Turkish Journal of Engineering
https://dergipark.org.tr/en/pub/tuje
e-ISSN 2587-1366

Comparison of CNN-based methods for yoga pose classification

Vildan Atalay Aydın *[1]

[1]*İzmir Demokrasi University, Department of Computer Engineering, Türkiye, vildan.atalayaydin@idu.edu.tr*

Cite this study: Aydın, V. A. (2024). Comparison of CNN-based methods for yoga pose classification. Turkish Journal of Engineering, 8 (1), 65-75

Keywords
Image classification
Wavelet transform
Deep learning
Yoga posture
CNN

Research Article
DOI: 10.31127/tuje.1275826

Received:02.04.2023
Revised: 07.05.2023
Accepted:11.05.2023
Published:15.09.2023

check for updates

Abstract

Yoga is an exercise developed in ancient India. People perform yoga in order to have mental, physical, and spiritual benefits. While yoga helps build strength in the mind and body, incorrect postures might result in serious injuries. Therefore, yoga exercisers need either an expert or a platform to receive feedback on their performance. Since access to experts is not an option for everyone, a system to provide feedback on the yoga poses is required. To this end, commercial products such as smart yoga mats and smart pants are produced; Kinect cameras, sensors, and wearable devices are used. However, these solutions are either uncomfortable to wear or not affordable for everyone. Nonetheless, a system that employs computer vision techniques is a requirement. In this paper, we propose a deep-learning model for yoga pose classification, which is the first step of a quality assessment and personalized feedback system. We introduce a wavelet-based model that first takes wavelet transform of input images. The acquired subbands, i.e., approximation, horizontal, vertical, and diagonal coefficients of the wavelet transform are then fed into separate convolutional neural networks (CNN). The obtained probability results for each group are fused to predict the final yoga class. A publicly available dataset with 5 yoga poses is used. Since the number of images in the dataset is not enough for a deep learning model, we also perform data augmentation to increase the number of images. We compare our results to a CNN model and the three models that employ the subbands separately. Results obtained using the proposed model outperforms the accuracy output achieved with the compared models. While the regular CNN model has 61% and 50% accuracy for the training and test data, the proposed model achieves 91% and 80%, respectively.

1. Introduction

Especially with the lockdowns caused by the Covid-19 pandemic, people started to tend towards exercises such as yoga and pilates that they can perform at home. Even though, when performed right, the health benefits of such exercises are undeniable, incorrect moves are known to cause serious injuries. Therefore, there is a need for an expert to give feedback to the exerciser. Nonetheless, since not everyone has time or money to attend a class, and the pandemic is the reason for exercising at home in the first place, the need for an expert's opinion is replaced with the requirement of a user application. With the help of an application, the exercisers can upload a picture of themselves performing a pose, and receive feedback on either the correctness or the quality of their asana.

Kinect cameras [1], wearable devices [2], Internet of Things (IoT) [3], and smart mats/pants are used for action recognition during exercises. However, these devices might be expensive or uncomfortable to wear during practice. As a result, researchers take advantage of computer vision techniques such as action recognition, pose estimation, and image classification. Several studies exploit deep learning and transfer learning methods in order to do so.

Deep learning is a powerful technique in many recent research areas. Researchers have also adopted deep learning methods in yoga pose classification. As an example, Jain et al. [4] design a 3D Convolutional Neural Network (CNN) to exploit the spatial-temporal relationship among Yoga poses for a real-time video dataset. Since the architecture of state-of-art image classifiers such as VGG and ImageNet is based on feature extraction by downsizing, these classifiers require high-resolution images. Therefore, Gochoo et al. [5] propose a privacy-preserving classification method for low-resolution images using deep learning. They utilize a feature-preserving architecture for 16x16 images, with variable ReLU slopes and a custom loss function.

A large number of studies in yoga pose classification make use of pose estimation together with deep learning. Some studies propose pose estimation and feedback generation using extraction of body joints (e.g., OpenPose) and finding differences in angles between the expert's and the user's pose [6, 7]. An angle-likelihood mechanism is also used by [8]. The authors propose a coarse-to-fine algorithm to classify yoga poses by first training a DenseNet classifier to predict image class at a coarse level. Later, a pre-trained pose estimator is employed to find noisy keypoints from input images to generate joint angles of the human body. Finally, to predict the pose class, a K-Nearest Neighbors classifier is trained on the pose vectors generated using the joint angles and the output probability vector of DenseNet. Moreover, Wu et al. [9] propose a yoga pose grading system by first extracting the keypoints of a human skeleton and later training with contrastive triplet examples. Garg et al. [10] utilize MediaPipe library to skeletonize input images and later compare several deep learning methods with and without the skeletonization process to conclude that VGG16 outperforms others if images are not skeletonized, and their proposed deep learning method performs best when skeletonized. On the other hand, Swain et al. [11] also use MediaPipe library for keypoint extraction. Features of keypoints are extracted via CNN, and Long-Short Term Memory (LSTM) layer is used to understand the sequence of video frames for predictions. In a similar method, Rishan et al. [12] propose a technique that uses OpenPose to detect keypoints and employ CNN, Long Short Term Memory (LSTM), and SoftMax regression for classification. Furthermore, Yadav et al. [13] propose another model that uses OpenPose to extract keypoints. The authors collect a video dataset for six Yoga poses and later employ CNN and LSTM to extract keypoint features and obtain temporal predictions, respectively.

When the amount of data at hand is not enough for a deep learning solution, transfer learning comes into play. Long et al. [14] proposed a yoga posture coaching system based on transfer learning. They compared VGG16, VGG19, MobileNet, MobileNetV2, InceptionV3, and DenseNet201 for classification and concluded that the MobileNet model was optimal on their collected dataset consisting of 14 different yoga postures performed by eight participants 10 times each. On the other hand, Chasmai et al. [15] propose using transfer learning, AlphaPose, to be specific, to extract the human body keypoints for human pose estimation. Later, a random forest classifier is trained for the classification of an in-house collected video dataset of yoga poses.

While some studies work on publicly available datasets and some use their in-house generated ones, some of them create challenging datasets for yoga pose classification. Verma et al. [16] propose a large-scale challenging dataset including pose diversity, object occlusion, and viewpoints for 82 yoga pose classes. The dataset has a fine-grained hierarchical structure that separates classes by body posture variations. The authors compare Resnet, DenseNet, MobileNet, and ResNext models on their proposed dataset and conclude that DenseNet-201 outperforms other models. Moreover, Yadav et al. [17] creates a video dataset including 5484 videos in 82 classes of yoga poses. They also propose an architecture with three parallel components which use the part affinity fields model, optical flow, and raw RGB videos to classify yoga asanas. Finally, Li et al. [18] introduce a 3D yoga pose dataset with 117 classes, including 3792 action samples where each sample consists of an RGB image, a human skeleton, a pose label, and a quality score for both action recognition and activity quality assessment. The authors also provide a two-stream adaptive graph CNN to recognize and access the poses.

Image classification is a process where deep learning is widely and successfully used to categorize images. There are two main types of classification, namely, supervised and unsupervised. While supervised classification uses a training set of data to predict classes for new images, unsupervised classification clusters images based on their characteristics without the need for a training dataset. Image classification has numerous application areas, including but not limited to medical applications [19], object recognition and detection, face recognition, and image segmentation, to name a few [20]. Lu and Weng [21] provide a review of image classification methods, especially in the remote sensing area. For a successful classification of remote sensing data, the authors advise employing multiple features and a suitable classification technique.

