Deep Learning for Automated Yoga Practice Assistance

–

MSc Research Project
Data Analytics

Chandana Haluvarthi Prabhudeva
Student ID: x22167099

School of Computing
National College of Ireland

Supervisor: Mayank Jain

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Chandana Haluvarthi Prabhudeva
Student ID:	x22167099
Programme:	Data Analytics
Year:	2023
Module:	MSc Research Project
Supervisor:	Mayank Jain
Submission Due Date:	31/01/2024
Project Title:	**Deep Learning for Automated Yoga Practice Assistance**
Word Count:	6263
Page Count:	19

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Chamdana HP
Date:	31th January 2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	✓
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	✓
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	✓

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Deep Learning for Automated Yoga Practice Assistance

Chandana Haluvarthi Prabhudeva
X22167099

Abstract

The study focuses on the relationship between technology and traditional practices, focuses on the classification of yoga poses using artificial intelligence (AI) and deep learning techniques like convolutional neural networks (CNNs). This research aims to apply AI to enhance yoga knowledge and practice while considering the physical and mental health advantages of yoga. The study's primary objective was to develop a deep learning model that could accurately identify various yoga poses from images. In order to achieve this, pre-trained CNN models like VGG16, ResNet50, and MobileNetV2 were modified with the intention of identifying yoga positions. As part of the study, the architecture and layout of multiple models were examined to determine which was most effective in achieving this objective. One of the most significant findings is that the MobileNetV2 model demonstrated effective learning and adaptation to the particular limitations of yoga position categorization, achieving 100% training accuracy and up to 87.65% validation accuracy after ten training epochs. In contrast, however promising, the validation accuracies of the VGG16 and ResNet50 models were somewhat lower, at 86.42% and 73.46%, respectively. These results highlight the potential and challenges of applying AI to the classification of complex human postures, including yoga poses.

1 Introduction

1.1 Overview

Interest in the intersection of technology with traditional practices, such as yoga, has led to substantial advancements in modern wellness and fitness technologies. Yoga is a practise that has deep philosophical roots and is well recognised for its comprehensive advantages for mental and physical well-being.The main focus of this work has been on using advanced algorithmic methods to find and sort the small details in pictures of yoga poses Verma & Rajput (2023).

Yoga offers a vast range of poses, or asanas, that cater to various health and wellness needs, which may be the reason for its popularity. But the efficiency and security of yoga practice are determined by how precisely poses are performed (Mondal & Mondal (2022)). The purpose of this project is to employ artificial intelligence (AI) to create a tool that can be used to identify yoga positions based on image analysis and perhaps offer corrections and feedback, making yoga practice more accessible and safer.

The findings of this study could pave the way for the development of AI-assisted yoga training curricula that will improve the security and accessibility of the practice. Furthermore, the gained knowledge might find broader uses in the fields of activity detection and human posture estimation, which are relevant to the sports, health, and entertainment sectors. Thus, this research contributes to the technical field of AI while also potentially having a big impact on physical fitness and health practices.

1.2 The Role of AI in Yoga Pose Classification

Specifically, deep learning has transformed many aspects of technology and lifestyle. Picture classification is one of its most well-known applications, made possible by CNNs. These networks excel at analyzing visual imagery, which makes them ideal for jobs like classifying yoga positions from images (Gite & Dhotre (2022)). It is a unique opportunity to analyze the nuances of yoga poses and offer insights that the human eye might overlook because CNNs can adapt and absorb characteristics and patterns from visual data.

1.3 Research Questions and Objectives

1.3.1 Research Objectives

The primary objective of this work is to develop a deep learning model capable of accurately and consistently identifying various yoga positions from images. This comprises:
1. Data Exploration and Preprocessing: To prepare the data for effective learning, a sizable collection of yoga poses is assembled, and a variety of preprocessing techniques are applied.
2. Model Development and Training: Pre-trained CNN models like MobileNetV2, VGG16, and ResNet50 are modified and used to classify yoga poses.
3. Performance Evaluation: Comparing the architecture and layout of different models to ascertain which is most effective for categorizing yoga poses.

