Journal of Artificial Intelligence and Data Science (JAIDA)
e-ISSN: 2791-8335
https://dergipark.org.tr/pub/jaida

Vol. 6, No. 1, pp. 22-30, 2026
Research Article
Received: 17/04/2026 Accepted: 25/06/2026

Dual-View Deep Learning for Tibial Plateau Fracture Detection Using Input-Level Fusion of X-Ray Images

Özlem Karabiber CURA[1,*], Fatma Günseli Yaşar ÇIKLAÇANDIR[2], Mehmet Göktuğ EFGAN[3]

Abstract
Tibial plateau fractures are serious injuries of the knee joint. If they are not diagnosed and treated on time, they can lead to important functional problems. Therefore, accurate detection from X-ray images is very important. In this study, a deep learning-based method is proposed for detecting tibial plateau fractures using dual-view X-ray images. Anterior–posterior and lateral images are combined at the input level to create a 6-channel representation. This simple design allows the model to learn from both views simultaneously. In experiments, five pre-trained convolutional neural network models (ResNet18, ResNet50, DenseNet201, EfficientNet-B0, and MobileNetV2) are used and evaluated under the same conditions. The results show that all models perform well, and EfficientNet-B0 achieved the best balance between sensitivity (87.23%) and specificity (94.23%) on the test set. In addition, Grad-CAM and t-SNE analyses are used to understand model decisions better. Overall, the proposed method provides a simple and effective approach for dual-view fracture detection and can support clinical diagnosis.
Keywords: Convolutional neural networks; deep learning; dual-view x-ray; grad-cam; input-level fusion; tibial plateau fracture.

1. Introduction

Tibial plateau fractures are among the most severe intra-articular injuries affecting the knee joint [1]. They typically occur due to high-energy trauma or low-energy mechanisms such as osteoporosis in elderly patients [2]. If not diagnosed and treated appropriately, they may lead to serious short- and long-term complications [3]. Accurate and timely diagnosis of tibial plateau fractures remains a critical challenge in emergency and orthopedic practice. Missed or delayed diagnosis may result in severe complications such as post-traumatic osteoarthritis [4,5], joint instability [6], and long-term functional impairment [7–9]. Moreover, variability in radiographic interpretation among clinicians further increases the risk of diagnostic errors, particularly in subtle or non-displaced fractures.

Recent advancements in artificial intelligence, particularly deep learning, have significantly improved automatic fracture diagnosis in radiographic images. Earlier studies primarily focused on detecting general bone fractures [10]. More recently, research has shifted toward clinically challenging and specific fracture types, such as tibial plateau fractures [11,12]. For instance, van der Gaast et al. developed CNN-based models for both detection and classification of tibial plateau fractures according to the Schatzker system, demonstrating the effectiveness of deep learning in identifying fracture presence [11]. In addition to classification-based approaches, object detection frameworks have also been explored to localize fracture regions. Huo et al. proposed a hybrid deep learning model integrating YOLOv8 and MobileNetV3, achieving approximately 91% accuracy, particularly for detecting occult fractures in multi-center datasets [12]. Similarly, Wang et al. evaluated multiple YOLO-based architectures trained on different radiographic views and showed that model performance varies depending on projection type, highlighting the importance of view selection in fracture detection [13]. Assink et al. developed machine learning models to predict the risk of total knee arthroplasty following tibial plateau fractures, demonstrating the broader applicability of artificial intelligence in clinical decision support [14].

Despite these advancements, effectively integrating multi-view radiographic information remains an open challenge. Existing approaches often rely on multi-branch architectures or feature-level fusion strategies [15], which increase model complexity and computational cost while requiring careful synchronization between views. However, in clinical practice, both anterior–posterior and lateral X-ray views are routinely used to improve diagnostic accuracy. Effectively leveraging complementary information from multiple radiographic views in a computationally efficient manner remains an open research problem.

*Corresponding author

ÖZLEM KARABIBER CURA*; Izmir Katip Celebi University, Faculty of Engineering And Architecture, Department of Biomedical Engineering, Türkiye: e-mail: ozlem.karabiber@ikcu.edu.tr 0000-0001-8650-1137

FATMA GÜNSELI YAŞAR ÇIKLAÇANDIR; Izmir Katip Celebi University, Faculty of Engineering And Architecture, Department Of Computer Sciences, Türkiye; e-mail: fatma.gunseli.yasar@gmail.com; 0000-0001-6182-7173

MEHMET GÖKTUĞ EFGAN; Izmir Katip Celebi University, Faculty of Medicine, Department of Internal Medical Sciences, Türkiye; e-mail: goktugefgan@gmail.com 0000-0002-0794-1239

To address these limitations, this study proposes a computationally efficient dual-view learning framework based on input-level fusion. By concatenating anterior–posterior and lateral X-ray images into a unified multi-channel representation, the proposed approach enables standard CNN architectures to learn cross-view relationships without requiring additional architectural complexity.

