
Düzce University
Journal of Science & Technology

Research Article

# Classification of Knee X-rays That Can Be Diagnosed Radiographically Using Deep Learning and Machine Learning Methods

Kemal ÜRETEN [a,*], Semra DURAN [b], Yüksel MARAŞ [c], Ebru ATALAR [c], Kevser ORHAN [c], Hadi Hakan MARAŞ [d]

[a] *Department of Computer Engineering, PhD, Çankaya University, Ankara, TURKEY*
[b] *Department of Radiology, Ankara City Hospital, Ankara, TURKEY*
[c] *Department of Rheumatology, Ankara City Hospital, Ankara, TURKEY*
[d] *Department of Computer Engineering, Faculty of Engineering, Çankaya University, Ankara, TURKEY*
* *Corresponding author's e-mail address: kemalureten@yahoo.com*
DOI: 10.29130/dubited.1626406

ABSTRACT

The aim of this study is to classify knee osteoarthritis, synovial chondromatosis, Osgood-Schlatter disease, os fabella pathologies that can be diagnosed with plain knee X-rays, and normal knee radiographs with deep learning and machine learning methods. This study was performed on 540 knee osteoarthritis, 151 Osgood_Schlatter disease, 191 knee chondromatosis, 152 os fabella and 523 normal knee X-ray images. First, classification was performed with the VGG-16 network, which is a pre-trained deep learning model. Then, the features extracted with the VGG-16 convolution layer were classified with random forest, support vector machines, logistic regression and decision tree machine learning algorithms. With VGG-16 model, 95.3% accuracy, 95.1% sensitivity, 98.7% specificity, 96.8% precision, and 95.9% F1 score results were obtained. In classifying the features extracted from the VGG-16 convolution layer with machine learning algorithms, 98.2% accuracy, 99.0% sensitivity, 98.9% specificity, 98.2% precision and 98.5% F1 score results were obtained with the logistic regression classifier. In this study, which was conducted to classify radiographically detectable knee pathologies, successful results were obtained with the VGG-16 network. The features extracted from the convolution layer of the VGG-16 model were reclassified with machine learning algorithms, logistic regression, support vector machines and random forest classifiers, and improvements in performance metrics were obtained compared to the VGG-16 model. With this proposed method, the performance of deep learning models can be further improved.

**Keywords:** *Knee osteoarthritis, knee chondromatosis, osgood-schlatter disease, Os fabella, deep learning, machine learning*

Received: 24/01/2025, Revised: 25/05/2025, Accepted: 26/05/2025

## Radyografik Olarak Tanı Konulabilen Diz Röntgenlerinin Derin Öğrenme ve Makine Öğrenmesi Yöntemleri ile Sınıflandırılması

ÖZ

Bu çalışmanın amacı, düz diz röntgenleriyle tanısı konulabilen diz osteoartriti, sinovyal kondromatozis, Osgood-Schlatter hastalığı, os fabella patolojileri ve normal diz radyografilerini derin öğrenme ve makine öğrenmesi yöntemleriyle sınıflandırmaktır. Bu çalışma 540 diz osteoartriti, 151 Osgood_Schlatter hastalığı, 191 diz kondromatozisi, 152 os fabella ve 523 normal diz röntgen görüntüsü üzerinde gerçekleştirildi. Öncelikle önceden eğitilmiş derin öğrenme modeli olan VGG-16 ağı ile sınıflandırma yapıldı. Daha sonra VGG-16 evrişim katmanı ile çıkarılan özellikler, rastgele orman, destek vektör makineleri, lojistik regresyon ve karar ağacı makine öğrenmesi algoritmalarıyla sınıflandırıldı. VGG-16 modeli ile %95,3 doğruluk, %95,1 duyarlılık, %98.7 özgüllük, %96,8 kesinlik ve %95,9 F1 skoru sonuçları elde edildi. VGG-16 evrişim katmanından çıkarılan özelliklerin makine öğrenmesi algoritmaları ile sınıflandırılmasında lojistik regresyon sınıflandırıcısı ile %98,2 doğruluk, %99,0 duyarlılık, %98.9 özgüllük, %98,2 kesinlik ve %98,5 F1 skoru sonuçları elde edilmiştir. Radyografik olarak tanısı konulabilen diz patolojilerinin sınıflandırılması amacıyla yapılan bu çalışmada, VGG-16 ağı ile başarılı sonuçlar elde edilmiştir. VGG-16 modeli evrişim katmanı üzerinden çıkarılan özellikler makine öğrenmesi algoritmaları ile yeniden sınıflandırılmış, lojistik regresyon, destek vektör makineleri ve rastgele orman sınıflandırıcıları ile VGG-16 modeline kıyasla performans metriklerinde iyileşmeler elde edilmiştir. Önerilen bu yöntemle, derin öğrenme modellerinin performansı daha da iyileştirilebilir.