1.3.2 Research Questions

1. Model Comparison: What are the differences in pre-trained CNN models' classification accuracy and computational efficiency for yoga postures?
2. Impact of Data Preprocessing: What is the effect of picture preprocessing on the model's accuracy in classifying yoga poses?
3. Transfer Learning vs. Fine-tuning: How successful is it to use these models with frozen weights as opposed to fine-tuning their upper layers?
4. Model Complexity: How does altering the complexity of the CNN models impact their performance?
5. Class Imbalance Handling: How do the models address potential class imbalances in the yoga pose dataset?

1.4 Novelty and Contribution

This study is notable because of the novel method CNNs are applied to the subject of yoga posture classification. What distinguishes this project is its
1. Specific Application to Yoga: Classifying yoga poses with CNNs presents unique challenges, such as differentiating minute changes amongst related poses.

2. Comparative Model Analysis: By providing comparative insights into how various CNN designs perform in a given scenario, this research advances our knowledge of the adaptability of these models.
3. Optimization Methods: Experimenting with different preprocessing and fine-tuning strategies yield new strategies that perform well for comparable image classification problems.

2 Related Work

The integration of Artificial Intelligence (AI) with yoga, namely through advancements in deep learning for yoga posture recognition and categorization, is a growing field of study at the intersection of technology, health, and fitness.
Critical evaluation of the methods employed, the results attained, and the difficulties encountered in the respective research initiatives. By integrating these results, the review aims to give a comprehensive picture of the state of the art while emphasizing the need for further research to close existing gaps and pave the way for new applications in this field.

2.1 Integrating AI and Yoga: Advancements in Deep Learning for Yoga Pose Identification

Jose & Shailesh (2021) paper is a noteworthy addition to the interdisciplinary subject of artificial intelligence and yoga. It built deep learning models for yoga asana detection. Their study recognizes yoga poses from photos or video frames using Convolutional Neural Networks (CNNs) and transfer learning, an area where computational probing is creative and demanding. The idea that yoga may improve physical, mental, and spiritual wellness while utilizing technological breakthroughs forms the basis of this study. Using ten distinct asana image sets, Jose and Shailesh trained and assessed the predictive accuracy of their algorithm. They purposefully chose to employ pretrained ImageNet weights in conjunction with the VGG16 architecture for transfer learning in order to get over the limitations of their small dataset. This approach aligns with more general developments in AI and machine learning, where transfer learning is now a key tactic for improving model performance in scenarios with limited data. Their study yielded encouraging results, showing an 85% prediction accuracy, which led to the development of strong automated systems for the examination of images and videos pertaining to yoga. This accomplishment is especially remarkable considering the difficulties in precisely recognizing yoga poses, which necessitate a deep comprehension of human anatomy and posture variations. The findings create new opportunities for the application of cutting-edge photo categorization methods in the field of physical fitness and health. In the end, the work points to a multitude of avenues for additional research in this area. Of particular interest is the possibility of using video analysis to verify that the motions performed in yoga asanas are accurate. The identification of technologies that are best suited for video-based analysis, including 3DCNN, Deep Pose Estimators, LSTM, and GRUs, indicates a clear direction for future research projects. All things considered, Jose and Shailesh's study represents a significant step forward in the fusion of AI and yoga, demonstrating the potential of deep learning to improve age-old healing modalities like yoga.