The main contributions of this study can be summarized as follows:

- A novel input-level fusion strategy for dual-view X-ray analysis that eliminates the need for complex multi-branch architectures.
- A systematic comparison of five state-of-the-art CNN architectures under identical experimental conditions.
- An explainability-driven analysis using Grad-CAM and t-SNE to provide insights into model decision-making.
- A statistical evaluation of projection-wise contributions using the Wilcoxon signed-rank test.,

The remainder of this paper is organized as follows. Section 2 describes the materials and methods, including dataset preparation, model architectures, and training procedures. Section 3 presents the experimental results along with performance evaluation and explainability analyses. Finally, Section 4 concludes the study and discusses potential future research directions.

2. Materials and Methods

In this work, a deep learning-based framework for automatically detecting tibial plateau fractures from dual-view X-ray images is proposed. Anterior and lateral projections are merged at the input level by creating a single 6-channel representation, as opposed to using distinct processing branches or explicit feature fusion techniques. To establish an impartial evaluation methodology, the dataset is randomly split into training, validation, and test subsets. Only the training and validation data are used for developing models, and the independent test set is used to report final performance. Several widely used convolutional neural network models are modified to accept the 6-channel input and trained under the same conditions. To further examine the learned feature representations and the relative contribution of each projection, feature space visualization and Grad-CAM-based explainability analyses are performed in addition to traditional performance evaluation metrics. The general framework of the proposed tibial plateau fracture detection approach is presented in Figure 1.

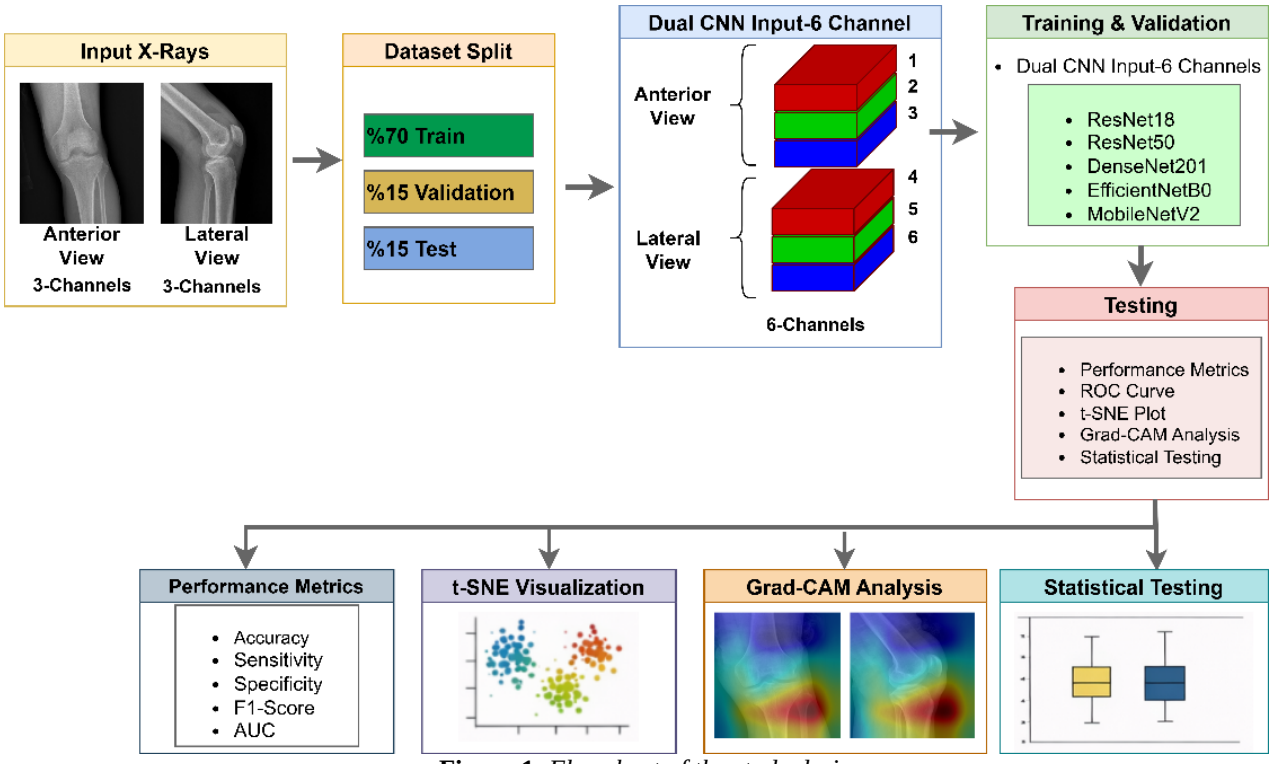

Figure 1. *Flowchart of the study design.*

2.1. Dataset and preprocessing

This study utilized retrospective tibial plateau radiographs obtained from patients who presented to the Emergency Department of Izmir Katip Celebi University Atatürk Training and Research Hospital between 2024 and 2026. Ethical approval for the study was granted by the Izmir Katip Celebi University Health Research Ethics Committee on February 12, 2026 (Decision No. 0103).