Deep learning techniques are employed effectively in image classification [22, 23], as in many other areas [24]. An interested reader can find a detailed review of the application of convolutional neural networks (CNN) to image classification task in [25]. In their paper, the authors provide an explanation of the CNN architecture, their development and early successes, and their application to image classification by reviewing over 300 publications. Su et al. [26] also provide a comprehensive work on the robustness of several deep learning methods for image classification. Due to the fact that the accuracy metric is used as a comparison technique for classification methods and the lack of robustness of these methods are studied in the literature, the authors compare 18 models using more robust comparison metrics, including distortion, success rate, and transferability of adversarial examples. They conclude that the empirical distortion metrics scale linearly with the logarithm of the error; model architecture is more important than its size; and increasing the depth of the architecture only slightly affects robustness.

Wavelets have also been employed in many areas such as super-resolution, image registration, video coding, etc. [27–34] due to their nature of local extraction of spectral and temporal information of images.

Wavelet-based methods are applied to image classification problem successfully. As an example, Li et al. [35] propose replacing the max pooling, strided-convolution, and average pooling layers with Discrete Wavelet Transform in order to reduce the noise-prone nature of CNNs. They decompose the feature maps into low- and high-frequency subbands. Later, the high-frequency subbands are dropped to avoid the noise included, and low-frequency subbands with the information on the basic data structures in images are fed into the next layers. Their comparisons on ImageNet [36] and ImageNet-C [37] confirm that their proposed wavelet-based method outperforms state-of-the-art techniques.

Image classification is widely used in medical applications in order to help experts in diagnostic radiology and disease detection. Wavelet-based deep learning methods are also used in medical image classification tasks. For example, Mallick et al. [38] propose a deep wavelet-autoencoder-based neural network architecture for brain MRI image classification. The paper aims to build a system for cancer detection using image classification. The authors use the deep wavelet autoencoder for image compressing in order to combine the properties of feature reduction and image decomposition of the autoencoder and wavelet decomposition, respectively. They use the autoencoder together with a Deep Neural Network for brain MRI image classification. Their comparisons prove that the proposed wavelet-based method outperforms existing classifiers based on accuracy.

Khatami et al. [39] also propose a wavelet decomposition method for medical X-ray image classification. Their proposed method employs the approximation (low-pass) subbands of the wavelet decomposition only, and provide promising results on disease detection.

Said et al. [40] combine wavelet network with deep learning for first supervised and later unsupervised classification. They experiment on COIL-100 and MNIST datasets to present that the proposed work outputs promising results.

Even though wavelet-based deep learning is widely and effectively used in medical image classification literature, its benefits have not yet been exploited for yoga pose classification. Motivated by the conclusion made by Fujieda et al. [41] on wavelet-CNN for texture classification which stated that the wavelet-CNN achieves higher accuracy results together with having a lower number of parameters than conventional CNN models, our goal in this paper is to explore the effects of wavelet transform on a CNN-based image classification technique applied to yoga pose images. Several deep-learning methods use the wavelet transform for numerous vision problems [30, 42, 43]. To the best of our knowledge, this study is the first attempt to employ wavelet transform-based deep learning in yoga pose classification. Inspired by the work of Serte and Demirel [44], instead of classifying original images, we perform wavelet transform on the images to obtain wavelet subbands. Later, our model is trained using the subbands only. Predictions obtained with approximation, horizontal, and vertical subbands are later fused together in order to achieve the final classes of images. We compare our results to a CNN model trained on the original data to conclude that the proposed method overperforms the CNN model.

The paper is organized as follows. In Section 2, we introduce the methodology proposed. We represent the results obtained in Section 3, discuss the application areas of our method in Section 4, and finally, concluding remarks are provided in Section 5.

2. Method

In this section, we introduce our proposed model for yoga pose classification.

Inspired by the work of Serte and Demirel [44], we investigate the effect of wavelet transform on classifying yoga poses. In their paper, Serte and Demirel [44] proposes taking wavelet transform of input skin lesion images and training their transfer learning (i.e., ResNet) models on two sets of data. The first dataset consists of the original skin images, together with one-level approximation, horizontal and vertical wavelet subbands; while the second dataset includes the original images together with approximation coefficients on the first, second, and third decompositions. Their proposed method outperforms the method based only on the original image dataset. The authors also present a method that uses Gabor wavelets in [43], which decomposes the image data into seven directional wavelet subbands.

In this paper, we propose a CNN model that is trained on only the wavelet subbands instead of a combination with the original data. Our goal is to demonstrate that training on the wavelet subbands only results in higher accuracy than training on the original yoga pose images. Unlike Serte and Demirel's work in [44], our model does not use the original images in training; instead, it employs one-level approximation, horizontal, vertical, and diagonal detail coefficients. The probabilities obtained for all subbands are later fused to reach the final classification of the images, as [44].

In Figure 1, we demonstrate the proposed wavelet CNN model for yoga pose classification. As seen in the figure, we first decompose input images into wavelet coefficients, i.e., approximation, horizontal, vertical, and diagonal subbands. Later, each subband dataset is fed into a separate CNN to be trained. After the training step, prediction scores for all subbands are fused together in order to achieve the final output class [44]. The sum of probabilities formula used is given in Equation 1.

$$s_i = \frac{\sum_{j=1}^{n} s_{ij}}{\sum_{i=1}^{c} \sum_{j=1}^{n} s_{ij}} \tag{1}$$

where s represents the score, c is the number of classes, and n is the number of CNN models.

We also show the CNN model used to predict the 5 yoga classes in Figure 2. We utilize a 3x3 convolution kernel in the convolution layer. Our previous work on the comparison of activation functions for yoga pose classification shows that the ReLU function outperforms the tanh and leaky ReLU functions widely used in CNN models for image classification [45]. The ReLU function acquires higher accuracy in less time compared to the methods mentioned before. Therefore, we employ the ReLU activation function in this work as well. After each pooling layer with max pooling, we employ a dropout layer which helps prevent the overfitting problem in many optimization tasks. Finally, one of the five classes is predicted for each input after a fully connected layer. When the prediction probability for each subband is achieved, the probabilities are fused in order to obtain the final class prediction for each input image.

For our studies, we used a dataset for yoga classification provided in Kaggle [46]. A set of example images is presented in Figure 3. The dataset consists of RGB images in 5 classes, i.e., Downdog, Goddess, Plank, Tree, and Warrior, with varying numbers of images for each class. While the training dataset has 1081 images in total, including 223 images for downdog, 180 for goddess, 266 for plank, 160 for tree, and 252 for warrior pose; the test dataset contains 470 in total, 97 images for downdog, 80 for goddess, 115 for plank, 69 for tree, and 109 for warrior pose. We separate our training dataset into two parts for training and validation by reserving 20% of the training set for validation.

The next section will display the results obtained using the proposed method.