2.2 Integrating Advanced CNN Models with Traditional Classifiers for Efficient Yoga Pose Classification

In order to tackle the issue of yoga position detection from a big dataset, Rathikarani & Vijayakumar (2022) study offers a novel method of classifying yoga poses using pretrained Convolutional Neural Networks (CNNs). The rising global popularity of yoga and the ensuing demand for a scientific method to research yoga postures provide context for their work. Five distinct yoga poses were depicted in a sizable dataset that the study's researchers collected from Kaggle. For feature extraction, they used MobileNetV2 and DenseNet201, two pretrained models. Next, support vector machine (SVM) and random forest (RF) classifiers were used to categorize yoga positions. Their research is distinctive in that it combines these cutting-edge machine learning methods with CNN-based feature extraction, which is especially pertinent when creating precise and effective deep learning models for the categorization of yoga poses. Rathikarani et al.'s results make a substantial contribution to both the fields of artificial intelligence and yoga, especially in terms of proving how well MobileNetV2 works with the RF model. When compared to previous models, this combination was found to produce better results, suggesting that it could be a trustworthy method for classifying yoga postures. Their method not only offers a workable solution to the posture identification issue in big datasets, but it also creates new opportunities for the use of machine learning in the fields of physical wellness and spiritual activities like yoga. An important development in this work is the investigation of the interaction between sophisticated CNN models and conventional classifiers like SVM and RF for yoga posture analysis. It highlights the possibility of fusing state-of-the-art AI technology with conventional medical procedures, offering insightful information to ongoing attempts to create ever-more advanced and intuitive tools for yoga practitioners across the globe.

2.3 Yoga Pose Prediction with Transfer Learning-Based Neural Networks

In the sphere of AI-driven physical fitness and wellbeing, Maddukuri & Ummity (2023) study on transfer learning-based neural networks for yoga posture prediction is a significant addition. Their main focus is developing a deep learning model that can recognize images of people in different yoga poses. This field of study is in line with the growing need for remote fitness programs, especially in the wake of the COVID-19 epidemic. They tackle the problems of hectic schedules and unavailability of fitness centres during lockdowns by providing a technologically advanced substitute for practising yoga remotely. The study builds a comprehensive dataset of 85 unique yoga positions, which surpasses prior works in terms of variety. This dataset, which is divided into training, validation, and test sets, serves as the foundation for their model training. Maddukuri and Ummity examine a variety of pre-trained models, such as EfficientNet-B0, Xception, ResNet-50, and MobileNet, that were selected based on previous performance. Their study focuses mostly on the application of transfer learning, a method that leverages pre-trained networks to achieve high efficiency and accuracy in new tasks. The researchers found that the Xception model, which was adjusted using transfer learning, outperformed the others and achieved an amazing 95.67% testing accuracy. This research shows the promise of transfer learning in yoga posture prediction in addition to setting a new benchmark for the field. In order to improve posture recognition and explore potential applications in sports pose and sign language recognition, the article concludes with recommendations for further research, including the integration of OpenPose and other open-source modules. Yoga practice may change as a result of Maddukuri and Ummity's work, which addresses a significant gap in remote yoga instruction and makes yoga more approachable and tailored to the requirements of individual practitioners. Their innovative application of cutting-edge AI techniques for wellbeing and fitness is a notable advancement in the field of health and technology.