***Anahtar Kelimeler****: Diz osteoartriti, diz kondromatozu, osgood-schlatter hastalığı, Os fabella, derin öğrenme, makine öğrenmesi*

# I. INTRODUCTION

Knee osteoarthritis (OA), synovial chondromatosis, Osgood-Schlatter disease and os fabella, which may cause pain and limitation of movement in the knees, are pathologies that can be recognized by plain knee radiographs. Knee osteoarthritis is an important cause of knee pain as well as limitation of movement and disability. Its incidence is increasing due to the increase in average life expectancy and the increase in the prevalence of obesity in societies. The patient's quality of life can be improved with weight loss, appropriate exercise programs and physical therapy modalities [1, 2]. The prevalence of knee osteoarthritis increases with age, is slightly more common in women, and a study found that the frequency of knee OA over the age of 60 was 37.4% [3].

Synovial chondromatosis is a rare, benign pathology of unknown etiology. The most commonly affected joints are the knees. It has a relatively non-specific presentation and, if symptomatic, occurs with insidious onset findings such as decreased range of motion, pain, swelling, knee crepitus or knee locking. Plain radiographs may be diagnostic and mineralized nodules are pathognomonic [3, 4]. Synovial chondromatosis is a rare pathology, seen in 1 in 100,000 [5].

Os fabella is a fibrocartilaginous or ossified sesamoid bone usually located in the lateral tendon of the gastrocnemius muscle. The os fabella is usually asymptomatic, but sometimes it can be associated with many pathological conditions as well as pain in different age groups [6, 7]. Os Fabella is more common in both knees than unilateral, and the prevalence rates vary from 3% to 87% [8].

Osgood-Schlatter disease (OSD), caused by inflammation of the patellar tendon at its attachment point proximal to the tibia, is common in active adolescents and is a common cause of knee pain. OSD may also be a cause of knee pain after adolescence. The prevalence in children aged 12 to 15 years was reported as 9.83%, more common in boys [9-11].

These knee pathologies can be diagnosed with antero-posterior and lateral knee radiographs. X-ray imaging is an easily accessible, cost-effective, easy and fast imaging modality. Computed tomography (CT) and magnetic resonance imaging (MRI) methods are not available in every center, especially in rural areas, and CT has a radiation risk, MRI is an expensive method and is not suitable for every

patient. X-ray imaging protocols are well-structured and standardized. For this reason, skeletal imaging usually begins with plain X-rays.
In recent years, many successful studies have been carried out in the medical field, especially in image processing, with deep learning and machine learning methods. Convolutional neural networks (CNN), a deep learning method, are very successful in object detection, object tracking and image classification [12-15]. The CNN structure includes some convolution layers, pooling layers and fully connected layers and activation functions such as sigmoid, softmax and Rectified Linear Unit (ReLU). In convolution layers, features are extracted with the help of filters applied to the image. In pooling layers, subsampling is applied on the image size to reduce the spatial dimension, further reducing the number of parameters required for computation. Fully connected layers use the features learned through previous convolution layers and the classification task is performed [16].
Transfer learning method is applied to improve the performance of the CNN model and to achieve better generalization ability of the model. Medical images are created with special equipment such as X-ray machine, CT machine and MRI and labeled by experienced medical professionals. Therefore, collecting sufficient training data is expensive and difficult in most cases. When there is not enough data for training, medical image analysis can be done with the help of transfer learning method. The ImageNet dataset contains more than 20,000 categories and more than 14 million labeled images [16], [17]. The model trained with ImageNet can be used for transfer learning in a related, similar task. In this way, successful results can be obtained with less data in areas where there is not enough labeled data or where it is difficult to obtain data [18]. Some of the pre-trained networks are AlexNet [19], GoogLeNet [20], VGG-16 [21], ResNet-50 [22] and MobileNet [23]. Many successful studies have been conducted using transfer learning with medical images using CNNs [24]. There are many studies performed on plain X-ray images, computed tomography CT) images and magnetic resonance imaging (MRI) images [25, 26].
Plain X-ray imaging is the most commonly used imaging method in the diagnosis of knee OA. Grading is done with the Kellgren-Lawrence grading scheme. Some studies have been done for knee OA diagnosis and grading with CNNs. Tuilpin A et al. studied with the Multicenter Osteoarthritis Study (MOST) dataset and Osteoarthritis Initiative (OAI) dataset for knee OA diagnosis, graded with the Deep Siamese CNN model and the Kellgren-Lawrence grading system, and obtained an average multi-class accuracy of 66.71% [27]. Wang Y et al. used the Osteoarthritis Initiative (OAI) database in their study, worked with 4506 samples, applied the YOLO algorithm for knee joint area detection, and graded with the Kelgren-Lawrence grading system and achieved 69.18% classification accuracy [28]. In their study to classify knee OA, Lui B et al. used the Osteoarthritis Initiative (OAI) database, graded with the Kellgren-Lawrence grading system, and achieved 71.1% accuracy with faster RCNN [29].
The aim of this study is to classify radiographically detectable knee OA, Osgood-Schlatter disease, synovial chondromatosis, os fabella, and normal knee X-ray images using deep learning and machine learning methods. It was aimed to develop a machine learning model that physicians who do not have sufficient experience in evaluating plain knee radiographs can use in their daily work.