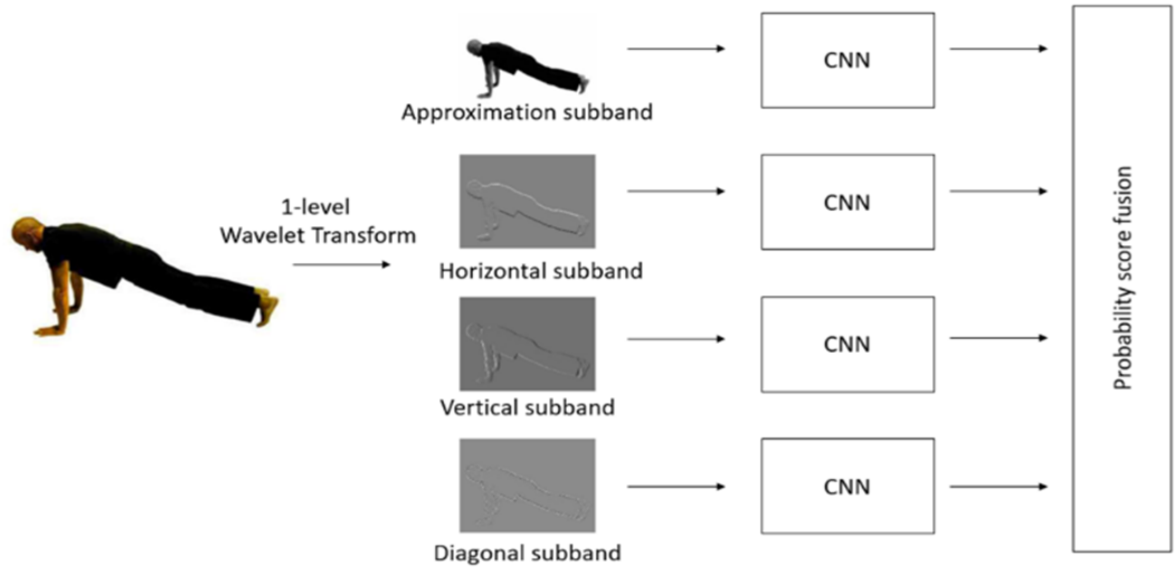

Figure 1. Proposed wavelet CNN model for yoga pose classification.

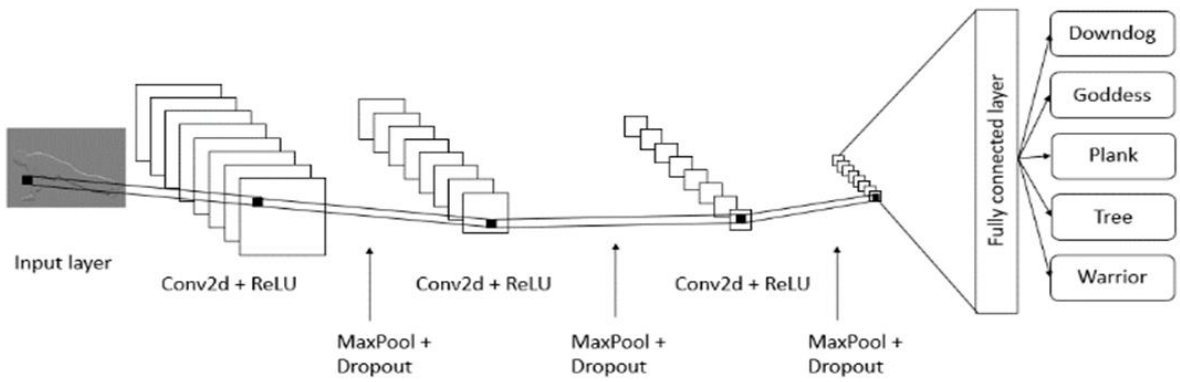

Figure 2. CNN model.

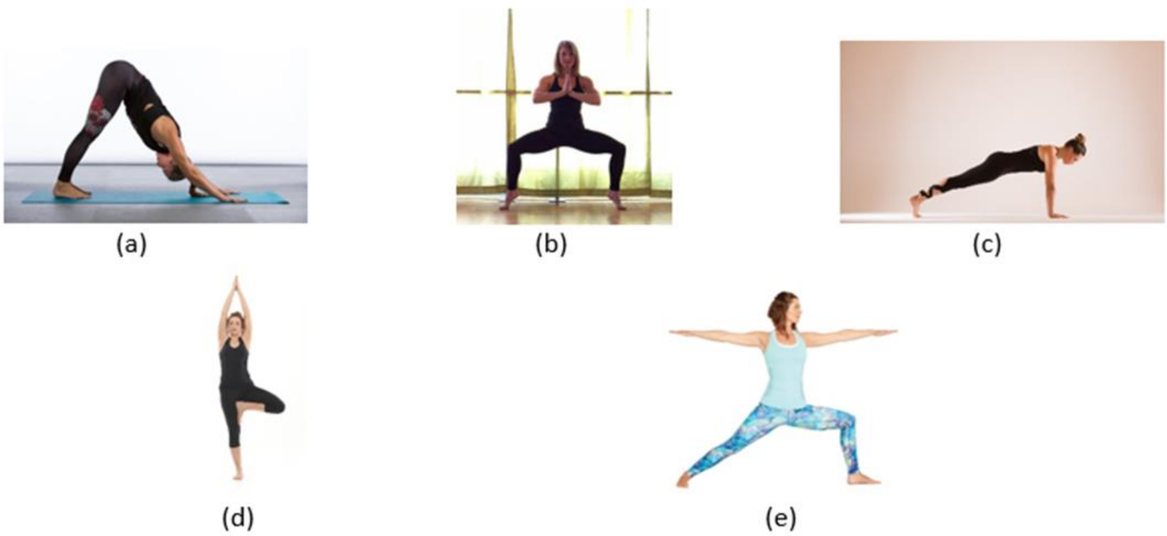

Figure 3. Sample images in training dataset (a) Downdog (b) Goddess (c) Plank (d) Tree (e) Warrior.

3. Results

We demonstrate the results using the proposed method in this section.

The experiments are run on MATLAB R2021a on a laptop computer with 8 GB RAM and a single CPU @1.60 GHz. The goal of this paper is to compare the accuracy results obtained using the proposed combined method to

a CNN-based method and the subbands separately, instead of a comparison on the time consumption. Since the proposed method requires training on all four subbands, including the first level approximation, horizontal, vertical, and diagonal coefficients of the wavelet transform, the runtime of the proposed combined method is higher than the training times for each subband separately.

Since this work aims to demonstrate the effect of wavelet transform on yoga pose classification, we do not employ a transfer learning-based method. Therefore, the dataset used in this study must be augmented to increase the number of images. Only horizontal flip is used in data augmentation. When the employed dataset is investigated more in detail, one can see that the images do not rotate. However, as in Figure 4, horizontally flipped images exist. Therefore, we apply horizontal flip for data augmentation but not rotation. The augmentation step is not applied to validation and test datasets.

Figure 4. Horizontally flipped images from the dataset.

All images are resized to 224x224 before training. Since Discrete Wavelet Transform (DWT) is a 2D transform, all input data is converted to grayscale before the training step, and DWT is applied later.

Due to the fact that we employ wavelet transform in this work, we have four separate sets of data for each class mentioned above. Meaning, our training dataset for the approximation subband, as an example, has 1081 images in total for all classes, and the test dataset has 470 images, as well as the horizontal, vertical, and diagonal datasets.