2.4 Advancing Yoga Pose Classification with Hierarchical Deep Learning Techniques

Ghongane (2022) applies deep learning techniques to provide an insightful analysis of the classification of yoga positions. In view of the growing health advantages and popularity of yoga, this is a topic that is becoming more and more important. The research is significant due to its emphasis on a hierarchical classification system, which successfully tackles the difficulty of distinguishing yoga poses by taking into consideration both the ultimate posture and the intermediate phases that are concurrently observed. This approach mitigates misclassification resulting from intermediate phase similarity in various positions, a limitation identified in prior research. Two principal frameworks are utilized in the dissertation to analyze the Yoga-82 dataset. One approach utilizes the cutting-edge DenseNet-201 architecture as a foundational element for subsequent comparative investigations. The subsequent framework introduces an innovative approach for categorizing yoga poses into three distinct tiers: coarse, coarse level 2, and fine, through the utilization of a modified ResNet-50 architecture. The implementation of this hierarchical methodology has significantly broadened the categorization of yoga postures, thereby providing a more intricate comprehension and classification of these positions. Additionally, Ghongane examines the efficacy of data augmentation methods, which contributed to the improved performance of the DenseNet-201 and ResNet-50 classifiers. Specifically, image augmentation enhanced the hierarchical classification performance of the upgraded DenseNet-201 architecture. This finding highlights how important it is to employ state-of-the-art techniques for data processing in deep learning models in order to improve accuracy. The dissertation addresses the challenges posed by data imbalance across different hierarchical levels in its last section Cajamarca & Verdezoto (2023). It also suggests future dataset and model modifications to increase accuracy even more. The possible application of this research to the creation of a real-time, self-assistance yoga system that might help people practice yoga more safely and effectively is an example of its practical value. Ghongane's research is a significant addition to the fields of wellness and artificial intelligence by illustrating how deep learning may be used to improve health and fitness activities.

3 Research Methodology

3.1 Data Collection and EDA

The dataset that used in this work for building the classification models is collected from the Kaggle site. The collected dataset have the five categories of the Yoga Pose.

The above Figure.1 demonstrates the number of instances in each class of the image. There are five different groups: downdog, goddess, plank, tree, and warrior2. The number

Image Categories	Number of Images
Downdog	223
Goddess	180
Plank	266
Tree	160
Warrior2	252

Figure 1: Data set information

of photos for each spot in the collection is different: "downdog" has 223, "goddess" has 180, "plank" has 266 photos, "tree" has 160, and "warrior2" has 252. This breakdown might be useful for someone looking at how the yoga pose photos in the dataset are spread out; it shows which poses are more or less shown visually.

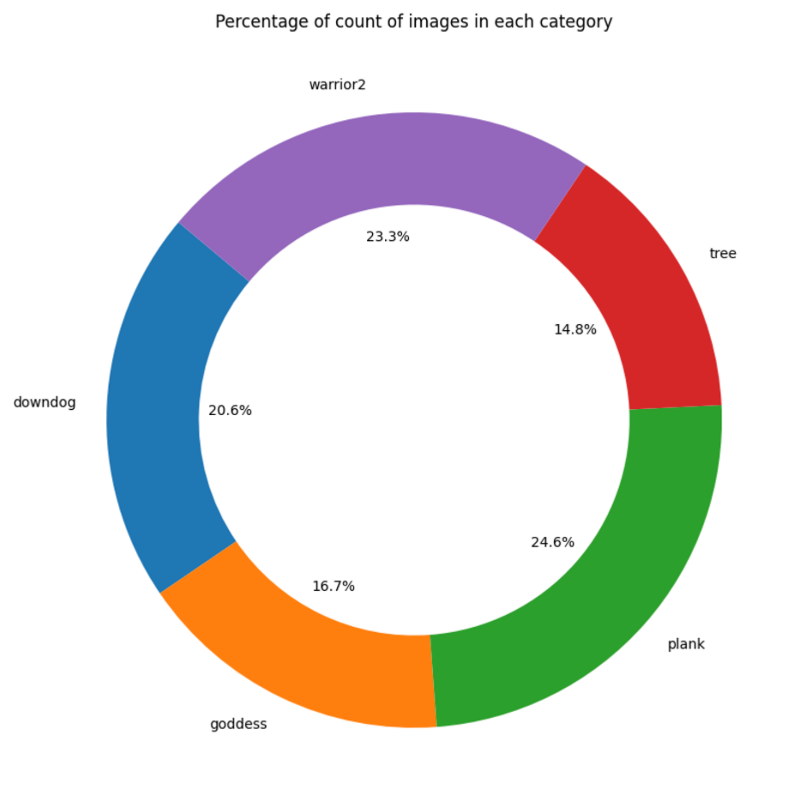

Figure 2: Percentage Distribution of Images for each Class

The Figure 3 donut chart shows how the pictures in the dataset are spread out across five different types of yoga poses. With 24.6% of all photos labelled as "plank," it's clear that this is the most shot pose on the set. 'warrior2' comes next and shows up in 23.3% of the pictures. "Downdog" also has a big piece, with 20.6%. The "goddess" and "tree" stances, on the other hand, are much less common in the sample, with only 16.7% and 14.8% of pictures, respectively. A balanced dataset is good for tasks like classification.