The dataset consisted of 656 patients, including 317 patients diagnosed with tibial plateau fractures and 339 healthy controls. Since each patient contributed one anteroposterior (AP) and one corresponding lateral (Lat) radiograph from the same examination, the total dataset comprised 1,312 X-ray images (656 AP and 656 Lat). Fracture labels are established based on clinical diagnoses and radiological evaluations recorded in the hospital information system by experienced emergency medicine specialists and orthopedic residents. Only radiographs with sufficient image quality and complete AP–Lat projection pairs are included in the study. Images with severe artifacts, incomplete anatomical coverage, duplicated examinations, or missing paired views are excluded. All images are anonymized and stored in .jpeg format before analysis. For the deep learning framework, AP and Lat radiographs belonging to the same patient are paired and treated as a single sample. Thus, the effective number of samples used for model development is equal to the number of patients (n = 656). The paired samples are randomly shuffled using a fixed random seed (rng(42)) and divided into mutually exclusive training (70%), validation (15%), and test (15%) subsets. To preserve the AP–Lat correspondence, identical indices are applied to both image folders during data partitioning. Consequently, both radiographic views from a given patient are always assigned to the same subset, preventing patient overlap and eliminating the possibility of data leakage between training, validation, and test sets

2.2. Deep learning models

Five pre-trained convolutional neural network (CNN) architectures are utilized: ResNet18, ResNet50, DenseNet201, EfficientNet-B0, and MobileNetV2.

The residual learning framework, which uses skip connections to enable effective training of deep models by learning residual mappings rather than direct transformations, is the foundation of the convolutional neural network architectures ResNet18 and ResNet50. With its 18 layers and simple residual blocks composed of two 3×3 convolutional layers, ResNet18 is a relatively shallow network that is less likely to overfit, especially with smaller datasets. ResNet50, on the other hand, is a deeper design with 50 layers that uses bottleneck residual blocks, where dimensionality is reduced and subsequently restored using 1×1, 3×3, and 1×1 convolutions to enable more intricate and hierarchical feature extraction. Because of its greater depth, ResNet50 has greater representational power than ResNet18, which delivers quicker training and cheaper computing [16,17]. Apart from ResNet architectures, a variety of deep learning models with different design principles and computational complexity are also used, including DenseNet201, EfficientNet-B0, and MobileNetV2. Dense connection, the foundation of DenseNet201, allows each layer to get feature maps from all previous layers, encouraging feature reuse and enhancing gradient flow. This leads to good performance but at a comparatively high computational cost [16,18]. In contrast to conventional deep networks, EfficientNet-B0 achieves a solid balance between accuracy and efficiency with fewer parameters by using a compound scaling strategy that equally scales network depth, width, and resolution [18]. Using depthwise separable convolutions and inverted residual blocks with linear bottlenecks, MobileNetV2's computation is lightweight and effective, making it ideal for settings with limited resources [19]. These architectures are modified to accept 6-channel inputs to train joint representations from AP and Lat X-ray images. This allowed for a thorough comparison of the classification performance of heavy, balanced, and lightweight models.

2.3 Network adaptation for dual-view learning

To incorporate complementary information from both radiographic projections, AP and Lat X-ray images belonging to the same patient are paired and merged at the input level. Specifically, each AP image occupied channels 1–3, while the corresponding Lat image occupied channels 4–6, resulting in a unified six-channel input representation of size 224 × 224 × 6. Unlike conventional multi-view architectures that employ separate feature extraction branches followed by feature-level, attention-based, or decision-level fusion mechanisms, the proposed framework integrates both views before feature extraction, allowing the network to jointly learn view-specific and complementary anatomical characteristics from the earliest layers of the model.

Since all evaluated CNN architectures are originally pre-trained on ImageNet using three-channel RGB images, their input layers are modified to accept six-channel inputs. For each backbone network, the original first

convolutional layer weights are adapted by duplicating and scaling the pre-trained filters across six channels. Specifically, the original three-channel convolutional weights are divided equally and copied into the AP and Lat channel groups, thereby preserving the initialization learned from ImageNet while enabling simultaneous processing of both radiographic projections. This strategy allows the model to retain the advantages of transfer learning without requiring random initialization of additional channels. The final fully connected and classification layers of each network are replaced to match the number of target classes in the tibial plateau fracture dataset. A softmax activation function is subsequently used to generate class probabilities.

2.4. Training configuration and performance evaluation

All models utilized the Adam optimizer, with a mini-batch size of 32, trained for a maximum of 20 epochs, and had an initial learning rate of 1e-4. Data is shuffled every epoch. Validation data is employed to monitor performance and prevent overfitting. To ensure a fair comparison among architectures, identical training, validation, and testing partitions, as well as the same optimization settings, are used for all evaluated models.

Model performance is evaluated on both validation and test sets, with final performance reported solely on the test set to avoid bias. Several metrics—accuracy (ACC), sensitivity (SEN), specificity (SPE), precision (PRE), and area under the ROC curve (AUC)—are computed, and confusion matrices along with ROC curves are generated for further analysis [10,19].

