Dicle University
Journal of Engineering
https://dergipark.org.tr/tr/pub/dumf
duje.dicle.edu.tr

Research Article

Detection of Flatfoot Deformity from X-Ray Images using Image Filtering and Transfer Learning Approaches

Merve KOKULU[1], Hanife GÖKER[2*], Ömer KASIM[3]

[1] Dumlupınar University, Electrical-Electronics Engineering Department, merve.kokulu0@ogr.dpu.edu.tr, Orcid No: 0009-0007-3593-9666
[2] Gazi University, Health Services Vocational College, gokerhanife@gazi.edu.tr, Orcid No: 0000-0003-0396-7885
[3] Dumlupınar University, Electrical-Electronics Engineering Department, omer.kasim@dpu.edu.tr, Orcid No: 0000-0003-4021-5412

ARTICLE INFO

Article history:

Received 1 January 2025
Received in revised form 27 February 2025
Accepted 15 March 2025
Available online 26 March 2025

Keywords:

Deep learning, Transfer learning, Image processing, Flatfoot, Image filter.

Doi: 10.24012/dumf.1611410

* Corresponding author

ABSTRACT

Flatfoot (pes planus) is a condition characterized by the flattening of the foot's arch due to the collapse of the foot structure or the weakening of the ligaments and muscles that maintain the alignment and curvature of the bones and tissues in the foot. If left untreated, flatfoot can lead to calf, knee, hip, and lower back pain, as well as postural disorders resulting from foot deterioration. This study presents a transfer learning-based approach for flatfoot detection from X-ray images using the Dilation filter. The dataset consists of 402 flatfoot images and 440 control images. Dilation filtering is employed for image preprocessing, enhancing the images through this method. After preprocessing, the performance of several transfer learning models, including DarkNet19, GoogLeNet, DenseNet-201, ResNet-101, and MobileNetV2, is compared. The holdout method was used for performance evaluation. Experimental results demonstrate that the DenseNet-201 transfer learning model with Dilation filtering outperforms the others, achieving an overall accuracy of 0.9802 and a Cohen's Kappa value of 0.96. These results indicate that the combination of dilation filtering and transfer learning offers an effective solution for automatic flatfoot detection. Compared to similar studies in the literature, the DenseNet-201 transfer learning model with Dilation filtering exhibits significantly higher accuracy.

Introduction

Flatfoot is a foot deformity that occurs when the arch of the foot deteriorates and becomes flat as a result of various disorders in the bones and tissues in the foot. It can be seen in children as well as adults. It can also be congenital or may occur due to causes such as insufficient tendons, injuries, arthritis or connection disorders between the back bones in the foot (tarsal coalition) [1]. Flatfoot can develop due to a number of factors, including obesity, shoe wear, the child's sitting and sleeping habits, structural defects in the lower limbs, muscle and ligament weaknesses and tendon ruptures [2]. In addition to these causes, pregnancy, diabetes and advanced age are also possible causes. Symptoms include pain in the ankle, cramps in the feet and legs, muscle aches and foot fatigue, outwardly shifted appearance of the toes, and gait disturbances. Pain is one of the most important symptoms of musculoskeletal disorders that reduce quality of life [3]. Although the exact prevalence in children is not known, it is reported to be 20-25% in adults [4]. In a screening of 377 preschool children, flatfoot was found in 57% of 2–3-year-olds, 28% of 4–5-year-olds and 21% of 5–6-year-olds [5]. In one study, it was found to be 18.1% in men and 14.6% in women [6]. Flatfoot, if left untreated, can lead to increased pain and joint deterioration. When the correct and necessary treatments are not applied to flatfoot patients, their mobility will decrease and their quality of life will be negatively affected.

Flatfoot can be detected by achilles shortness, Jack test, pedobarographic examination, and radiological evaluations (ultrasonographic examination, computed tomography, magnetic resonance). Among these evaluations, achilles shortness and Jack test evaluations may be difficult to detect mild shortness and may give misleading results. In addition, its reliability in children may be variable. Pedobarographic examinations and radiologic evaluations can be costly and time-consuming. Computed tomography can be a difficult process for pediatric patients with its high radiation properties. Since the feet and the arch of the foot are still in the developmental stage, flat feet in children can in some cases resolve without treatment. Flat feet in adults are usually permanent as they develop later. Treatment is not required in painless, asymptomatic cases that do not cause problems. Shoe modifications and soft silicone insoles can be used [4]. However, soft silicone insoles are short-lived due to their structure. They can also cause hygiene problems in long-term use. There are both conservative and surgical options for treatment; surgery is the last treatment

option. Barefoot walking exercises and other recommended activities are widely practiced in the treatment of flat feet. While some researchers believe that strengthening exercises may not be beneficial when muscle weaknesses are unclear, Riccio et al. have shown that appropriate rehabilitation exercises can improve the effectiveness of pediatric flexible flatfoot therapy [2]. Painkillers and anti-inflammatory drugs can be used to alleviate pain, but these drugs only relieve symptoms and do not eliminate the cause of flatfoot. Obesity is an important factor that increases the risk of flatfoot. Therefore, weight control and achieving a healthy body weight can alleviate the symptoms of flatfoot. If flatfoot is severe and does not respond to other treatment methods, surgical intervention may be required. Surgical methods aim to reconstruct the arch of the foot or correct the foot structure.