We demonstrate the accuracy and loss graphs generated in MATLAB in Figure 5-9. Figure 5 presents the accuracy and loss results for the CNN model in Figure 2, while Figure 6 demonstrates the accuracy and loss results for approximation, Figure 7 shows horizontal, Figure 8 displays vertical, and finally, Figure 9 presents diagonal subbands, separately. Investigating the graphs, one can see that the CNN model does not provide highly accurate results when used alone because the dataset is not large enough for a deep learning method. The approximation subband, like the regular CNN model applied to the original images, also has limitations on achieving high accuracies, since the approximation subband is actually a low-resolution version of the original ones. Due to the fact that the horizontal, vertical, and diagonal subbands consist of the details in the images, compared to the approximation subbands and the original images, these subbands have higher accuracy and lower loss results.

Table 1 summarizes the obtained results using the proposed method compared to the CNN model and the subbands separately. For the comparison to be fair, we use the same model for CNN, shown in Figure 2, for all models. For the CNN model, all images are resized to 224x224x3, and data augmentation is also employed with a horizontal flip.

One can see from Table 1 that, the combined (i.e., proposed) model outperforms the CNN model, and the model's using approximation, horizontal, vertical, and diagonal subbands one by one separately. Since the approximation subband is a low-resolution representation of the original images, the accuracy achieved using only the approximation subband cannot match the results obtained with the CNN model. We can observe that the detail subbands, i.e., the horizontal, vertical, and diagonal ones, have higher accuracies than the CNN and approximation counterparts because the high-frequency subbands have detailed information on images. While the accuracy results achieved using the subbands separately are not high enough, we can see that the combination of the probabilities achieved with the subbands provides a promising outcome.

Table 2 demonstrates the comparison of the proposed method to the transfer learning methods based on the accuracy results of the training data. It can be observed from the table that the proposed technique outperforms the transfer learning methods of ResNet18, ResNet50, and GoogleNet. All compared transfer learning models are fed the original images as input and the same data augmentation step is also applied before training.

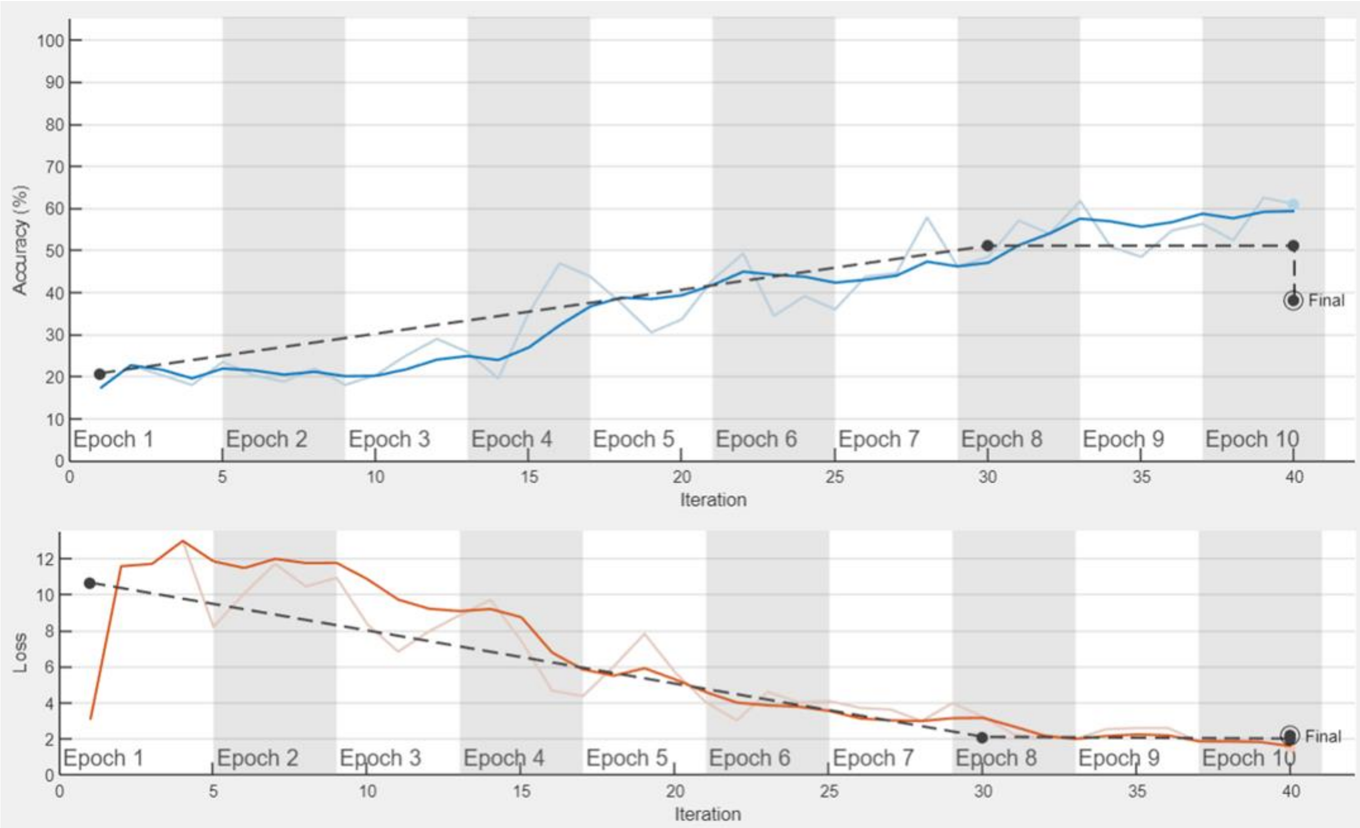

Figure 5. Accuracy and loss graphs for the CNN model.