# II. MATERIAL AND METHOD

## A. DATASET

This retrospective study was conducted using picture archiving and communication systems (PACS) radiographs from Ankara Bilkent City Hospital Radiology Department. An ethical approval certificate was obtained from the local ethic commitee. Knee radiographs taken between January-2020 and May-2023 were scanned retrospectively by a radiologist experienced in skeletal radiographs. Radiographs detecting knee OA, Osgood-Schalatter disease, synovial chondromatosis and os fabella were recorded with patient age and gender. The demographic characteristics of the patients are shown in Table 1. The dataset included 540 knee osteoarthritis, 151 Osgood_Schlatter, 191 knee chondromatosis, 152 os fabella and 523 normal knee X-ray images.

***Table 1.*** *Demographic characteristics of patients*

| | Female | Male | Total | Mean age |
|---|---|---|---|---|
| Osgood-Schlatter | 36 | 115 | 151 | 37 |
| Osteoarthritis | 357 | 183 | 540 | 65 |
| Condromatosis | 85 | 106 | 191 | 49 |
| Os fabella | 79 | 73 | 152 | 44 |
| Normal | 290 | 233 | 523 | 40 |

## B. DATA PREPROCESSING

The knee X-rays were of different sizes and some of the radiographs had the patient's name, directional sign and various artifacts on them. Therefore, the radiographs were cropped from 1/4 distal femur and 1/4 proximal tibia to include the knee. The study was done with these X-rays. All of the images were resized to 224 x 224 pixels. Figure 1 shows knee X-ray image samples used in the study. Thousands of images are required to train CNN from scratch. If the number of data is insufficient, successful results can be obtained with the transfer learning method. In our previous studies, we achieved successful results with pre-trained VGG networks trained on the ImageNet dataset [30, 31], [32]. Therefore, in this study, we performed transfer learning with VGG-16. In the transfer learning process, the first 13 layers of the VGG-16 architecture, corresponding to the initial convolutional blocks (blocks 1–4), were frozen to retain the pre-trained ImageNet features. The final fully connected layer was retrained and replaced with a new dense layer to perform 5-class classification specific to the task. The dataset was divided into 85% training and 15% testing. 15% of the training dataset was used as the validation set. Test data was not used for training or validation at any stage. The validation dataset was used for hyperparameter tuning. Hyperparameter optimization was performed manually, inspired by similar studies in the literature and our previous studies, with the help of performance metrics obtained with the validation set during training, accuracy and loss graphs, and by trying. The training parameters of VGG-16 were set as follows; optimizer: adam, mini-batch size: 64, initial learning rate: 5e-5, validation frequency: 16, number of epochs: 25. Hyperparameters of machine learning algorithms are shown in Table 2. In this study, due to the limited and imbalanced nature of the dataset, data augmentation techniques were employed to improve model performance. The augmentation parameters included random rotation (±15 degrees), horizontal flipping, zoom range of 0.1, and width and height shift ranges of 0.1.

***Table 2.*** *Machine learning algorithms hyperparameters*

| Machine Learning Algorithm | Hyperparameters |
|---|---|
| Support Vector Machines | LinearSVC, C=1.0, degree = 3, gamma = 'scale' |
| Logistic Regression | penalty='l2', tol=0.0001, C= 0.5, solver= 'liblinear', max_iter=100, |
| Random Forest | criterion='gini', n_estimators=150, max_depth=None, max_leaf_nodes=None, |
| Decision Tree | criterion='gini', splitter='best', max_depth=None, min_samples_split=2, max_leaf_nodes=None, |

Osgood-Schlatter disease X-ray images

Knee osteoarthritis X-ray images

Knee chondromatosis X-ray images

Os fabella X-ray images

Normal knee X-ray images

***Figure 1.*** *Knee X-ray image samples*

## C. DATA PROCESSING ENVIRONMENT

This study was performed on a computer that has GeForce RTX2060 Graffic Processing Unit. Python 3.9 programming language was used, the required libraries were imported and worked on the tensorflow keras environment. The Sklearn library was used for statistical calculations.