Compared to the diagnosis of other orthopedic diseases, studies using transfer learning approaches for flatfoot detection are more limited. Wang et al. (2019) presented an image processing-based method for plantar pressure optical sensor image classification using CNN rotation invariance LBP (RI-LBP) algorithm, which is a local binary pattern method and achieved 82% success [7]. Kim et al. (2021) achieved 84.09% success using VGG16 architecture for flatfoot detection [8]. Jung et al. (2022) presented a decision tree-based method that can prescribe foot orthotics for flatfoot patients and achieved 80.16% success [9]. Lauder et al. (2022) used a Random Forest algorithm to evaluate foot collapse in lateral weight-bearing foot radiographs using a point finder, achieving a median point-to-point error of 2.2 mm [10]. Ryu et al. (2022) used a cascaded convolutional neural network (FlatNet) for deformity diagnosis and improved the mean absolute distance X and Y coordinates for the expert orthopedic surgeon to 0.87 ± 1.21, 0.69 ± 0.74 and 1.24 ± 1.31, respectively, and the mean absolute distance X and Y coordinates for the general practitioner to 0.74 ± 0.73, 0.57 ± 0.63 and 1.04 ± 0.85, respectively [11]. Ryu et al. (2022) achieved a Dice similarity coefficient of 0.964, a Hausdorff distance of 1.292 mm, a minimum moment of inertia of 0.828°, and an ellipsoidal alignment of 1.186° using active learning with U-Net-based semantic segmentation for flatfoot assessment in the long axis of the tarsal and metatarsal bones [12]. Gül et al. (2023) achieved 95.14% success using MobileNetV2 architecture for flatfoot detection [13]. Alsaidi et al. (2023) achieved 93.13% success with CNN and Random Forest algorithm [14]. Doğan et al. (2024) achieved 96.80% success in flatfoot detection using VGG16 architecture [15]. As yet, the development of transfer learning approaches and image processing techniques for the automatic detection of flatfoot has not been addressed in sufficient depth. In transfer learning architectures, high success rates can be achieved with image filtering methods. The need for an effective detection system for fast and accurate detection of such conditions is increasing.

In this study, an image filtering technique is applied to achieve high performance. The preferred filtering method for X-ray images is the dilation filter. Furthermore, the performances of transfer learning approaches DarkNet19, DenseNet-201, GoogLeNet, MobileNetV2 and ResNet-101 are analyzed. The main contributions of the study are summarized below:

a) An image processing-based solution using transfer learning and image filtering is proposed for flatfoot detection from X-ray images.

b) To determine the best classification performance, the performances obtained with transfer learning approaches are compared.

c) The DenseNet-201 transfer learning model with Dilation filtering can support experts in detecting flatfoot deformity in the early stages and during busy work schedules.

d) The images were filtered and high accuracy rates were obtained with transfer learning methods.

Materials and Methods

Proposed Method

In this study, a solution for flatfoot detection from X-ray images is developed using transfer learning approaches and image preprocessing methods. X-ray images are reconstructed in appropriate dimensions for each transfer learning technique. These images are preprocessed by applying a dilation filter. The success of the model is calculated with various evaluation metrics and compared between different approaches. The process flow of the DenseNet-201 model with Dilation filtering is shown in Figure 1.

Dataset

In this study, a publicly available dataset was used [13]. The flatfoot radiographic images used in this study were obtained from radiographic images of patients who came to the Radiology Department of Elazığ Fethi Sekin City Hospital for routine military health screening or with suspicion of flatfoot. The age range of the patients whose X-Ray images were obtained was 14-47 years. Images were collected after approval from the ethics committee of Fırat University, Turkey. There are X-Ray images of 439 patients in total. The X-ray images in the dataset were obtained with a Philips dual detector digital X-ray device (65 kV, 6.3 mAs). X-Ray images of 18 patients with poor quality and resolution were not used and are not included in the dataset. The remaining 842 X-ray images were labeled by two expert radiologists by measuring the calcaneal tilt angle and completed by a third expert radiologist by examining complex specimens. Of the 842 X-Ray images, 402 were labeled as flatfooted and 440 as normal [13]. The training set of 589 images was divided into a training set for the training phase and a test set of 253 images for the testing phase. Figure 2 shows the images in the dataset.