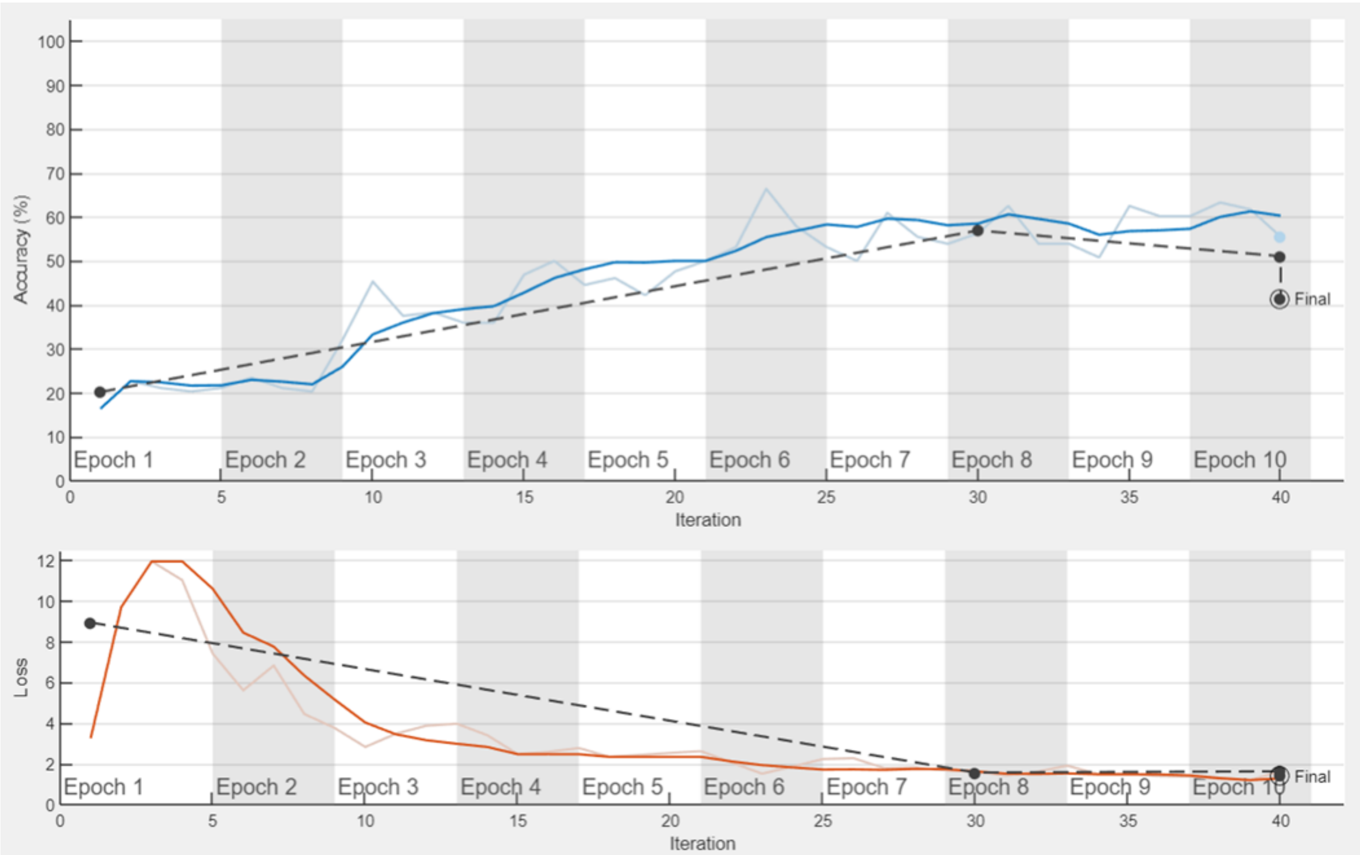

Figure 6. Accuracy and loss graphs for approximation subband.

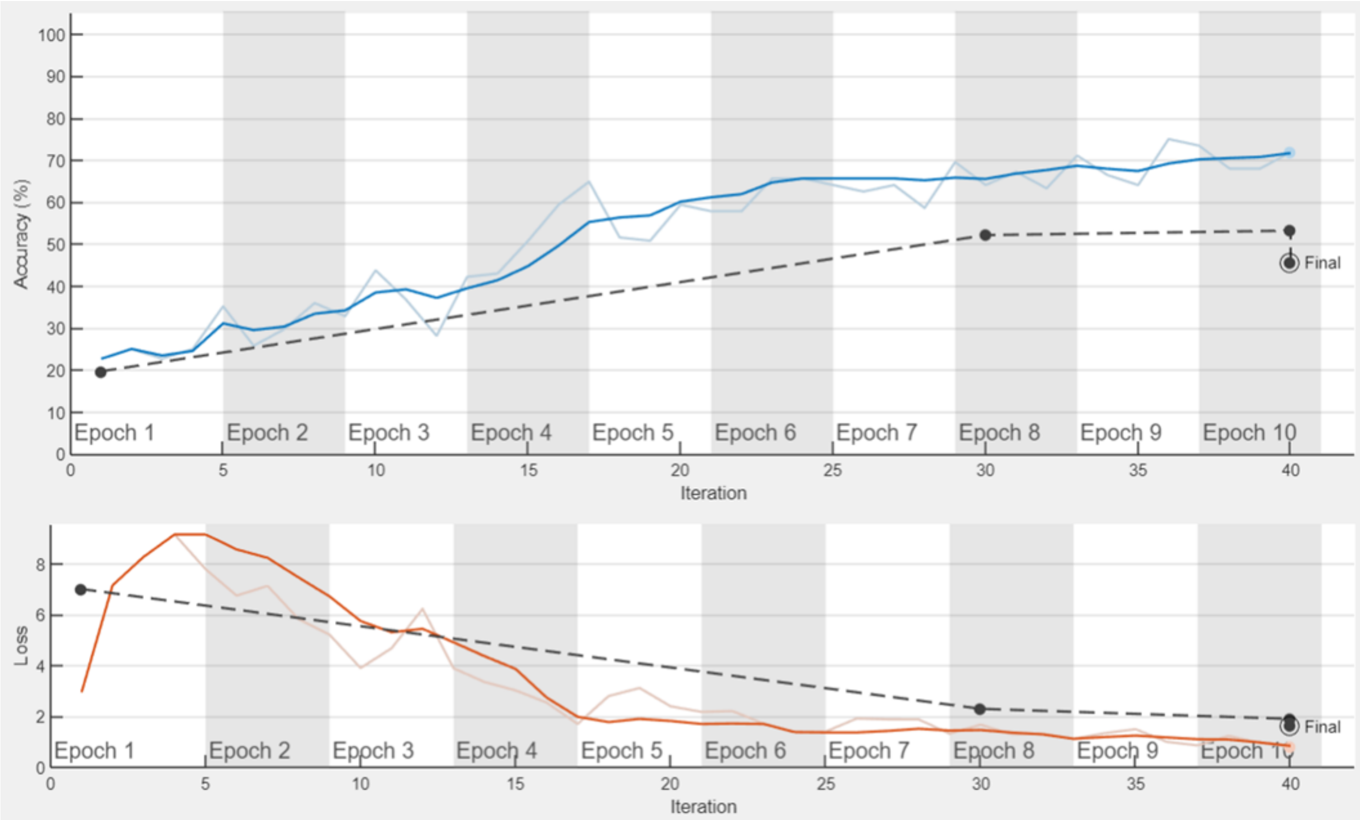

Figure 7. Accuracy and loss graphs for horizontal subband.