## D. FEATURE EXTRACTION

Feature extraction from images is to obtain the image data as quantitative data in accordance with the structure of machine learning algorithms. Feature extraction is the process of extracting features such as edges, shapes, textures, colors, which are used to describe the content of the image. The major difference between CNNs and traditional machine learning methods is that CNN's directly extract image features without the need for manual feature extraction. CNN's is very successful in feature extraction [33], feature extraction is done through the convolution layer. In this study, feature extraction was performed from VGG-16 model intermediate layer, and 512 features were extracted. These features were used in machine learning algorithms.

## E. STATISTICAL ANALYSIS

Confusion Matrix is a table that summarizes the prediction results obtained with test data in a classification problem. Predicted values can be compared with actual values. Thus, accuracy, sensitivity, specificity, precision and F1 scores are calculated with the TP (True Positive), FP (False Positive), TN (True Negative), and FN (False Negative) numbers obtained. Using these parameters, the performance metrics were calculated using equations (1) through (5) below.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$sensitivity\ (recall) = \frac{TP}{TP + FN} \tag{2}$$

$$specificity = \frac{TN}{TN + FP} \tag{3}$$

$$precision = \frac{TP}{TP + FP} \tag{4}$$

$$F1\ score = \frac{2 * (precision * recall)}{(precision + recall)} = \frac{2TP}{2TP + FP + FN} \tag{5}$$

# III. RESULTS AND DISCUSSION

In this study, which was conducted to classify knee X-rays with deep learning methods, the transfer learning method was applied with pre-trained VGG-16 network. With pre-trained VGG-16, 95.3% accuracy, it ranged from 0.92 to 0.98 at the 95% confidence level (α :0.95, CI :0.92 - 0.98), 95.1% sensitivity (α :0.95, CI :0.92 - 0.98), 98.7% specificity (α :0.95, CI :0.97 - 1), 96.8% precision (α :0.95, CI :0.94 - 0.99), and 95.9% F1 score (α :0.95, CI :0.93 - 0.98) results were obtained. Figure 2 shows the confusion matrix and ROC curve obtained with the VGG-16 network. Feature extraction was performed from the intermediate layer of the pre-trained VGG-16 network. 512 features were obtained.

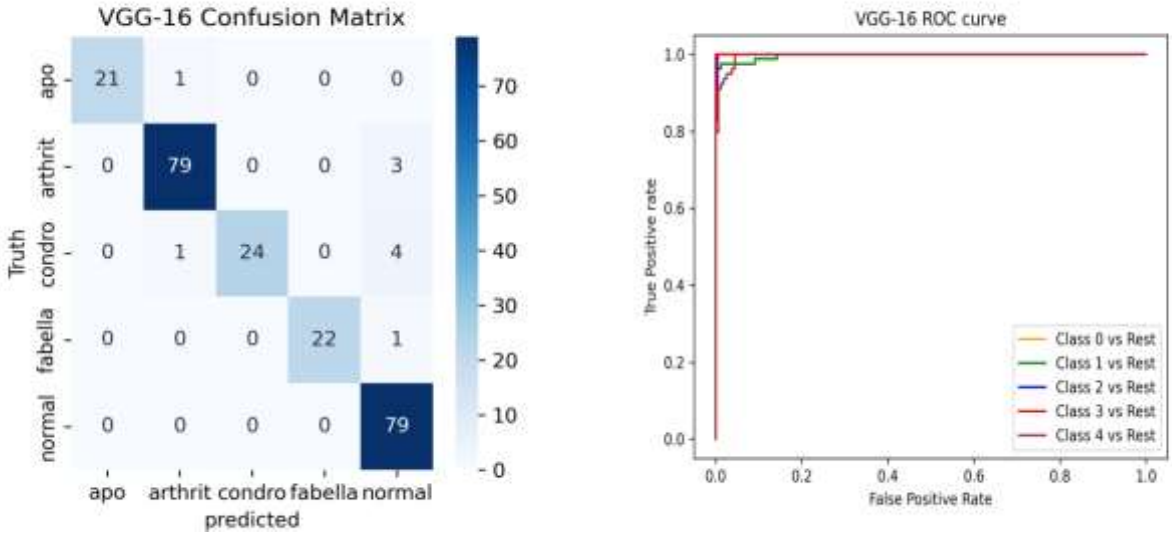


***Figure 2.*** *Confusion matrix and ROC curve obtained with test data using VGG-16 network. (apo: Osgood_Schlatter disease, arthrit : osteoarthritis, condro: condromatozis)*

These extracted features were reclassified with machine learning algorithms. For this, 5-fold cross validation was performed, the accuracy results obtained at each step with logistic regression, SVM, decision trees and random forest algorithms are shown in Table 3. It is seen that the accuracy results obtained at each step are consistent and the models perform well.