Using the t-distributed stochastic neighbor embedding (t-SNE) method, we display the two-dimensional feature maps to exhibit the capabilities of the extracted features from one of the top-performing models [18]. This approach reveals class separability in the learned feature space.

Grad-CAM usually generates heatmaps in a supervised manner by propagating gradients from the target class score back to the network's final convolutional layer. By highlighting the areas that significantly affect the model's prediction for the target class, these heatmaps offer coarse localization [20,21]. In our proposed study, Grad-CAM, which computes activation maps from the final convolutional layers for successfully categorized test data, is used to evaluate model decisions. The analysis comprised calculating mean activation scores for both projections (μ_{AP}), and (μ_{LAT}), weighting the Grad-CAM activation maps by view, and dividing the 6-channel input into AP and Lat components.

To determine whether there are statistically significant differences between the extracted deep features of groups, the Wilcoxon signed-rank test is utilized with a significance level of 0.05. This non-parametric alternative for the paired t-test is ideal for comparing paired observations because it does not assume normality of differences [21]. In our experiment, a Wilcoxon signed-rank test is applied to Grad-CAM activation scores to determine whether AP and Lat contributions differed significantly. The mean activation values and effect size (ES, "Cohen's d") are then computed.

$$ES = \frac{\mu_{AP}+\mu_{Lat}}{\sigma_{AP-Lat}} \quad (1)$$

3. Experimental Results and Discussions

The multi-view deep learning framework for tibia plateau fracture detection is evaluated in this study. Standard classification criteria are used to evaluate the performance of different CNN models on both separate test and validation datasets. The relative contributions of anterior and lateral projections to the model predictions are examined using a Grad-CAM-based analysis in addition to standard performance evaluation. In addition to comparing model performance, this evaluation approach provides information about the decision-making behavior of different architectures.

Table 1. *Performance evaluation results of pre-trained CNN architectures for validation and test datasets.*

Model	Validation					Test				
	ACC	**SEN**	**SPE**	**PRE**	**AUC**	**ACC**	**SEN**	**SPE**	**PRE**	**AUC**
ResNet18	94.90	93.88	95.92	95.83	97.63	90.91	85.11	96.15	95.24	97.50
ResNet50	94.90	93.88	95.92	95.83	97.48	90.91	85.11	96.15	95.24	97.46
DenseNet201	95.92	93.88	97.96	97.87	97.13	90.91	82.98	98.08	97.50	97.67
EfficientNet-B0	**96.94**	**95.92**	**97.96**	**97.92**	**97.46**	**90.91**	**87.23**	**94.23**	**93.18**	**95.95**
MobilenetV2	95.92	93.88	97.96	97.87	98.67	90.91	85.11	96.15	95.24	98.04

Table 1 displays the quantitative performance of the examined CNN models. With validation ACC ranging from 94.90% to 96.94% and test ACC is 90.91, all models achieved high classification performance on both validation and test sets. With the greatest test AUC (95.95%) and balanced sensitivity (87.23%) and specificity (94.23%), EfficientNet-B0 outperformed the other models that are assessed. ResNet18 and ResNet50 produced similar outcomes, with consistent generalization on the test set and marginally lower AUC values (97.48% and 97.46% in validation, respectively). DenseNet201 demonstrated a bias toward negative class prediction with good specificity (98.08%) but relatively poor sensitivity (82.9%). While MobileNetV2 performed competitively, especially in validation (AUC: 98.67%), its generalization is marginally worse than that of EfficientNet-B0.

Table 2. *Statistical comparison of AP and Lat projection contributions based on Grad-CAM activation scores.*

Model	μ_{AP}	μ_{LAT}	ES	p-value
ResNet18	0.097	0.091	0.195	0.034
ResNet50	0.072	0.069	0.093	0.625
DenseNet201	0.097	0.092	0.194	0.056
EfficientNet-B0	0.062	0.061	0.071	0.761
MobilenetV2	0.108	0.101	0.198	0.043

A Grad-CAM-based statistical analysis is carried out to further examine the influence of various imaging projections, as shown in Table 2. The average activation scores show that, in every model, the AP projection produced higher activation values than the Lat projection. Nevertheless, only ResNet18 (p = 0.034) and MobileNetV2 (p = 0.043) showed statistically significant variations across projections. While ResNet50 (p = 0.625) and EfficientNet-B0 (p = 0.761) do not exhibit any statistically significant differences, DenseNet201 revealed a marginal significance (p = 0.056). These results are further supported by the calculated effect sizes (ES, "Cohen's d"), which show that all models have small effect sizes (varying from 0.071 to 0.198). Interestingly, the biggest effect sizes are shown by MobileNetV2 (d = 0.198) and ResNet18 (d = 0.195), suggesting a comparatively greater dependence on projection-specific properties. EfficientNet-B0, on the other hand, showed the smallest effect size (d = 0.071), indicating a more equal use of data from both projections.