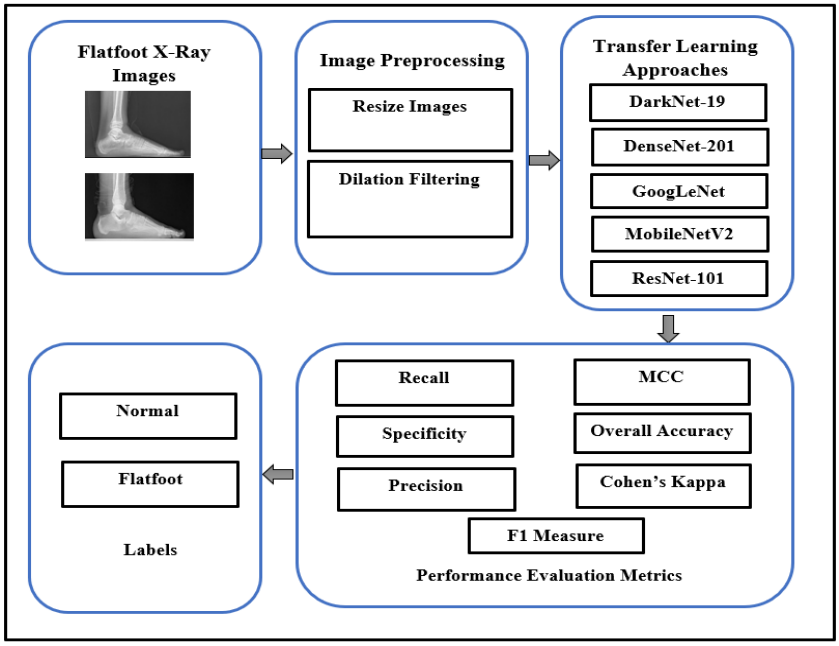


Figure 1. Flowchart illustrating the proposed method

(i) (ii) (iii) (iv)

Figure 2. Samples of X-Ray images per label, (i) normal (left), (ii) normal (right), (iii) flatfoot (left), (iv) flatfoot (right)

Holdout Method

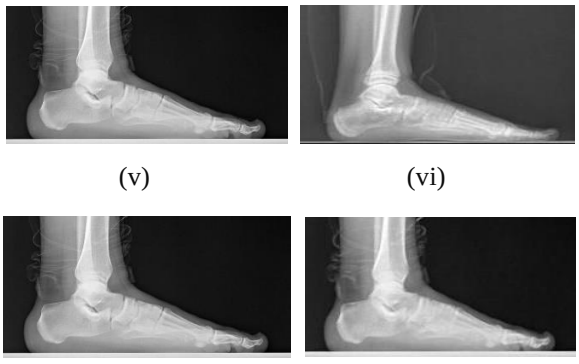
In this study, the holdout method is used, which allows the dataset to be divided into two main parts as training and test. In the training set, the learning process of the model was performed and the model was created. The test set was used to assess the performance of the model. 70% of the dataset was allocated for training and 30% for test (Table 1).

Table 1. Division of data

	Number of Dataset	Training Data	Test Data
Normal	440	308	132
Flatfoot	402	281	121
Total	842	589	253

Image Preprocessing

All the images in the dataset were resized and then dilation filtering was applied. Dilation is a basic morphological operation. It is used to extend the boundaries of the region under consideration and to prevent the same object from appearing as two separate objects by thinly dividing it with a noise. It performs a pixel layer addition to both the inner and outer boundaries of the regions in the image. The output pixel value is set as the highest value of its neighboring pixels. In a binary image, a pixel is set to 1 if any of its neighbors has a value of 1. Morphological magnification makes objects more distinct and fills in small gaps. Figure 3 shows the images without and with image preprocessing.



(v) (vi)

(vii) (viii)

Figure 3. Examples of image X-Ray images per label, (i) normal, (ii) normal (with image preprocessing), (iii) flatfoot, (iv) flatfoot (with image preprocessing)

Transfer Learning Approaches

DarkNet-19

The DarkNet19 architecture has been proposed as a classification model based on the YOLOv2 model [16]. Similar to the VGG model, 3x3 convolution filters are used and the number of filters is doubled after each maximum pooling layer. In the classification layer, a convolution filter of size 1x1 is applied between the average pooling and the 3x3 convolution layer to make features more salient. DarkNet19 has a structure with 19 convolution layers and five maximum pooling layers [17]. Since the input size of the DarkNet-19 architecture is 256x256, the flatfoot images were resized. Figure 4 shows the DarkNet-19 architecture.