***Table 3****. Machine learning Algorithms Cross-Validation Accuracy Scores*

| | Machine learning Algorithms Cross-Validation Accuracy Scores | | | | | |
|---|---|---|---|---|---|---|
| | Step 1 | Step 2 | Step 3 | Step 5 | Step 5 | Mean |
| SVM | 0.9911 | 0.9955 | 0.9911 | 1. | 0.9910 | 0.9937 |
| Logistic regression | 0.9955 | 0.9955 | 0.9866 | 0.9955 | 0.9910 | 0.9928 |
| Random forest | 0.9955 | 0.9866 | 0.9777 | 0.9910 | 0.9955 | 0.9893 |
| Decision Tree | 0.9555 | 0.9466 | 0.9688 | 0.9642 | 0.9732 | 0.9617 |

With the test data that was not used in either training or validation, 98.2% accuracy (α : 0.95, CI: 0.96 - 0.99), 99.0% sensitivity (α : 0.95, CI: 0.97 - 1), 98.9% specificity (α : 0.95, CI: 0.97 - 1), 98.2% precision (α : 0.95, CI: 0.96 - 0.99) and 98.5% F1 score (α : 0.95, CI: 0.96 - 1) results were obtained with the logistic regression classifier. Table 4 shows the accuracy, sensitivity, specificity, precision and F1 scores obtained with the VGG-16 network and machine learning algorithms . Figure 3 shows the confusion matrix obtained with the logistic regression algorithm.

***Table 4.*** *Performance metrics obtained with VGG-16 network and machine learning algorithms*

| | Accuracy (%) | Sensitivity (%) | Specificity (%) | Precision (%) | F1 score (%) |
|---|---|---|---|---|---|
| VGG-16 model | 95.3 | 95.1 | 98.7 | 96.8 | 95.9 |
| Support Vector Machines | 97.8 | 98.7 | 99.1 | 97.9 | 98.3 |

**Table 4 (cont).** *Performance metrics obtained with VGG-16 network and machine learning algorithms*

| | | | | | |
|---|---|---|---|---|---|
| Logistic Regression | 98.2 | 99.0 | 98.9 | 98.2 | 98.5 |
| Random Forest | 97.8 | 98.3 | 98.8 | 98.3 | 98.3 |
| Decision Tree | 94.0 | 93.9 | 96.9 | 93.8 | 93.8 |

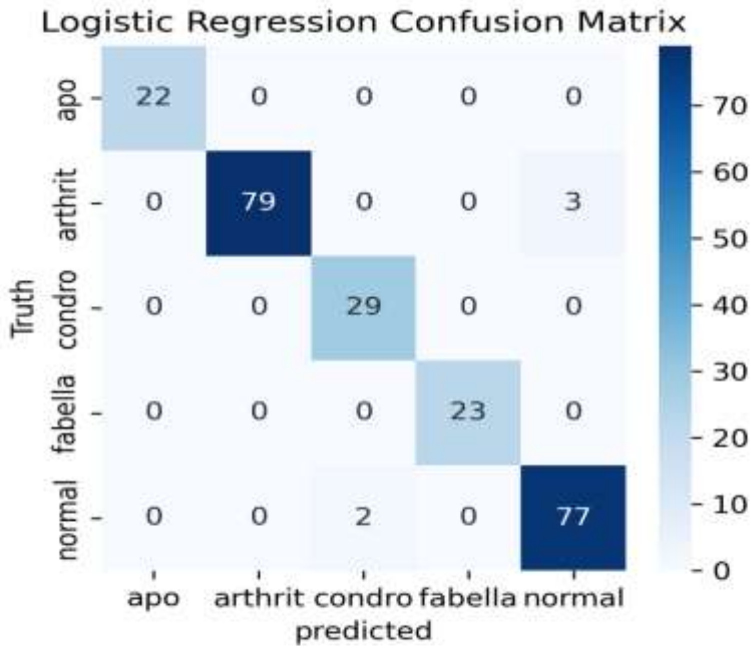


***Figure 3.*** *Confusion matrix obtained with Logistic Regression algorithm with test data (apo: Osgood_Schlatter disease, arthrit : osteoarthritis, condro: condromatozis)*