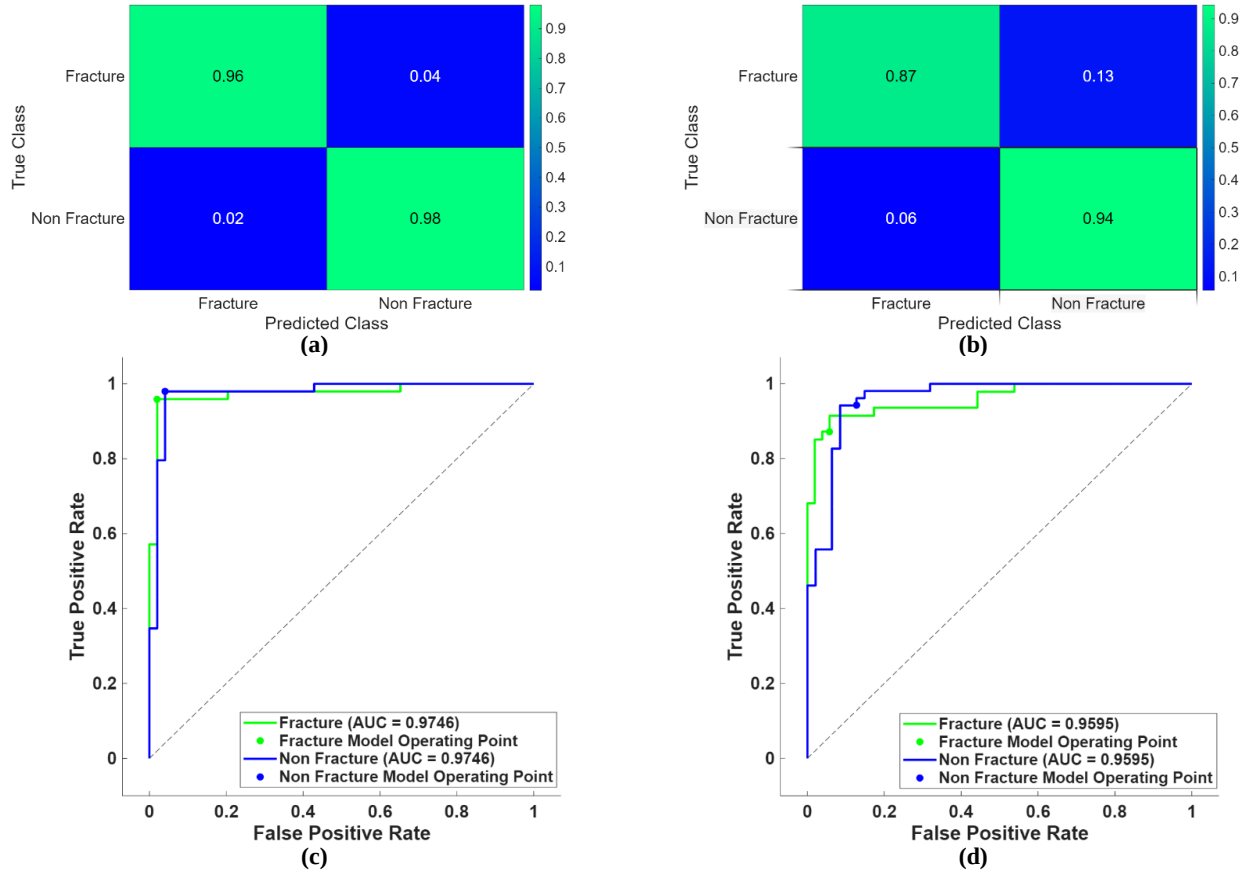

Figure 2. *Confusion matrices (a, b) and Receiver Operating Characteristic (ROC) curves (c, d) for the EfficientNet-B0 model. Figures (a) and (c) represent the validation dataset performance, while (b) and (d) illustrate the test dataset performance.*

Confusion matrices and ROC curves for both validation and test datasets are used in Figure 2 to demonstrate the classification performance of the EfficientNet-B0 model. Strong sensitivity and specificity are demonstrated by the model's high accuracy on the validation set, as seen in Figure 2(a), which properly classifies 96% of fracture instances and 98% of non-fracture cases. Despite a minor decline in comparison to validation findings, the model continues to perform robustly on the independent test set (Figure 2(b)), accurately identifying 87% of fracture instances and 94% of non-fracture cases, indicating high generalization capability. The ROC curves in Figures 2(c) and (d) provide additional evidence of the model's discriminative ability. With an AUC of 0.9746 on the validation set and 0.9595 on the test set, the EfficientNet-B0 demonstrates exceptional class separability. The model successfully balances true positive and false positive rates across various thresholds, as seen by the ROC curves' closeness to the upper-left corner. Overall, Figure 2 shows that the suggested model retains consistent and dependable performance when tested on unobserved data in addition to achieving high classification accuracy.