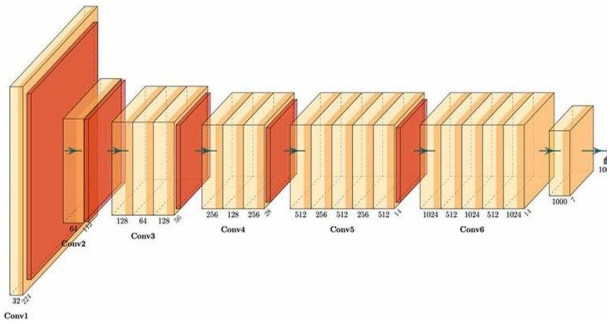


Figure 4. DarkNet-19 architecture [18]

DenseNet-201

DenseNet integrates all layers to create an efficient and deep model. It is very similar to ResNet, but with some important differences [19]. DenseNet is similar to ResNet in the sense that each layer feeds the next layer. Thanks to these feed-forward connections, the number of layers is increased from L to L (L + 1)/2 [20]. The input of each layer contains the feature maps for all previous levels. DenseNet has many advantages, such as the ability to resolve vanishing gradients [21]. Since the input size of the DenseNet-201 architecture is 224x224, the flatfoot images were resized. Figure 5 shows the DenseNet-201 architecture.

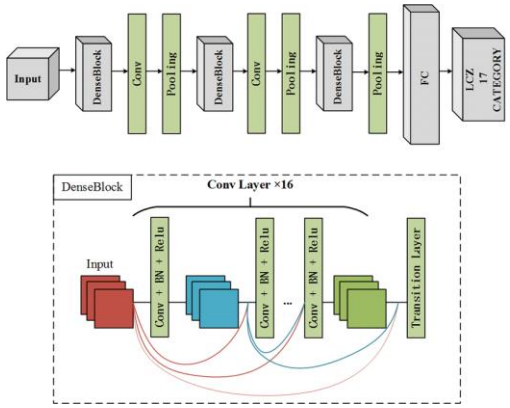


Figure 5. DenseNet-201 architecture [22]

GoogLeNet

2014 winner of the ILSVRC competition. GoogLeNet architecture uses methods such as 1x1 convolution and global average pooling. These methods enable the creation of a deep architecture. Therefore, the GoogLeNet architecture is different from other architectures. GoogLeNet is a complex structure due to its initial modules. The initialization module is a small initialization module that helps to limit the scale of parameters and model complexity. of a few evolved nuclei. [23]. Since the input size of the GoogLeNet architecture is 224x224, the flatfoot images were resized. Figure 6 shows the GoogLeNet architecture.

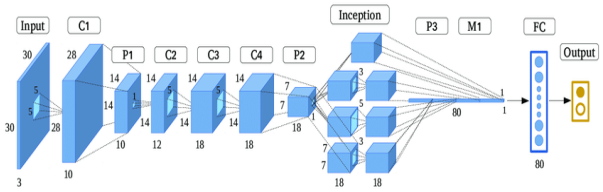


Figure 6. GoogLeNet architecture [24]

MobileNetV2

MobileNetV2 architecture is an architecture with flexibility and scalability capabilities. It can also achieve lightweight, high-performance capabilities without sacrificing accuracy and generalizability. The MobileNetV2 architecture consists of inverted residual structures, fully connected layers, stacked convolutional layers, pooling layers and other basic network structures [25]. Since the input size of the MobileNetV2 architecture is 224x224, the flatfoot images were resized. Figure 7 shows the MobileNetV2 architecture.

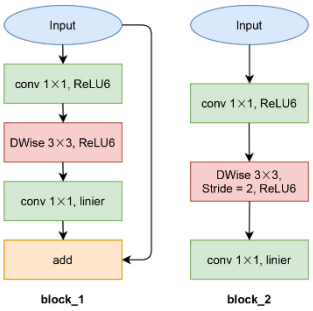


Figure 7. MobileNetV2 architecture [26]

ResNet-101

ResNet-101 consists of 101 layers is made up of. This network contains more than 1 million images Pre-trained on the ImageNet database. Education process, computers, printers, and various more than 1000 different classes such as animal species and more to have the ability to recognize. [27]. In the 2015 ILSVRC competition, it achieved a high degree of accuracy and won a great success [28]. ResNet is one of the pioneering algorithms in incorporating batch normalization. [29]. Since the input size of the ResNet-101 architecture is 224x224, the flatfoot images were resized. Figure 8 shows the ResNet-101 architecture.