There are some studies conducted to classify knee osteoarthritis using CNN methods. To the best of our knowledge, this is the first study to classify radiographically identifiable Osgood-Schalatter, os fabella, knee chondromatosis, knee osteoarthritis radiographs and normal knee radiographs. First, transfer learning was performed with pre-trained VGG-16 network to classify knee X-rays. With the pre-trained VGG-16 model, 95.3% accuracy, 95.1% sensitivity, 98.7% specificity, 96.8% precision, and 95.9% F1 score results were obtained. Then, feature extraction was performed from the intermediate layer of the VGG-16 network. These extracted features were classified with random forest, support vector machines, logistic regression and decision tree machine learning algorithms. With the logistic regression algorithm, 98.2% accuracy, 99.0% sensitivity, 98.9% specificity, 98.2% precision and 98.5% F1 score results were obtained.

Convolutional Neural Networks (CNNs) are highly effective for both feature extraction and classification tasks. In a CNN, feature extraction is performed by the convolutional layers, while classification is handled by the fully connected layers.There are two common approaches to using CNNs: 1. End-to-end training, where the CNN learns to extract features and classify them in a single process. 2. Using pre-trained CNNs for feature extraction. In this method, a CNN that has already been trained on a large dataset is used to extract features from new data. These features are then passed to traditional machine learning algorithms (like SVM, Random Forest, etc.) for classification.

The second approach can significantly reduce training time, minimize computation, and require fewer computational resources, while still achieving strong performance. For instance, there are some studies conducted for the diagnosis of COVID-19 from chest X-rays. In these studies, feature extraction was first done with deep learning methods, and these extracted features were classified with support vector machines or XGBoost machine learning algorithm [34, 35]. In this study, feature extraction was performed with the VGG-16 network from knee X-rays. These extracted features were classified with machine learning algorithms. Successful results were obtained with logistic regression, support vector machines and random forest classifiers.

In this study, where feature extraction was performed using deep learning algorithms and the dataset exhibited class imbalance, the logistic regression algorithm outperformed SVM, random forest, and decision tree classifiers. This may be attributed to the nature of deep learning-extracted features, which often produce linearly separable representations that are well-suited for linear models. Previous studies have shown that logistic regression can perform competitively when such features are informative and of reduced dimensionality. Additionally, the relatively small sample sizes in some classes may have adversely affected tree-based methods and SVMs, which are more sensitive to class imbalance and sparse distributions without appropriate tuning or resampling [36-38].

All models in this study demonstrated strong performance despite the presence of class imbalance. Logistic regression achieved the highest sensitivity (99.0%) and F1-score (98.5%), highlighting its effectiveness when combined with deep learning-based feature extraction. VGG-16 also performed well with 95.3% accuracy, while SVM and Random Forest yielded similarly high results across multiple evaluation metrics. Considering the confusion matrices and key performance metrics such as sensitivity, specificity, precision, and especially F1-score, the overall performance of the developed models was found to be satisfactory. These findings suggest that both traditional machine learning algorithms and deep learning-based methods can achieve effective classification performance on imbalanced medical imaging datasets when supported by informative feature representations.

One of the limitations of this study is that all radiographic images were obtained from a single center (Ankara Bilkent City Hospital), which may introduce potential dataset bias and limit the generalizability of the results. Although the dataset size is methodologically reasonable, relying on a single source may lead the model to learn center-specific features related to imaging protocols, equipment, or population characteristics. To improve external validity, future studies should incorporate data from multiple centers and machines, allowing for better assessment of the model's robustness across diverse imaging conditions and patient populations. Additionally, external validation on independent datasets would further strengthen confidence in the model’s generalizability.

Patients who apply to general practitioners with complaints such as pain, limitation of movement and swelling in the knees can be diagnosed correctly with physical examination and simple radiological imaging methods. The method we propose may be particularly helpful to physicians who do not have sufficient experience to evaluate plain knee radiographs. Thus, there is no need for advanced imaging methods. The patient is treated with methods such as weight loss, exercise, and lifestyle changes, and the patient is referred to the appropriate specialist when necessary.

# IV. CONCLUSION

Successful results were obtained in this first deep learning and machine learning study for the classification of Osgood-Schalatter, os fabella, knee chondromatosis, knee osteoarthritis and normal knee X-rays, which can be diagnosed radiographically. The performance of deep learning models can be improved with machine learning algorithms. If this model is developed through multicenter studies and more X-rays, it may assist clinicians in their daily work, especially physicians who do not have sufficient experience to evaluate plain knee X-rays. Thus, an accurate diagnosis can be made in most patients without the need for advanced imaging methods such as CT or MRI.