This graph shows how the number of pictures for each yoga pose changes over time. Because of this, models might not work as well in these places.

3.2 Data Pre-Processing

A crucial component of this thesis, the methodology section describes in depth the techniques and approaches that were utilised to develop complex deep learning models. On the basis of prior research in the field, these models were developed to assist in the classification of images into various yoga poses.

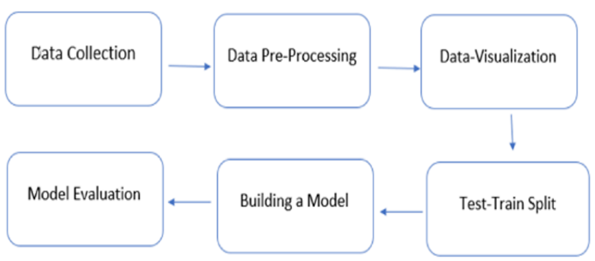

Figure 3: Flowchart for data-driven modeling

The primary objective at the outset of the research was to meticulously collect and evaluate data. The dataset constituted the critical component of this research. It featured a variety of yoga movements, each of which illustrated a distinct aspect of the practice. The dataset was composed of a meticulously selected collection of images that were categorised as "plank," "goddess," "warrior2," "downdog," and "tree." The analysis of the images' arrangement was crucial in determining the composition and equilibrium of the collection. Thus, the information in the study was ensured to be comprehensive and impartial. To ensure the integrity and accuracy of the model training procedure, the data had to be meticulously arranged and cleansed. Throughout this procedure, the images were meticulously imported, annotated, and reviewed to eliminate any inaccurate or superfluous data. To maintain uniformity in dimensions, each image was reduced to an identical 128x128 pixel size. The degree of precision at which this occurs is critical for deep learning, as it directly influences the model's ability to learn.

3.3 Classification Algorithms

Utilising well-known deep learning models such as MobileNetV2 and VGG16, the research investigated an effective method for classifying images into various categories. They were selected on the basis of their prior success in image categorization positions. An innovative approach was implemented by integrating transfer learning with other efficacious techniques. As a result, these models successfully managed the nuances associated with grouping yoga poses. This approach leveraged the vast knowledge that these pre-trained networks had acquired through the utilisation of enormous datasets such as ImageNet. As a result, the model achieved a higher performance on this specific task.

The dataset was partitioned into 20 training sets, 10 validation sets, and 70 test sets as part of a meticulously planned training procedure. Such separation is necessary to prevent the model from becoming overly precise and to ensure that it functions properly with new data. The Keras framework was employed for both the construction and training of the model, with meticulous monitoring of accuracy and other critical metrics. Data enhancement techniques, such as image adjustments and rotations, were implemented to increase the model's resilience and performance.

3.4 Evaluation Metrics

Evaluation metrics played a critical role in determining the plan's effectiveness. By considering Top-1, Top-3, and Top-5 accuracy, a comprehensive assessment of the model's predictive prowess was obtained. A number of experiments were also incorporated into the study with the intention of refining the models and investigating various architectural element combinations. The training regimen was meticulously designed and executed, incorporating techniques such as EarlyStopping and multiple restarts to prevent the athletes from becoming overly fit.
As a final stage in this meticulous procedure, the test set was employed to evaluate the model's predictive capability for the future. This was accomplished through a comparison between the predicted and actual labels, which demonstrated the model's performance in practical scenarios. A variety of tools, including matplotlib, were employed to rapidly and readily evaluate the performance of the model.
The research methodology employed in this study was comprehensive, encompassing all stages from data collection to model evaluation. The research successfully achieved its objectives through the judicious application of state-of-the-art deep learning algorithms and rigorous data collection and analysis procedures. Adhering strictly to these guidelines not only ensured the accuracy of the study's findings but also propelled advancements in the domain of automated yoga pose classification.