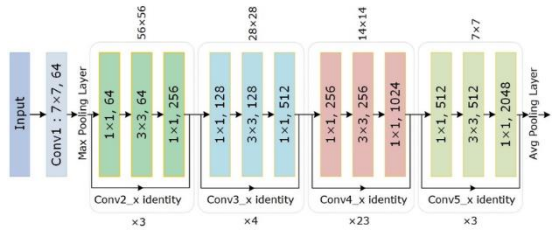


Figure 8. ResNet-101 architecture [30]

Performance Evaluation Metrics

Various evaluation metrics are used to evaluate the accuracy of models created by classification algorithms and to determine which model performs better. These metrics are usually derived from a table called the confusion matrix [31]. The confusion matrix contains four basic values: false negative (FN), false positive (FP), true positive (TP), and true negative (TN). It also shows the number of correct or incorrect predictions of the classification results [32]. Metrics commonly used to measure model performance include specificity, precision, sensitivity, overall accuracy, Matthews Correlation Coefficient (MCC), F1 measure, and Cohen's kappa statistic [33, 34].

$$Overall\ Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$F1 - score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (4)$$

$$MCC = \frac{(TN \times TP) - (FP \times FN)}{\sqrt{(TN + FN) \times (FP + TP) \times (TN + FP) \times (FN + TP)}} \quad (5)$$

$$Specificity = \frac{TN}{TN + FP} \quad (6)$$

The kappa statistic is a statistical calculation commonly used to determine the agreement between values. It was developed to measure the level of agreement between two different values when analyzing classification results. One of the most important advantages of Kappa is that it is very simple to calculate and its results can be easily interpreted. In addition, the most important feature is that the agreement can be evaluated by removing the effect of chance [35]. The mathematical formulas of Cohen's Kappa statistic are given below [36]:

$$P_0 = \frac{TP + TN}{TP + FN + FP + TN}$$

$$Pc_1 = \frac{(TP + FN) \times (TP + FP)}{[TP + FN + FP + TN]2}$$

$$Pc_0 = \frac{(FP + TN) \times (FN + TN)}{[TP + FN + FP + TN]2}$$

$$Pe = Pc_1 + Pc_0$$

$$Cohen's\ Kappa = \frac{P_0 - Pe}{1 - Pe} \quad (7)$$

The Kappa statistic of over 0.4 shows an acceptable fit, a value from 0.6 to 0.8 shows a significant fit, and a value between 0.8 and 1 shows a perfect fit.

Experimental Results and Discussion

In this study, the dataset contains radiographic images of individuals with flatfoot deformity and control group. X-Ray images were obtained from patients who came to the Radiology Department of Elazığ Fethi Sekin City Hospital for routine military health screening or with suspected flatfoot [7]. Transfer learning models were used to help detect this deformity. A total of 842 radiographic images were used in the study, 440 control images and 402 flatfoot images. The flatfoot images were resized to align with the input sizes of the transfer learning models. Dilatation filter, an image preprocessing method, was performed to resize images. The hyperparameters used for flatfoot detection are: maximum epochs 60, optimizer stochastic gradient descent with momentum (sgdm), initial learning rate 0.001, mini batch size 16, each epoch shuffle, bias learning rate factor 20, and weight learning rate factor 20. The confusion charts obtained from the experiments are shown in Figure 9.

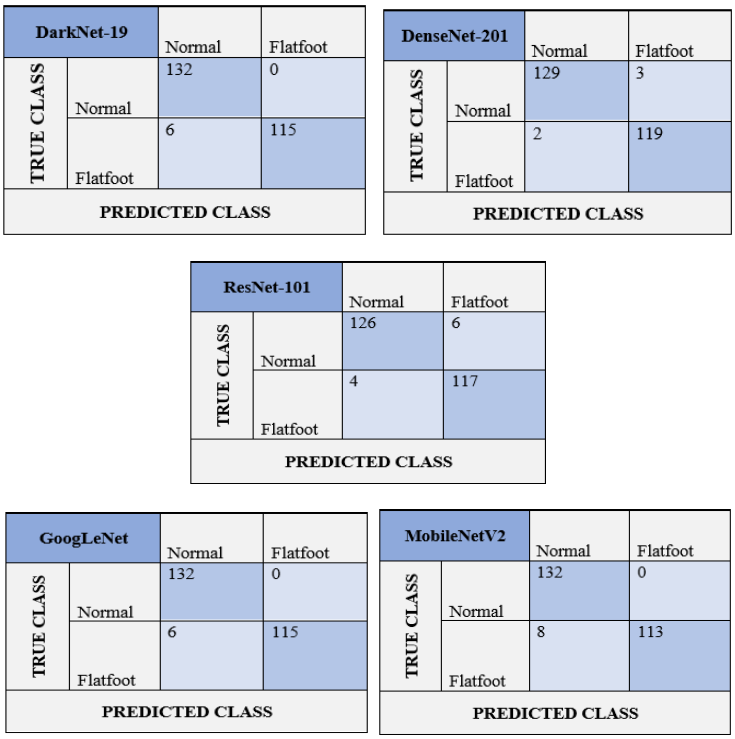


Figure 9. Confusion charts

Table 2. Confusion matrix-based performance metrics values and Cohen's kappa statistic values of transfer learning architectures