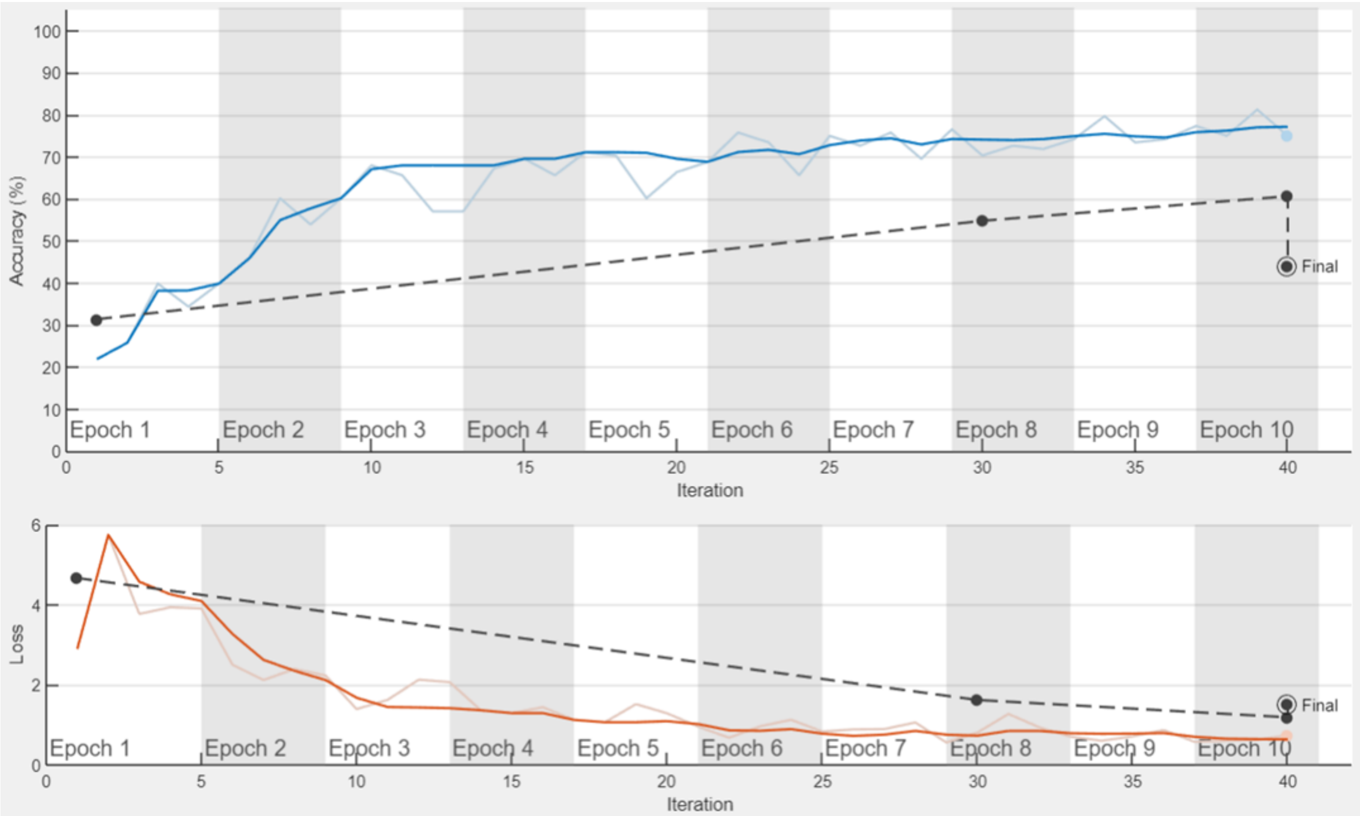

Figure 8. Accuracy and loss graphs for vertical subband.

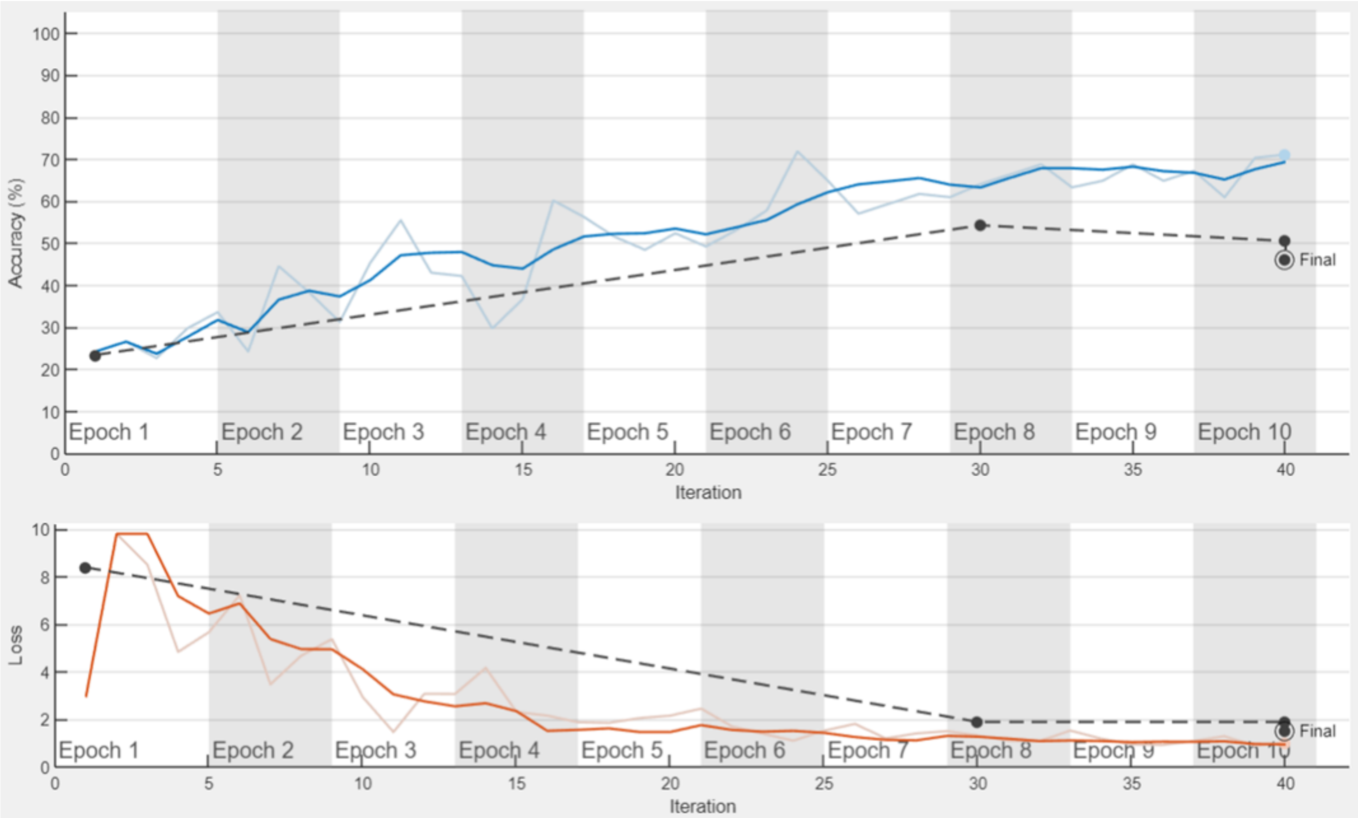

Figure 9. Accuracy and loss graphs for diagonal subband.

Table 1. Accuracy results obtained with the proposed method compared to a CNN model and the CNN model applied to the subbands separately.

Model	Training data	Test data
CNN	0.61	0.51
Approx.	0.55	0.51
Horizontal	0.72	0.53
Vertical	0.75	0.61
Diagonal	0.71	0.51
Proposed	0.91	0.8

Table 2. A comparison between transfer learning methods and the proposed model.

Model	Training data
ResNet18	0.80
ResNet50	0.83
GoogleNet	0.86
Proposed	0.91

4. Discussion

This paper investigates the effects of employing wavelet subbands on a deep learning-based yoga pose classification problem. The proposed method aims at increasing the accuracy results obtained with a regular CNN method by employing wavelet subbands.

Even though the proposed CNN-based method outperforms the compared methods and does have promising results, since each subband of wavelet transform is fed into separate CNNs and the methodology is not parallelized, it requires higher run time compared to regular CNN, which is applied on the original images. However, since wavelet subband images have lower resolution compared to the original images, and each subband is separately fed into CNNs, employing parallelization would decrease the runtime to allow for the proposed method to be effectively used in real-time applications.