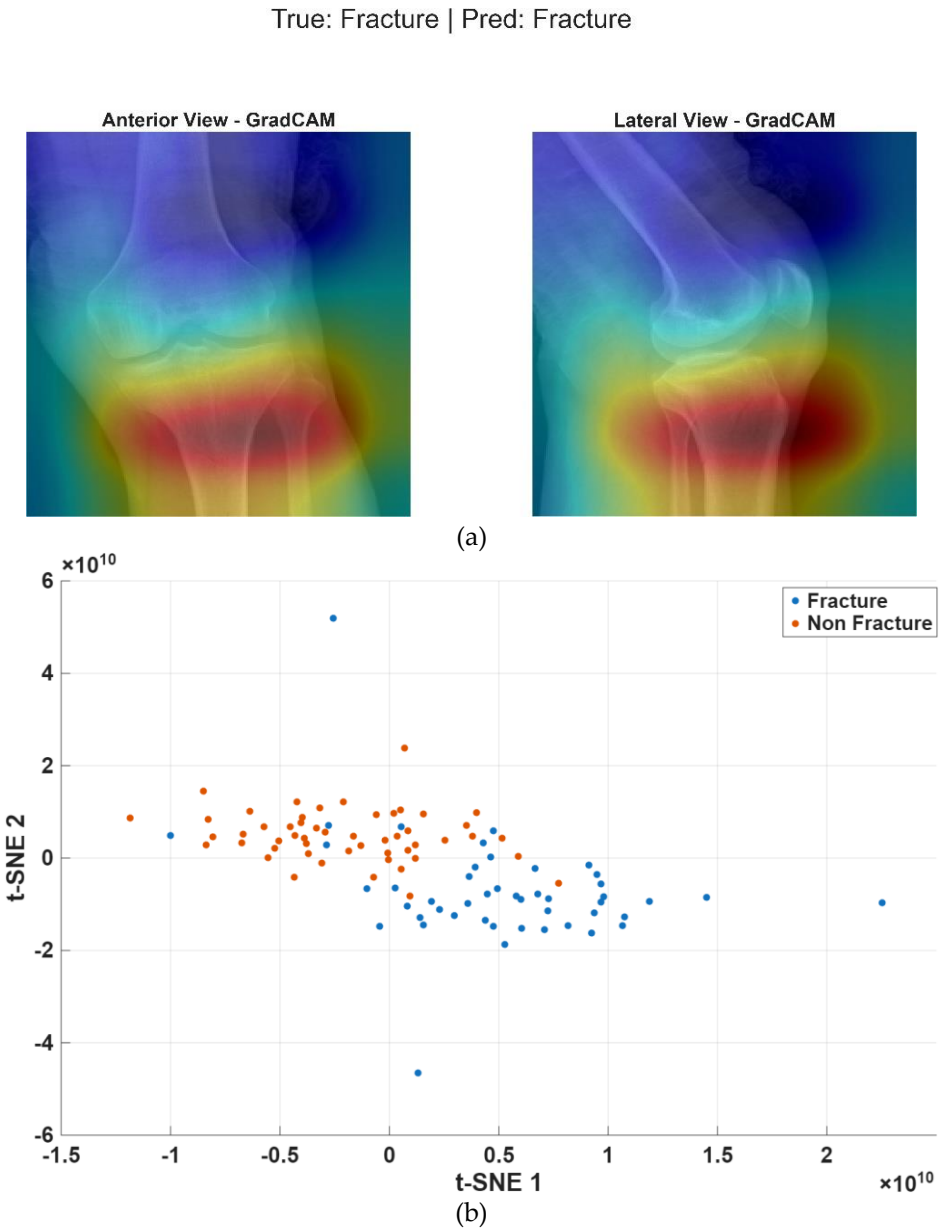

Figure 3. *Tibial plateau fracture detection Model interpretability and feature distribution. (a) Grad-CAM visualizations that highlight the proximal tibia in both anterior and lateral X-ray views. (b) t-SNE scatter plot showing how the high-dimensional latent space's fractured and non-fractured features are clustered.*

Gradient-weighted Class Activation Mapping (Grad-CAM) was used to examine the image regions that contributed most to the model predictions. In both the AP and Lat views, the model appropriately concentrates its "attention" toward the proximal tibia (the tibial plateau region), as seen in Figure 3 (a). The heatmaps show that

the dual-view method enables the model to efficiently capture 3D spatial context for diagnostic confirmation, with significant activation intensity (red zones) focused around the fracture site. Additionally, using t-SNE to analyze the model's internal representations, "Fracture" and "Non-Fracture" clusters are clearly separable (given in Figure 3 (b)). A few outliers close to the decision border may suggest minor or non-displaced fractures, but the EfficientNet-B0 backbone effectively learns distinct morphological signatures for tibial plateau fractures. Notably, the accuracy and biological significance of the model in detection are reflected in the connection between Grad-CAM heatmaps and t-SNE clustering.