Models	Recall	Specificity	Precision	F1 measure	MCC	Overall Accuracy	Cohen's Kappa
DarkNet-19	0.9565	1.0	1.0	0.9778	0.9535	0.9763	0.952
DenseNet-201	0.9847	0.9754	0.9773	0.9810	0.9604	0.9802	0.96
GoogLeNet	0.9565	1.0	1.0	0.9778	0.9535	0.9763	0.952
ResNet-101	0.9692	0.9512	0.9545	0.9618	0.9210	0.9605	0.921
MobileNetV2	0.9429	1.0	1.0	0.9706	0.9384	0.9684	0.936

The comparison of transfer learning approaches was carried out in MATLAB environment. All transfer learning approaches were evaluated using the same hyperparameter values. After the training and testing processes were completed, confusion matrices were obtained for each transfer learning approach. These matrices enable the calculation of different metrics and a detailed analysis of model accuracy by process. Table 2 show the performance metric values obtained from the transfer learning approaches.

The calculated performance evaluation metrics of the confusion matrices obtained were compared with the performances of the transfer learning models. When the values of the calculated performance evaluation metrics are close to 1, it shows that the fit of the model is high. It can be seen in Table 2 that the transfer learning model with the best performance is the DenseNet-201. 0.9847 sensitivity, 0.9754 specificity, 0.9773 precision, 0.9604 MCC, 0.9802 overall accuracy and 0.96 Cohen's kappa statistic values are the results of the model.

The performance rate of the DenseNet-201 model with Dilation filtering and the performance rates of previous studies in the literature are shown in Table 3. Considering the results obtained in relevant studies with the datasets used in the study, Wang et al. obtained 82% accuracy using local binary pattern and CNN [7]. Kim et al. obtained 84.09% accuracy using data augmentation and VGG16 architecture [8]. Jung et al. achieved 80.16% accuracy using decision tree method [9]. Gül et al. obtained 95.14% accuracy using data augmentation, image resizing and MobileNetV2 architecture [13]. Alsaidi et al. achieved 93.13% accuracy by adding filler, resizing images and using random forest algorithm [14]. Doğan et al. obtained 96.80% accuracy using data augmentation, image resizing and VGG-16 architecture [15]. In the DenseNet-201 model with Dilation filtering, a 98.02% accuracy rate was achieved on the DenseNet-201 architecture by adopting a different approach than the transfer learning approaches and image preprocessing methods used in previous studies.

Table 3. A comparative analysis of relevant studies

Research	Image Preprocessing	Dataset	Number of Class	Best Model	Performance
Wang et. al. (2019) [7]	Local Binary Pattern	21 × 20 matrix consisting of 20 testee (normal/flatfoot/talipes equinovarus)	3	CNN	Acc. 0.82
Kim et. al. (2021) [8]	Data Augmentation	176 images: 88 normal, 88 flatfoot	2	VGG-16	Acc. 0.8409
Jung et. al. (2022) [9]	-	100 images: 40 normal, 60 flatfoot	2	Decision Tree	Acc. 0.8016
Gül et. al. (2023) [13]	Data Augmentation Resizing Images	842 images: 440 normal, 402 flatfoot	2	MobileNetV2	Acc. 0.9514
Alsaidi et. al. (2023) [14]	Adding Filler Resizing Images	Training dataset: 8000 images (normal/ slightly flatfoot/ medium flatfoot) Test dataset: 240 images (normal/ slightly flatfoot/ medium flatfoot)	3	Random Forest	Acc. 0.9313
Doğan et. al. (2024) [15]	Data Augmentation Resizing Images	842 images: 440 normal, 402 flatfoot	2	VGG-16	Acc. 0.9680
Proposed method	Resizing Images Dilation Filtering	842 images: 440 normal, 402 flatfoot	2	DenseNet-201	Acc. 0.9802

Conclusion

In this study, a method for detecting flatfoot deformity using flatfoot X-ray images containing 2 different classes, control and flatfoot, is proposed. In this method, flatfoot X-ray images are resized for each transfer learning architecture. Dilatation filtering, an image preprocessing method, is applied to the resized images. The performances of images with dilation filter applied as image preprocessing method are compared between transfer learning approaches. DarkNet-19, GoogLeNet, DenseNet-201, ResNet-101, and MobileNetV2 were used as transfer learning architectures. Recall, specificity, overall accuracy, F1 measure, MCC, and Cohen's Kappa statistic were used. The highest performing architecture among the transfer learning approaches was DenseNet-201 with an overall accuracy of 0.9802 and Cohen's Kappa statistic of 0.96. This study presents a method that can support experts in the early detection of flatfoot deformity in busy work schedules. Experimental results show that the DenseNet-201 model with Dilation filtering outperforms previous works by using different image processing techniques and transfer learning approaches. This approach provides an effective solution for automatic detection of flatfoot deformity.