## Article Information

**Ethics Approval:** An ethical approval certificate was obtained from the local ethic commitee, Ankara Bilkent City Hospital, Clinical Research Ethics Committee No. 1, date : 20.04.2022, no : E1-22-2420.

**Informed Consent:** The study is a retrospective study, archival radiographs were studied, no data indicating patient identity was used, ethics committee approval was obtained, therefore informed consent form was not obtained.

**Funding:** None.

**Author Contributions:** Data collection S.D., Y.M., E.A. and K.O., data analysis K.Ü., S.D. and K.O., writing—original draft, K.Ü., H.H.M. and Y.M., writing—review and editing, K.Ü., E.A., and H.H.M.. All authors have read and approved the final version of it.

**Artificial Intelligence Statement:** No any Artificial Intelligence tool is used in this paper.

**Conflict of Interest Disclosure**: The authors declare that they have no conflicts of interest.

**Plagiarism Statement:** This article was scanned by the plagiarism program.

# V. REFERENCES

[1] M. J. Lespasio, N. S. Piuzzi, M. E. Husni, G. F. Muschler, A. Guarino and M. A. Mont, "Knee osteoarthritis: a primer," *The Permanente Journal*, vol. 21, no. 4, pp. 16–183, 2017.

[2] B. Heidari, "Knee osteoarthritis prevalence, risk factors, pathogenesis and features: Part I.," *Caspian Journal of Internal Medicine*, vol. 2, no. 2, pp. 205–12, 2011.

[3] R. C. Lawrence et al., "Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part II.," *Arthritis & Rheumatism*, vol. 58, no. 1, pp. 26–35, 2008.

[4] J. A. Neumann, G. E. Garrigues, B. E. Brigman and W. C. Eward, "Synovial chondromatosis," *JBJS Reviews*, vol. 4, no. 5, 2016, Art. no. e2.

[5] M. A. Adelani, R. M. Wupperman and G. E. Holt, "Benign synovial disorders," *Journal of the American Academy of Orthopaedic Surgeons*, vol. 16, no. 5, pp. 268–275, 2008.

[6] A. Robertson, S. C. E. Jones, R. Paes and G. Chakrabarty, "The fabella: a forgotten source of knee pain?," Knee, vol. 11, no. 3, pp. 243–245, 2004.

[7] O. Unluturk, S. Duran and H. Yasar Teke, "Prevalence of the fabella and its general characteristics in Turkish population with magnetic resonance imaging," *Surgical and Radiologic Anatomy*, vol. 43, no. 12, pp. 2047–2054, 2021.

[8] M. A. Berthaume, E. Di Federico and A. M. J. Bull, "Fabella prevalence rate increases over 150 years, and rates of other sesamoid bones remain constant: a systematic review," *Journal of Anatomy*, vol. 235, no. 1, pp. 67–79, 2019.

[9] N. H. Sahar, A. S. Ramli, S. F. Badlishah-Sham and M. F. Mohd Miswan, "Osgood-Schlatter disease in adult: would early diagnosis and treatment ımprove the prognosis?," *Journal of Clinical and Health Sciences*, vol. 8, no. 1, pp. 76–81, 2023.

[10] J. Høgh and B. Lund, "The sequelae of Osgood-Schlatter's disease in adults," *International Orthopaedics*, vol. 12, no. 3, pp. 213–215, 1988.

[11] G. L. De Lucena, C. Dos Santos Gomes and R. Oliveira Guerra, “Prevalence and associated factors of osgood-schlatter syndrome in a population-based sample of brazilian adolescents,” *The American Journal of Sports Medicine*, vol. 39, no. 2, pp. 415–420, 2011.

[12] M. Mazzei, "A change detection with machine learning approach for medical image analysis". in *MICAD 2022. Lecture Notes in Electrical Engineering*, vol. 810. R. Su, Y. Zhang, H. Liu and A. F. Frangi, Eds. Singapore: Springer, 2023, pp. 203-229.

[13] S. Suganyadevi, V. Seethalakshmi and K. Balasamy, “A review on deep learning in medical image analysis,” *International Journal of Multimedia Information Retrieval*, vol. 11, no. 1, pp. 19–38, 2022.

[14] P. K. Sethy, S. K. Behera, P. K. Ratha and P. Biswas, “Detection of coronavirus disease (COVID-19) based on deep features and support vector machine,” *International Journal of Mathematical, Engineering and Management Sciences,* vol. 5, no. 4, pp. 643–651, 2020.

[15] S. S. Abdullah and M. P. Rajasekaran, “Automatic detection and classification of knee osteoarthritis using deep learning approach,” *La Radiologia Medica*, vol. 127, no. 4, pp. 398–406, 2022.