In addition to the classification performance achieved in this study, the proposed dual-view framework offers several methodological advantages. Existing multi-view fracture detection studies commonly rely on feature-level fusion, decision-level fusion, or attention-based fusion mechanisms to combine information from different radiographic projections. In contrast, the proposed approach employs an early input-level fusion strategy by integrating AP and Lat radiographs into a unified six-channel representation before feature extraction. This design allows complementary anatomical information from both projections to be jointly processed from the first convolutional layer onward, potentially facilitating the learning of fracture-related features from both projections. Furthermore, unlike more complex fusion architectures, the proposed method maintains a relatively simple network structure, requires no additional fusion modules, and can be readily adapted to a variety of convolutional neural network backbones. The favorable performance obtained across multiple CNN architectures suggests that early fusion of paired AP and Lat radiographs represents a practical and computationally efficient strategy for automated tibial plateau fracture detection.

4. Limitations and Future Work

Several limitations of this study should be acknowledged.

- Although the dataset included 656 patients and 1,312 radiographs, all images are collected retrospectively from a single medical center. Consequently, the imaging protocols, patient population, and acquisition conditions may not fully represent the variability encountered in different hospitals and clinical environments. This may limit the generalizability of the proposed framework to broader patient populations.
- External validation using an independent dataset from another institution is not performed. While the proposed dual-view approach demonstrated high performance on the internal test set, its robustness across different imaging devices, healthcare centers, and demographic populations remains to be investigated. Future studies should evaluate the model on multi-center datasets to better assess its clinical applicability and generalizability.
- The ground-truth labels are derived from clinical diagnoses and radiological evaluations documented in the hospital information system. Although these diagnoses are established by experienced physicians, a dedicated multi-reader consensus study involving orthopedic surgeons and musculoskeletal radiologists has not been conducted. Future studies may further strengthen the reliability of the reference standard through expert consensus-based annotation protocols.
- The proposed framework utilizes only AP and Lat radiographs. In clinical practice, additional imaging modalities such as computed tomography (CT) and magnetic resonance imaging (MRI) may provide complementary information regarding fracture morphology and severity. Integrating multi-modal imaging data may further improve diagnostic performance and clinical decision support capabilities.
- The study focused on an input-level fusion strategy that combines AP and Lat projections into a six-channel representation. Although this approach demonstrated promising results while maintaining architectural simplicity, other fusion strategies, including feature-level fusion, attention-based fusion, and decision-level fusion, have not been investigated. Future research may compare different fusion paradigms to determine the most effective strategy for multi-view fracture assessment.
- Despite the encouraging performance achieved by the evaluated convolutional neural network architectures, prospective validation studies and real-world clinical deployment are beyond the scope of this work. Future investigations should include prospective multi-center studies, larger patient cohorts, and external validation datasets to confirm the robustness, reliability, and clinical utility of the proposed dual-view deep learning framework.

5. Conclusions

A multi-view deep learning framework for automatic tibia plateau fracture detection employing anterior and

lateral X-ray views is presented in this work. Several state-of-the-art CNN architectures are thoroughly examined, and both traditional classification metrics and explainability-driven analyses are used to assess their performance. EfficientNet-B0 outperformed the other models in the evaluation, showing excellent generalization ability with high accuracy values on both the independent test and validation datasets.

In addition to performance evaluation, the impact of each forecast on the decision-making process is examined using a Grad-CAM-based methodology. The findings showed that although multi-view input is advantageous for all models, there are variations in determining whether they use complementary information. In particular, EfficientNet-B0 showed a more balanced integration of AP and Lat views, whereas lighter frameworks such as ResNet18 and MobileNetV2 showed a statistically significant dependence on a dominating view. These results demonstrate that higher-capacity models are more adept at fusing multi-view data, which results in predictions that are more reliable and understandable, in addition to being more accurate.

Overall, the proposed architecture suggests that combining multi-view imaging with deep learning may improve fracture detection performance. Additionally, combining statistical analysis with explainability methodologies yields insightful information about model behavior that aids in the creation of more dependable and clinically significant decision-support systems. To further enhance multi-view feature fusion and interpretability, future research will concentrate on expanding this method to bigger and more varied datasets, adding new imaging modalities, and investigating sophisticated multi-branch structures.

Declaration of Interest

The authors declare that there is no conflict of interest.

Acknowledgements

This study was supported by the Izmir Katip Celebi University Scientific Research Projects Coordination Unit. Project number 2025-KDP-MÜMF-0003).

References

[1] S. E. Karapinar, S. Korkmaz, and R. Dincer, “An Evaluation of Intra-articular Pathologies Accompanying Tibial Plateau Fractures Postoperatively,” Cureus, 16(10), 2024.

[2] J. L. Marsh, “Tibial plateau fractures,” In Management of Fractures in Severely Osteoporotic Bone: Orthopedic and Pharmacologic Strategies (pp. 296-308), 2000. London: Springer London.

[3] D. J. Haslhofer, N. Kraml, P. W. Winkler, T. Gotterbarm, and A. Klasan, “Risk for total knee arthroplasty after tibial plateau fractures: a systematic review,” Knee Surgery, Sports Traumatology, Arthroscopy, 31(11), 5145-5153, 2023.