References

[1] T. Çit, D. Yılmaz, T. Çaviş, E. Alp, and H. Aydın, "A Software Tool Developed for Automatic Diagnosis of Pes Planus," in 11th International May 19 Innovative Scientific Approaches Congress, Samsun, Turkey, 2024, pp. 303–311.

[2] A. Azizov and Ö. Şevgin, "Pediatrik Pes Planus ve Fizyoterapi."

[3] N. Katz, "The impact of pain management on quality of life," J. Pain Symptom Manage., vol. 24, no. 1 Suppl, pp. S38–S47, Jan. 2002.

[4] S. Erkuş and Ö. Kalenderer, "Pes planovalgus," Totbid Dergisi, vol. 16, pp. 413–425, 2017.

[5] C. J. Lin, K. A. Lai, T. S. Kuan, and Y. L. Chou, "Correlating factors and clinical significance of flexible flatfoot in preschool children," J. Pediatr. Orthop., vol. 21, no. 3, pp. 378–382, 2001.

[6] D. Bordin, G. De Giorgi, G. Mazzocco, and F. Rigon, "Flat and cavus foot, indexes of obesity and overweight in a population of primary-school children," Minerva Pediatr., vol. 53, no. 1, pp. 7–13, 2001.

[7] C. Wang et al., "An efficient local binary pattern based plantar pressure optical sensor image classification using convolutional neural networks," Optik, vol. 185, pp. 543–557, 2019.

[8] Y. Kim and N. Kim, "Deep learning-based pes planus classification model using transfer learning," Journal of The Korea Society of Computer and Information, vol. 26, no. 4, pp. 21–28, 2021.

[9] J. Y. Jung, C. M. Yang, and J. J. Kim, "Decision Tree-Based Foot Orthosis Prescription for Patients with Pes Planus," International Journal of Environmental Research and Public Health, vol. 19, no. 19, p. 12484, 2022. [Online]. Available: http://www.journalwebsite.com DOI: 10.3390/ijerph191912484.

[10] J. Lauder et al., "A fully automatic system to assess foot collapse on lateral weight-bearing foot radiographs: A pilot study," Computer Methods and Programs in Biomedicine, vol. 213, p. 106507, 2022. DOI: 10.1016/j.cmpb.2022.106507.

[11] S. M. Ryu et al., "Automated landmark identification for diagnosis of the deformity using a cascade convolutional neural network (FlatNet) on weight-bearing lateral radiographs of the foot," Computers in Biology and Medicine, vol. 148, p. 105914, 2022. DOI: 10.1016/j.compbiomed.2022.105914.

[12] S. M. Ryu, K. Shin, S. W. Shin, S. Lee and N. Kim, "Enhancement of evaluating flatfoot on a weight-bearing lateral radiograph of the foot with U-Net based semantic segmentation on the long axis of tarsal and metatarsal bones in an active learning manner," Computers in Biology and Medicine, vol. 145, p. 105400, 2022. DOI: 10.1016/j.compbiomed.2022.105400.

[13] Y. Gül, S. Yaman, D. Avcı, A. H. Çilengir, M. Balaban, and H. Güler, "A novel deep transfer learning-based approach for automated Pes Planus diagnosis using X-ray image," Diagnostics, vol. 13, no. 9, p. 1662, 2023. DOI: 10.3390/diagnostics13091662. Dataset from: https://www.kaggle.com/datasets/suleyman32/pesplanus-two-class-dataset

[14] F. A. Alsaidi and K. M. Moria, "Flatfeet Severity-Level Detection Based on Alignment Measuring," Sensors, vol. 23, no. 19, p. 8219, 2023. DOI: 10.3390/s23198219.

[15] K. Doğan, T. Selçuk, and A. Yılmaz, "A Novel Model Based on CNN–ViT Fusion and Ensemble Learning for the Automatic Detection of Pes Planus," Journal of Clinical Medicine, vol. 13, no. 16, p. 4800, 2024. DOI: 10.3390/jcm13164800.

[16] J. Redmon and A. Farhadi, "YOLO9000: better, faster, stronger," in IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 7263–7271.

[17] M. Lin, "Network in network," arXiv preprint arXiv:1312.4400, 2013. [Online]. Available: http://arxiv.org/abs/1312.4400.

[18] Z. Zhou, Y. Hu, Z. Zhu, and Y. Wang, "Fabric wrinkle objective evaluation model with random vector function link based on optimized artificial hummingbird algorithm," Journal of Natural Fibers, vol. 20, no. 1, p. 2163026, 2023. DOI: 10.1080/15440478.2023.2163026.

[19] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 4700–4708.