The proposed method can also be employed in 3D pose estimation. Chen and Ramanan [47] argue that instead of predicting 3D poses from image measurements, it is more accurate to use 2D pose estimation with 3D mocap libraries. They experiment with off-the-shelf 2D pose estimation systems to prove their argument. Our proposed 2D pose estimation system can also be used in such scenarios.

The proposed system is the first step of a personalized feedback system. As in the previously mentioned publications, the image classification method can be incorporated into a feedback-providing application where the users either upload their pictures or perform their asanas in front of a camera to rate their performance.

Misclassification of images occurs due to the fact that some poses look alike, as shown in Figure 10. The image on the left is in the Warrior dataset, while the one on the right is in Goddess. One can see the similarity between poses, which results in misclassified images in the datasets. Incorporating skeletonization and angle-likelihoods into the proposed method would decrease the rate of misclassified images as in previously mentioned publications.


Figure 10. Misclassified images from the dataset.

5. Conclusion

Yoga is an ancient exercise performed by people for its many benefits. Although it helps gain strength in the body and mind, and its benefits are innumerable when performed correctly, an inaccurate move can cause serious health problems. With the increase in the number of people exercising at home after the COVID-19 pandemic, the need for systems to determine whether the poses are done correctly has increased. Since the commercial products produced are not accessible by everyone; computer vision methods such as image classification came in handy. In this paper, we propose a wavelet-based CNN model for yoga pose classification. To the best of our knowledge, this work is the first attempt at exploring the effects of wavelet decomposition in the yoga pose detection problem. Since wavelet transform exhibits the details in images, as expected, decomposing images into wavelet subbands to train the models, in fact, increases the accuracy compared to the one acquired by training on the original images. We propose first decomposing input images into wavelet subbands, later these subbands are fed into separate CNN models. When the output probabilities for each subband are achieved, we fuse the probability results in order to find the final prediction of the yoga pose class for an input image. Demonstrated results outperform the training outcomes achieved using the original images or the separate subbands while having promising accuracy percentages. Since the datasets used are not large enough for a deep learning method to have high accuracies, the next step of this work will include employing a transfer learning method to increase the accuracy results obtained in this paper.

Conflicts of interest

The authors declare no conflicts of interest.

References

1. Chang, C. W., Da Nian, M., Chen, Y. F., Chi, C. H., & Tao, C. W. (2014, August). Design of a Kinect sensor based posture recognition system. In 2014 Tenth International Conference on Intelligent Information Hiding and Multimedia Signal Processing (pp. 856-859). IEEE. https://doi.org/10.1109/IIH-MSP.2014.216
2. Wang, J., Huang, Z., Zhang, W., Patil, A., Patil, K., Zhu, T., ... & Harris, T. B. (2016, December). Wearable sensor based human posture recognition. In 2016 IEEE International conference on big data (big data) (pp. 3432-3438). IEEE. https://doi.org/10.1109/BigData.2016.7841004
3. Gochoo, M., Tan, T. H., Huang, S. C., Batjargal, T., Hsieh, J. W., Alnajjar, F. S., & Chen, Y. F. (2019). Novel IoT-based privacy-preserving yoga posture recognition system using low-resolution infrared sensors and deep learning. IEEE Internet of Things Journal, 6(4), 7192-7200. https://doi.org/10.1109/JIOT.2019.2915095
4. Jain, S., Rustagi, A., Saurav, S., Saini, R., & Singh, S. (2021). Three-dimensional CNN-inspired deep learning architecture for Yoga pose recognition in the real-world environment. Neural Computing and Applications, 33, 6427-6441. https://doi.org/10.1007/s00521-020-05405-5
5. Gochoo, M., Tan, T. H., Alnajjar, F., Hsieh, J. W., & Chen, P. Y. (2020, October). Lownet: Privacy preserved ultra-low resolution posture image classification. In 2020 IEEE International Conference on Image Processing (ICIP) (pp. 663-667). IEEE. https://doi.org/10.1109/ICIP40778.2020.9190922
6. Anand Thoutam, V., Srivastava, A., Badal, T., Kumar Mishra, V., Sinha, G. R., Sakalle, A., ... & Raj, M. (2022). Yoga pose estimation and feedback generation using deep learning. Computational Intelligence and Neuroscience, 4311350. https://doi.org/10.1155/2022/4311350
7. Kumar, D., & Sinha, A. (2020). Yoga pose detection and classification using deep learning. International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 6 (6), 160-184. https://doi.org/10.32628/CSEIT206623