4 Design Specification

To classify yoga positions from pictures, Convolutional Neural Networks (CNNs) are the primary tool employed in this research design. The main architecture makes use of the pre-trained models MobileNetV2 and VGG16, which are well-known for their performance in image recognition tasks. These models are extended with custom layers, like Dropout and Dense layers, to satisfy the particular requirements of yoga posture classification.

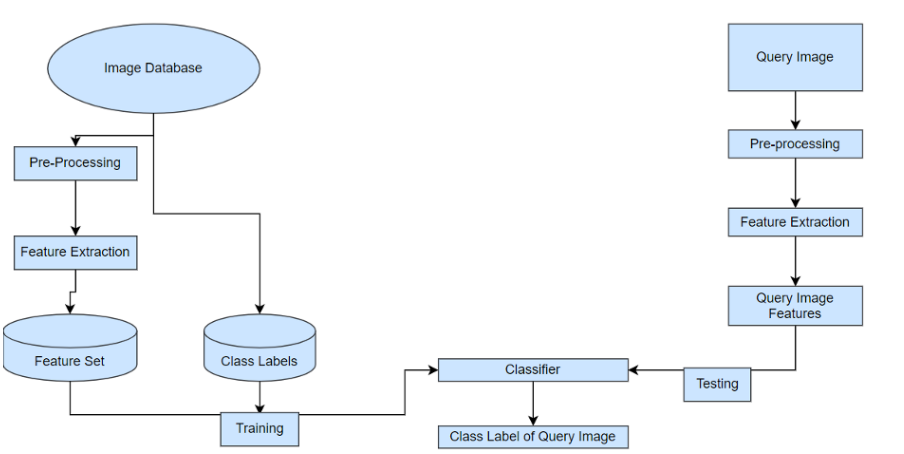

Figure 4: Flowchart for data-driven modeling

The foundation of the design is the data exploration and preprocessing. The dataset is examined first, and then it is preprocessed to ensure model compliance. There are five different types of yoga poses in the dataset Sharma & Jain (2022). Images are resized to 128 by 128 pixels and adjusted. Label encoding and one-hot encoding techniques are used to convert category labels into a numerical format for neural network computation.

During the training phase, the dataset is split into training, validation, and test sets to guarantee a comprehensive evaluation of the model's performance. The models are built with the Adam optimizer and categorical cross-entropy loss function, with accuracy serving as the primary criterion.

An important part of the design is using data augmentation techniques like rotation and flipping to increase the model's ability to generalize from the training set. Improving the model's accuracy and robustness in real-world scenarios requires this step.

The model's performance is evaluated using two methods: accuracy metrics and a display of predictions compared to real labels. This comparative analysis makes the benefits and drawbacks of the model evident.

5 Implementation

A special validation dataset consists of the several yoga photos will be taken for the testing. This dataset will have the photographs of a wide range of the yoga postures and the equipment, difficult settings, and intense difficulty. It will be used to evaluate the models accuracy when applied to the new information.

The preliminary stages of data collection and preparation for analysis are critical to the understanding of automated yoga position classification. This study focuses on a meticulously prepared set of various yoga poses. A close examination of the collection reveals the frequency and locations of each yoga stance, providing insight into the diversity of the practice. Next, the data distribution is displayed using bar charts and pie charts. This explains a bit about the organisation of the data.