[16] K. O’Shea and R. Nash, “An introduction to convolutional neural networks,” *International Journal of Research in Applied Sciences and Engineering Technology*, vol. 10, no. 12, pp. 943–947, 2022.

[17] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg and L. Fei-Fei, “ImageNet large scale visual recognition challenge,” *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, 2015.

[18] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, “A comprehensive survey on transfer learning,” *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43–76, 2021.

[19] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” *Advances in Neural Information Processing Systems*, vol. 25, 2012.

[20] C. Szegedy *et al*., "Going deeper with convolutions," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, 2015, pp. 1-9.

[21] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” in *3rd International Conference on Learning Representations (ICLR)*, San Diego, CA, USA, 2015, pp. 1–14, 2014.

[22] K. He, X. Zhang, S. Ren and J. Sun, “Deep residual learning for image recognition,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 770–778.

[23] A. G. Howard et al., “MobileNets: efficient convolutional neural networks for mobile vision applications,” *arXiv: Computer Vision and Pattern Recognition*, 2017, [Online]. Available: http://arxiv.org/abs/1704.04861.

[24] A. W. Salehi et al., “A study of CNN and transfer learning in medical imaging: advantages, challenges, future scope,” *Sustainability*, vol. 15, no. 7, 2023, Art. no. 5930.

[25] P. M. Shakeel, M. A. Burhanuddin and M. I. Desa, “Automatic lung cancer detection from CT image using improved deep neural network and ensemble classifier,” *Neural Computing and Applications*, vol. 34, no. 12, pp. 9579–9592, 2022.

[26] P. Tiwari et al., “CNN based multiclass brain tumor detection using medical imaging,” *Computational Intelligence and Neuroscience*, vol. 2022, pp. 1–8, 2022.

[27] A. Tiulpin, J. Thevenot, E. Rahtu, P. Lehenkari and S. Saarakkala, “Automatic knee osteoarthritis diagnosis from plain radiographs: a deep learning-based approach,” *Scientific Reports*, vol. 8, no. 1, 2018, Art. no. 1727.

[28] Y. Wang, X. Wang, T. Gao, L. Du and W. Liu, “An automatic knee osteoarthritis diagnosis method based on deep learning: data from the osteoarthritis initiative,” *Journal of Healthcare Engineering*, vol. 2021, pp. 1–10, 2021.

[29] B. Liu, J. Luo and H. Huang, “Toward automatic quantification of knee osteoarthritis severity using improved Faster R-CNN,” *International Journal of Computer Assisted Radiology and Surgery*, vol. 15, no. 3, pp. 457–466, 2020.

[30] K. Üreten and H. H. Maraş, “Automated classification of rheumatoid arthritis, osteoarthritis, and normal hand radiographs with deep learning methods,” *Journal of Digital Imaging*, vol. 35, no. 2, pp. 193–199, 2022.

[31] S. Duran et al., “Automatic detection of spina bifida occulta with deep learning methods from plain pelvic radiographs,” *Research on Biomedical Engineering*, vol. 39, no. 3, pp. 655–661, 2023.

[32] K. Üreten, H. F. Sevinç, U. İğdeli, A. Onay and Y. Maraş, “Use of deep learning methods for hand fracture detection from plain hand radiographs.,” *Turkish Journal of Trauma & Emergency Surgery*, vol. 28, no. 2, pp. 196–201, 2022.

[33] S. Benyahia, B. Meftah and O. Lézoray, “Multi-features extraction based on deep learning for skin lesion classification,” *Tissue Cell*, vol. 74, 2022, Art. no. 101701.

[34] H. Nasiri and S. Hasani, “Automated detection of COVID-19 cases from chest X-ray images using deep neural network and XGBoost,” *Radiography*, vol. 28, no. 3, pp. 732–738, 2022.

[35] H. Nasiri and S. A. Alavi, “A novel framework based on deep learning and ANOVA feature selection method for diagnosis of COVID-19 cases from chest x-ray images,” *Computational Intelligence and Neuroscience*, vol. 2022, pp. 1–11, 2022.

[36] A. Mahbod, G. Schaefer, C. Wang, R. Ecker and I. Ellinge, "Skin lesion classification using hybrid deep neural networks," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brighton, UK, 2019, pp. 1229-1233.

[37] Y. Wang and X. S. Ni, “Predicting class-imbalanced business risk using resampling, regularization, and model emsembling algorithms,” *International Journal of Management and Information Technology*, vol. 11, no. 1, 2019.

[38] Y. Jiao, J. Yuan, Y. Qiang and S. Fei, “Deep embeddings and logistic regression for rapid active learning in histopathological images,” *Computer Methods and Programs in Biomedicine*, vol. 212, 2021, Art. no. 106464.