[4] T. Resch, F. Hartz, L. Faber, P. Zehnder, G. Römmermann, A. Ellafi, and F. Greve, “Low rate of secondary interventions for post-traumatic osteoarthritis and satisfactory mid-to-long-term outcomes following tibial plateau fractures,” BMC Musculoskeletal Disorders, 26(1), 427, 2025.

[5] N. Manidakis, A. Dosani, R. Dimitriou, D. Stengel, S. Matthews, and P. Giannoudis, “Tibial plateau fractures: functional outcome and incidence of osteoarthritis in 125 cases,” International orthopaedics, 34(4), 565-570, 2010.

[6] P. J. Papagelopoulos, A. A. Partsinevelos, G. S. Themistocleous, A. F. Mavrogenis, D. S. Korres, and P. N. Soucacos, “Complications after tibia plateau fracture surgery,” Injury, 37(6), 475-484, 2006.

[7] R. L. M. Van Dreumel, B. P. W. Van Wunnik, L. Janssen, P. C. G. Simons, and H. M. J. Janzing, “Mid-to long-term functional outcome after open reduction and internal fixation of tibial plateau fractures,” Injury, 46(8), 1608-1612, 2015.

[8] C. Neidlein, J. Watrinet, R. Pätzold, D. P. Berthold, W. C. Prall, W. Böcker, and Bormann, M., “Patient-Reported outcomes following tibial plateau fractures: Mid-to Short-Term implications for knee function and activity level,” Journal of Clinical Medicine, 13(8), 2327, 2024.

[9] A. S. García, D. G. Quevedo, M. S. Casado, D. M. Garceso, L. de la Puerta Migueles, D. M. Molinero, and I. Tamimi, “The functional outcome and complication rates in tibial plateau fractures: A comparison between three different classification systems,” The Knee, 55, 214-225, 2025.

[10] K. D. Ahmed, and R. Hawezi, “Detection of bone fracture based on machine learning techniques,” Measurement: Sensors, 27, 100723, 2023.

[11] N. Van der Gaast, P. Bagave, N. Assink, S. Broos, R. L. Jaarsma, R. L., M. J. R Edwards, and Machine Learning Consortium, “Deep learning for tibial plateau fracture detection and classification,” The Knee, 54, 81-89, 2025.

[12] T. Huo, P. Liu, M. Xue, J. Zhang, Y. Xie, H. Wang, Z. Ye, “Deep learning diagnosis of adult tibial plateau fractures: multicenter study with external validation,” Radiology Advances, 2(3), 2025.

[13] S. P. Wang, H. T. Shih, Y. X. Liao, C. H. Wei, J. C. Liu, E. Kristiani, and C. T. Yang, “On Construction of Tibial Plateau Fracture Detection in Different Radiographic Views Using YOLO Models,” Diagnostics, 16(2), 182, 2026.

[14] N. Assink, M. P. Gonzalez-Perrino, R. Santana-Trejo, J. N. Doornberg, H. Hoekstra, J. Kraeima, and F. F. IJpma, "Development of machine learning–based algorithms to predict the 2-and 5-year risk of TKA after tibial plateau fracture treatment," Clinical Orthopaedics and Related Research®, 483(9), 1731-1743, 2025.

[15] S. Abd El-Ghany, M. A. Mahmood, and A. A. Abd El-Aziz, "FracFusionNet: A Multi-Level Feature Fusion Convolutional Network for Bone Fracture Detection in Radiographic Images," *Diagnostics*, *15*(17), 2212, 2025.

[16] A. Subasi, "Applications of artificial intelligence in medical imaging," Academic Press, 2022.

[17] A. W. Salehi, S. Khan, G. Gupta, B. I. Alabduallah, A. Almjally, H. Alsolai, A. Mellit, A., "A study of CNN and transfer learning in medical imaging: Advantages, challenges, future scope," Sustainability, 15(7), 5930, 2023.

[18] A. Mahbod, N. Saeidi, S. Hatamikia, and R. Woitek, "Evaluating pre-trained convolutional neural networks and foundation models as feature extractors for content-based medical image retrieval," Engineering Applications of Artificial Intelligence, 150, 110571, 2025.

[19] A. Souid, N. Sakli, H. Sakli, "Classification and predictions of lung diseases from chest x-rays using mobilenet v2," Applied Sciences, 11(6), 2751, 2021.

[20] Z. Zhao, H. Chen, Y. P. Wang, D. Meng, Q. Xie, Q. Yu, and L. Wang, "Retinal disease diagnosis with unsupervised Grad-CAM guided contrastive learning," Neurocomputing, 593, 127816, 2024.

[21] B. M. Malau, and M. O. Olusanya, "Comparative Evaluation of Parallel and Sequential Hybrid CNN–ViT Models for Wrist X-Ray Anomaly Detection," Applied Sciences, 15(22), 11865, 2025.