[20] V. Miglani and M. P. S. Bhatia, "Skin lesion classification: A transfer learning approach using efficientnets," in International Conference on Advanced Machine Learning Technologies and Applications, Singapore, 2020, pp. 315–324.

[21] A. Lumini and L. Nanni, "Deep learning and transfer learning features for plankton classification," Ecological Informatics, vol. 51, pp. 33–43, 2019. DOI: 10.1016/j.ecoinf.2019.02.007.

[22] H. Kumar, A. Virmani, S. Tripathi, R. Agrawal and S. Kumar, "Transfer learning and supervised machine learning approach for detection of skin cancer: performance analysis and comparison," Transfer, vol. 10, no. 1, 2021. DOI: 10.1007/s11203-021-00207-3.

[23] M. B. Özküçük, Ö. F. Alçin and M. T. Gençoğlu, (2024). Transfer öğrenme yaklaşımı kullanılarak izolatör kusurlarının tespiti. Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi, 15(2), 323-330. DOI: 10.24012/dumf.1415322.

[24] Z. Guo, Q. Chen, G. Wu, Y. Xu, R. Shibasaki, and X. Shao, "Village building identification based on ensemble convolutional neural networks," Sensors, vol. 17, no. 11, p. 2487, 2017. DOI: 10.3390/s17112487.

[25] Y. Zou, L. Wu, C. Zuo, L. Chen, B. Zhou, and H. Zhang, "White blood cell classification network using MobileNetv2 with multiscale feature extraction module and attention mechanism," Biomedical Signal Processing and Control, vol. 99, p. 106820, 2025. DOI: 10.1016/j.bspc.2024.106820.

[26] F. Arnia, K. Saddami, and K. Munadi, "Dcnet: Noise-robust convolutional neural networks for degradation classification on ancient documents," Journal of Imaging, vol. 7, no. 7, p. 114, 2021. DOI: 10.3390/jimaging7070114.

[27] A. Ciran, and E. Özbay, "Derin Öğrenme ve Özellik Seçimi Yaklaşımları Kullanılarak Göz Hastalıkları Tespiti," Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi, vol. 15, no. 2, pp. 421-433, 2024. DOI: 10.24012/dumf.1465929

[28] S. Targ, D. Almeida, and K. Lyman, "Resnet in resnet: Generalizing residual architectures," arXiv preprint arXiv:1603.08029, 2016. [Online]. Available: http://arxiv.org/abs/1603.08029.

[29] S. Keskin, O. Sevli, and E. Okatan, "Comparative analysis of the classification of recyclable wastes," Journal of Scientific Reports-A, vol. (055), pp. 70-79, 2023. DOI: 10.59313/jsr-a.1335276

[30] S. Albahli, and T. Nazir, "AI-CenterNet CXR: An artificial intelligence (AI) enabled system for localization and classification of chest X-ray disease," Frontiers in Medicine, vol. 9, p. 955765, 2022. DOI: 10.3389/fmed.2022.955765.

[31] A. Alan and M. Karabatak, "Veri seti-sınıflandırma ilişkisinde performansa etki eden faktörlerin değerlendirilmesi," Fırat Üniversitesi Mühendislik Bilimleri Dergisi, vol. 32, no. 2, pp. 531–540, 2020.

[32] C. Hark, "Sahte haber tespiti için derin bağlamsal kelime gömülmeleri ve sinirsel ağların performans değerlendirmesi," Fırat Üniversitesi Mühendislik Bilimleri Dergisi, vol. 34, no. 2, pp. 733–742, 2022. DOI: 10.35234/fumbd.1126688

[33] S. Çelik and Ö. Kasım, "Detection of tumor slice in brain magnetic resonance images by feature optimized transfer learning," Aksaray University Journal of Science and Engineering, vol. 4, no. 2, pp. 187–198, 2020. DOI: 10.29002/asujse.820599.

[34] S. Akyol, M. Yıldırım, and B. Alataş, "Multi-feature fusion and improved BO and IGWO metaheuristics-based models for automatically diagnosing the sleep disorders from sleep sounds," Computers in Biology and Medicine, vol. 157, p. 106768, 2023. DOI: 10.1016/j.compbiomed.2023.106768.

[35] Ö. B. Bilgen and N. Doğan, "Puanlayıcılar arası güvenirlik belirleme tekniklerinin karşılaştırılması," Journal of Measurement and Evaluation in Education and Psychology, vol. 8, no. 1, pp. 63–78, 2017. DOI: 10.21031/epod.294847.

[36] N. W. S. Wardhani, M. Y. Rochayani, A. Iriany, A. D. Sulistyono, and P. Lestantyo, "Cross-validation metrics for evaluating classification performance on imbalanced data," 2019 International Conference on Computer, Control, Informatics, 2019.