, F1-score, and AUC-ROC values, confirming its reliability and generalizability. Its ability to accurately distinguish between mild, moderate, and severe pain levels highlights its potential utility in clinical

decision support, particularly for early intervention and personalized treatment planning. Although challenges such as data heterogeneity and multimodal integration remain, the findings establish a solid foundation for future refinement, external validation, and integration into electronic health record systems. Overall, the proposed hybrid approach represents a significant contribution to OA research and precision medicine, demonstrating the value of interdisciplinary, multimodal deep learning methods in advancing musculoskeletal healthcare.

REFERENCES

Guo, J., Zhang, Z., & Zhao, M. (2024). Deep learning methods for knee cartilage thickness grading in osteoarthritis assessment. *IEEE Journal of Biomedical and Health Informatics*. 56(2), 123-149. https://www.frontiersin.org/articles/10.3389/fmed.2024.1337993/full

Johnson, V. L., Hunter, D. J. (2019). The epidemiology of osteoarthritis. *Best Practice & Research Clinical Rheumatology*, 28(1), 5–15. https://doi.org/10.1016/j.berh.2014.01.004

Lazzarini, N., Runhaar, J., Agricola, R., & Bierma-Zeinstra, S. M. A. (2017). Machine learning algorithms applied to radiographs for the diagnosis of osteoarthritis: A systematic review. *Seminars in Arthritis and Rheumatism*, 47(5), 632-638. https://doi.org/10.1016/j.semarthrit.2017.08.005

Lin, Y., Teoh, Y. W., & Chen, M. (2023). Dynamic nomogram based on radiomics from MRI for prediction of knee pain improvement in osteoarthritis. *European Radiology, 33*, 2245–2256. https://doi.org/10.1109/ACCESS.2021.3062493

Liu, Q., Wang, Y., Lin, J., & Zhang, Y. (2022). Hybrid deep learning model using BioBERT and BiLSTM for biomedical text classification. *IEEE Journal of Biomedical and Health Informatics*, 26(2), 583-593. https://doi.org/10.1109/JBHI.2021.3078412

Morales, D. R., Duru, O. K., & Sharma, L. (2021). Imaging biomarkers of chronic pain in knee osteoarthritis: A data-driven analysis using quantitative MRI. *Pain Reports, 6*(3), e949. https://pubmed.ncbi.nlm.nih.gov/34753963/

Muhammad, L. J., Algehyne, E. A., & Mohammed, I. A. (2021). Ensemble machine learning methods for osteoarthritis severity prediction using radiograph and MRI data. *Health Information Science and Systems, 9*(1), 5. https://link.springer.com/article/10.1186/s13755-021-00168-2

Nguyen, H., Tran, Q., & Pham, B. (2020). Deep learning-based automated classification of knee osteoarthritis severity using semi-supervised learning. *Sensors, 20*(22), 6701. https://www.mdpi.com/1424-8220/20/22/6701

Osteoarthritis Research Society International (OARSI). (2020). Osteoarthritis: A serious disease, submitted to the U.S. Food and Drug Administration. *Osteoarthritis Research Society International (OARSI)*, 45(2), 35-67. https://oarsi.org/research/osteoarthritis-serious-disease-reports

Owoade, A. A., Kareem, M. A., Usman, O. L., Ogunbanwo, A. S., & Shosanya, S. O. (2025). An optimized method of brain tumor neural-biomarkers prediction using parallel 3D convolutional neural networks and contrast-enhanced normalized magnetic resonance imaging. *Dutse Journal of Pure and Applied Sciences, 11*(1d), 124–139. https://doi.org/10.4314/dujopas.v11i1d.13

Panfilov, E., Tiulpin, A., & Saarakkala, S. (2022). Deep learning-based progression prediction in knee osteoarthritis using structural MRI. *Computerized Medical Imaging and Graphics, 91*, 101963. https://www.sciencedirect.com/science/article/pii/S0895611122000089

Panwar, R., Kumar, S., & Patel, M. (2024). Deep stacked ensemble model for osteoarthritis severity prediction using MRI and radiographic data. *Artificial Intelligence in Health and Medicine Journal, 91*, 101963. https://pubmed.ncbi.nlm.nih.gov/34753963/

Pedoia, V., Majumdar, S., & Link, T. M. (2018). Automated segmentation and quantitative analysis of knee MRI using convolutional neural networks: Data from the Osteoarthritis Initiative. *Magnetic Resonance in Medicine*, 79(1), 128-137. https://doi.org/10.1002/mrm.26624

Smith, T., & Lee, J. (2021). Integration of clinical text and imaging data using NLP and deep learning: Application in osteoarthritis diagnosis. *Journal of Biomedical Informatics*, 115, 103690. https://doi.org/10.1016/j.jbi.2021.103690

Teoh, Y. W., Lin, Y., & Tan, M. (2023). Hybrid multitask deep learning for knee OA feature stratification using Osteoarthritis Initiative data. *Pattern Recognition Letters, 168*, 34–42. https://www.sciencedirect.com/science/article/pii/S016786552300473X

Thomas, K. A., Wilson, J. J., & Avery, A. (2020). Automated radiographic classification of knee osteoarthritis severity using deep neural networks. *Radiology: Artificial Intelligence, 2*(5), e200037. https://pubmed.ncbi.nlm.nih.gov/32280948/

Usman O.L, & Adeusi O.O. (2025). Optimization of an efficient net-based transfer learning model for automated pneumonia detection from chest x-ray images. *Dutse Journal of Pure and Applied Sciences*, *11*(4a), 397–411. Retrieved from https://www.ajol.info/index.php/dujopas/article/view/311486

Usman, O. L., Muniyandi, R. C., Omar, K., & Mohamad, M. (2021). Advanced machine learning methods for dyslexia biomarker detection: A review of implementation details and challenges. *IEEE Access*,*9*,36879–36894. https://doi.org/10.1109/ACCESS.2021.3062709

Usman, O. L., Muniyandi, R. C., Omar, K., Mohamad, M., Owoade, A. A., & Kareem, M. A. (2025). HoRNS-CNN model: An energy-efficient fully homomorphic residue number system convolutional neural network model for privacy-preserving classification of dyslexia neural-biomarkers. *Brain Informatics*, *12*(11), 1–28. https://doi.org/10.1186/s40708-025-00256-z