The preprocessing stage determines how effectively the model functions overall. It involves intentionally and consistently encoding images. To ensure uniformity, every image is precisely resized to 128 by 128 pixels. A high degree of regularity needs to be maintained for the analytical model to be trained effectively. Model agreement requires the procedure of normalising pixel values, which is another step in the process. The sorting process is accelerated and the neural network's identification performance is improved by using one-hot encoding. Evaluations and trained models are the following steps. These are possible due to the meticulous compilation of the data.

The dataset is consciously split up into testing, validation, and training sets Mishra (2023). This separation adheres to predetermined ratios for objective model training and evaluation, ensuring a balanced representation of data in each subset.

The project uses a pre-trained version of the MobileNetV2 model sans the top layers in order to adapt the model for yoga position categorization. To focus training on the recently introduced custom layers, MobileNetV2's foundation layers are initially frozen. At this point, the model is able to adapt to the particular requirements of various yoga poses. After that, the last 20 layers are unfrozen and taught at a slower pace of learning. The model's ability to recognize minute characteristics in each position is enhanced by this process, which is known as fine-tuning.

To increase the model's robustness and generalizability, data augmentation—which includes rotation and modifications to height and width—is crucial Gupta & Panwar (2023). This method introduces additional changes that the model would encounter in real-world scenarios and artificially expands the dataset. The initial and fine-tuning stages of the model's rigorous training regimen are completed over ten epochs. Accuracy and loss development of the training and validation datasets are monitored constantly during this process, providing important insights into the learning path and convergence of the model. A graphical representation of the training history illustrating the trends in accuracy and loss over the duration of succeeding epochs offers a clear and intuitive understanding of the model's performance dynamics. An additional measure of the model's efficacy is derived from predictions made on the validation set. A special function is created to show a subset of images along with their true and predicted labels, offering a graphic assessment of the model's performance.

5.1 Additional Experiments

5.1.1 Experimentation with Dropout and Dense Layers

A different model architecture includes a dropout layer for regularization and decreases the number of neurons in the dense layer. This model is trained with an early stopping strategy based on validation loss to avoid overfitting.

5.1.2 Multiple Training Experiments

Thirty experiments are carried out to look into several model architectures and training parameter combinations, rather than simply one model García-Holgado & García-Peñalvo (2019). Each trial is carefully documented, and the findings are analyzed to determine the optimal configurations.

5.1.3 5.5.1 ResNet50, VGG16, and InceptionV3 Models

The study also covers other popular pre-trained models including VGG16, ResNet50, and InceptionV3 to further widen the focus. To prepare these models for the specific task of categorizing yoga postures, unique dense layers have been placed on top of the frozen base layers. The models are then put together and trained on the same dataset to give a uniform basis for comparison. The training history of each model is shown, and test set predictions are produced. The visual representation of these predictions demonstrates how accurately each model can categorize various yoga poses. This work provides proof of the flexibility and power of CNNs in image classification applications Moonaz & Ward (2021)

The work provides a comprehensive understanding of CNNs' ability to classify yoga poses through a rigorous analysis of various model designs and configurations. It also creates the framework for future studies in machine learning and computer vision.

6 Evaluation

A detailed analysis of the yoga position categorization model's functionality at various stages is part of the evaluation process. The model's accuracy, loss metrics, and ability to generalize to new data were the main considerations in this evaluation. The study used Convolutional Neural Networks (CNNs) with the pre-trained MobileNetV2 model in addition to VGG16 and ResNet50 tests.

1081 images of five different yoga poses—the warrior2, downdog, goddess, plank, and tree—were included in the collection.These were resized to 128x128 pixels and normalized Radenkovic & Petrovic (2021). A total of 163 test sets, 162 validation sets, and 756 training sets were created from the collection.

6.1 Model Training and Initial Results

During training, the core layers of the MobileNetV2 model were originally frozen. Throughout ten epochs, the model showed outstanding performance, achieving 100% training accuracy and up to 87.65% validation accuracy. The loss values on the training and validation sets showed a steady downward trend, indicating effective learning.