8. Dittakavi, B., Bavikadi, D., Desai, S. V., Chakraborty, S., Reddy, N., Balasubramanian, V. N., ... & Sharma, A. (2022). Pose tutor: an explainable system for pose correction in the wild. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 3540-3549).
9. Wu, Y., Lin, Q., Yang, M., Liu, J., Tian, J., Kapil, D., & Vanderbloemen, L. (2021, December). A computer vision-based yoga pose grading approach using contrastive skeleton feature representations. Healthcare, 10(1), 36. https://doi.org/10.3390/healthcare10010036
10. Garg, S., Saxena, A., & Gupta, R. (2022). Yoga pose classification: a CNN and MediaPipe inspired deep learning approach for real-world application. Journal of Ambient Intelligence and Humanized Computing, 1-12. https://doi.org/10.1007/s12652-022-03910-0
11. Swain, D., Satapathy, S., Acharya, B., Shukla, M., Gerogiannis, V. C., Kanavos, A., & Giakovis, D. (2022). Deep Learning Models for Yoga Pose Monitoring. Algorithms, 15(11), 403. https://doi.org/10.3390/a15110403
12. Rishan, F., De Silva, B., Alawathugoda, S., Nijabdeen, S., Rupasinghe, L., & Liyanapathirana, C. (2020, December). Infinity yoga tutor: Yoga posture detection and correction system. In 2020 5th International conference on information technology research (ICITR) (pp. 1-6). IEEE. https://doi.org/10.1109/ICITR51448.2020.9310832
13. Yadav, S. K., Singh, A., Gupta, A., & Raheja, J. L. (2019). Real-time Yoga recognition using deep learning. Neural Computing and Applications, 31, 9349-9361. https://doi.org/10.1007/s00521-019-04232-7
14. Long, C., Jo, E., & Nam, Y. (2022). Development of a yoga posture coaching system using an interactive display based on transfer learning. The Journal of Supercomputing, 78, 5269–5284. https://doi.org/10.1007/s11227-021-04076-w
15. Chasmai, M., Das, N., Bhardwaj, A., & Garg, R. (2022). A View Independent Classification Framework for Yoga Postures. SN computer science, 3(6), 476. https://doi.org/10.1007/s42979-022-01376-7
16. Verma, M., Kumawat, S., Nakashima, Y., & Raman, S. (2020). Yoga-82: a new dataset for fine-grained classification of human poses. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops (pp. 1038-1039).
17. Yadav, S. K., Singh, G., Verma, M., Tiwari, K., Pandey, H. M., Akbar, S. A., & Corcoran, P. (2022). YogaTube: a video benchmark for Yoga action recognition. In 2022 International Joint Conference on Neural Networks (IJCNN) (pp. 1-8). IEEE. https://doi.org/10.1109/IJCNN55064.2022.9892122
18. Li, J., Hu, H., Li, J., & Zhao, X. (2022). 3D-Yoga: A 3D Yoga Dataset for Visual-Based Hierarchical Sports Action Analysis. In Proceedings of the Asian Conference on Computer Vision (pp. 434-450).
19. Fu, Y., Lei, Y., Wang, T., Curran, W. J., Liu, T., & Yang, X. (2020). Deep learning in medical image registration: a review. Physics in Medicine & Biology, 65(20), 20TR01. https://doi.org/10.1088/1361-6560/ab843e
20. Yang, W., Zhang, X., Tian, Y., Wang, W., Xue, J. H., & Liao, Q. (2019). Deep learning for single image super-resolution: A brief review. IEEE Transactions on Multimedia, 21(12), 3106-3121. https://doi.org/10.1109/TMM.2019.2919431
21. Lu, D., & Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance. International journal of Remote sensing, 28(5), 823-870. https://doi.org/10.1080/01431160600746456
22. Gülgün, O. D., & Hamza, E. R. O. L. (2020). Classification performance comparisons of deep learning models in pneumonia diagnosis using chest x-ray images. Turkish Journal of Engineering, 4(3), 129-141. https://doi.org/10.31127/tuje.652358
23. Zeybek, M. (2021). Classification of UAV point clouds by random forest machine learning algorithm. Turkish Journal of Engineering, 5(2), 48-57. https://doi.org/10.31127/tuje.669566
24. Öztürk, A., Allahverdi, N., & Saday, F. (2022). Application of artificial intelligence methods for bovine gender prediction. Turkish Journal of Engineering, 6(1), 54-62. https://doi.org/10.31127/tuje.807019
25. Rawat, W., & Wang, Z. (2017). Deep convolutional neural networks for image classification: A comprehensive review. Neural computation, 29(9), 2352-2449. https://doi.org/10.1162/neco_a_00990
26. Su, D., Zhang, H., Chen, H., Yi, J., Chen, P. Y., & Gao, Y. (2018). Is Robustness the Cost of Accuracy?--A Comprehensive Study on the Robustness of 18 Deep Image Classification Models. In Proceedings of the European conference on computer vision (ECCV), 631-648.
27. Aydin, V. A., & Foroosh, H. (2017, September). Motion compensation using critically sampled dwt subbands for low-bitrate video coding. In 2017 IEEE International Conference on Image Processing (ICIP), 21-25). https://doi.org/10.1109/ICIP.2017.8296235
28. Aydin, V. A., & Foroosh, H. (2017). In-band sub-pixel registration of wavelet-encoded images from sparse coefficients. Signal, Image and Video Processing, 11, 1527-1535. https://doi.org/10.1007/s11760-017-1116-5
29. Aydin, V. A., & Foroosh, H. (2018). A linear well-posed solution to recover high-frequency information for super resolution image reconstruction. Multidimensional Systems and Signal Processing, 29, 1309-1330. https://doi.org/10.1007/s11045-017-0499-3
30. Huang, H., He, R., Sun, Z., & Tan, T. (2017). Wavelet-srnet: A wavelet-based cnn for multi-scale face super resolution. In Proceedings of the IEEE international conference on computer vision (pp. 1689-1697).
31. Le Moigne, J., Campbell, W. J., & Cromp, R. F. (2002). An automated parallel image registration technique based on the correlation of wavelet features. IEEE Transactions on Geoscience and Remote Sensing, 40(8), 1849-1864. https://doi.org/10.1109/TGRS.2002.802501

32. Postalcıoğlu, S., Erkan, K., & Bolat, E. D. (2005). Comparison of wavenet and neuralnet for system modeling. In Knowledge-Based Intelligent Information and Engineering Systems: 9th International Conference, KES 2005, Melbourne, Australia, September 14-16, 2005, Proceedings, Part II 9 (pp. 100-107). Springer Berlin Heidelberg. https://doi.org/10.1007/11552451_14
33. Postalcioglu, S., & Becerikli, Y. (2007). Wavelet networks for nonlinear system modeling. Neural Computing and Applications, 16, 433-441. https://doi.org/10.1007/s00521-006-0069-3
34. Robinson, M. D., Toth, C. A., Lo, J. Y., & Farsiu, S. (2010). Efficient Fourier-wavelet super-resolution. IEEE Transactions on Image Processing, 19(10), 2669-2681. https://doi.org/10.1109/TIP.2010.2050107
35. Li, Q., Shen, L., Guo, S., & Lai, Z. (2020). Wavelet integrated CNNs for noise-robust image classification. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 7245-7254.
36. Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009, June). Imagenet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition, 248-255. https://doi.org/10.1109/CVPR.2009.5206848
37. Hendrycks, D., & Dietterich, T. (2019). Benchmarking neural network robustness to common corruptions and perturbations. arXiv preprint arXiv:1903.12261. https://doi.org/10.48550/arXiv.1903.12261
38. Mallick, P. K., Ryu, S. H., Satapathy, S. K., Mishra, S., Nguyen, G. N., & Tiwari, P. (2019). Brain MRI image classification for cancer detection using deep wavelet autoencoder-based deep neural network. IEEE Access, 7, 46278-46287. https://doi.org/10.1109/ACCESS.2019.2902252
39. Khatami, A., Nazari, A., Beheshti, A., Nguyen, T. T., Nahavandi, S., & Zieba, J. (2020, July). Convolutional neural network for medical image classification using wavelet features. In 2020 International Joint Conference on Neural Networks (IJCNN), 1-8. https://doi.org/10.1109/IJCNN48605.2020.9206791
40. Said, S., Jemai, O., Hassairi, S., Ejbali, R., Zaied, M., & Amar, C. B. (2016, October). Deep wavelet network for image classification. In 2016 IEEE International conference on systems, man, and cybernetics (SMC), 000922-000927. https://doi.org/10.1109/SMC.2016.7844359
41. Fujieda, S., Takayama, K., & Hachisuka, T. (2017). Wavelet convolutional neural networks for texture classification. arXiv preprint arXiv:1707.07394. https://doi.org/10.48550/arXiv.1707.07394
42. Postalcioglu, S. (2022). Design of Automatic Tool for Diagnosis of Pneumonia Using Boosting Techniques. Brazilian Archives of Biology and Technology, 65, e22210322.
43. Serte, S., & Demirel, H. (2019). Gabor wavelet-based deep learning for skin lesion classification. Computers in biology and medicine, 113, 103423. https://doi.org/10.1016/j.compbiomed.2019.103423
44. Serte, S., & Demirel, H. (2020). Wavelet-based deep learning for skin lesion classification. IET Image Processing, 14(4), 720-726. https://doi.org/10.1049/iet-ipr.2019.0553
45. Aydin, V. A. (2022). CNN Tabanlı Yoga Pozu Sınıflandırmasında Aktivasyon Fonksiyonu Karşılaştırması. In Proceedings of IES'22 International Engineering Symposium, Engineering Applications in Industry, 117-122.
46. https://www.kaggle.com/datasets/niharika41298/yoga-poses-dataset
47. Chen, C. H., & Ramanan, D. (2017). 3d human pose estimation= 2d pose estimation+ matching. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7035-7043).