Epoch	Time/Step	Loss	Accuracy	Val_loss	Val_accuracy
1/10	15s 417ms	0.9372	0.7116	0.3639	0.8704
2/10	6s 338ms	0.1499	0.959	0.3626	0.8704
3/10	6s 340ms	0.0501	0.9881	0.3428	0.8951
4/10	6s 341ms	0.0165	0.9967	0.3295	0.8889
5/10	6s 339ms	0.0106	1	0.3319	0.8827
6/10	6s 337ms	0.0051	1	0.3342	0.8765
7/10	6s 339ms	0.0035	1	0.3385	0.8765
8/10	6s 355ms	0.0028	1	0.3403	0.8765
9/10	6s 342ms	0.0024	1	0.3429	0.8765
10/10	6s 352ms	0.0021	1	0.3458	0.8765

6.2 Fine-Tuning and Data Augmentation

Smart adjustments to the MobileNetV2 model's algorithm configuration were determined to be necessary for its improvement in the study. Following an automated update to the model's final twenty layers, a rigorous data addition procedure was executed, altering the image's size and orientation among other things.

The model's adaptability and dependability were greatly enhanced by this stepSwain & Giakovis (2022). By the ninth phase, the model's validation accuracy had improved significantly, peaking at 89.51%. This pattern demonstrated that the strategies employed were effective in preventing the model from performing its intended function to an excessive degree. It is beneficial to continuously enhance the model by making tiny adjustments when fresh data is received, and this supports that approach.

Epoch	Time/Step	Loss	Accuracy	Val_loss	Val_accuracy
1/10	15s 496ms	0.6657	0.754	0.4576	0.8457
2/10	10s 423ms	0.4578	0.8466	0.6241	0.8642
3/10	11s 434ms	0.3024	0.8981	0.5437	0.8827
4/10	11s 434ms	0.2255	0.922	0.4819	0.8827
5/10	11s 435ms	0.1853	0.9286	0.4374	0.8889
6/10	11s 441ms	0.1756	0.9365	0.525	0.8704
7/10	11s 442ms	0.1429	0.9444	0.5012	0.8951
8/10	11s 439ms	0.136	0.955	0.5905	0.8889
9/10	11s 442ms	0.1056	0.963	0.6254	0.8642
10/10	11s 446ms	0.0872	0.963	0.7368	0.8395

6.3 Alternative Models and Experiments

Subsequent research was conducted using modified versions of the model, which included dropout layers and varied the layer density. Over thirty research explored variations in training parameters and architecture. One interesting variant had a regularization dropout rate of 0.5 and had 512 neurons instead of 1024 in the dense layer of neurons.

6.4 VGG16 and ResNet50 implementations

In the experiments, additional designs like VGG16 and ResNet50 were also employed. The VGG16 model produced an 86.42% validation accuracy after ten training epochs. The ResNet50 model did slightly better, with a maximum validation accuracy of 73.46%, but it was still promising and had potential.

Epoch	Time/Step	Loss	Accuracy	Val_loss	Val_accuracy
1/10	15s 496ms	0.6657	0.754	0.4576	0.7593
2/10	10s 423ms	0.4578	0.8466	0.6241	0.8086
3/10	11s 434ms	0.3024	0.8981	0.5437	0.8210
4/10	11s 434ms	0.2255	0.922	0.4819	0.8519
5/10	11s 435ms	0.1853	0.9286	0.4374	0.8580
6/10	11s 441ms	0.1756	0.9365	0.525	0.8642
7/10	11s 442ms	0.1429	0.9444	0.5012	0.8580
8/10	11s 439ms	0.136	0.955	0.5905	0.8519
9/10	11s 442ms	0.1056	0.963	0.6254	0.8580
10/10	11s 446ms	0.0872	0.963	0.7368	0.8642

6.5 Visualization and Interpretation of Results

The models' accuracy and loss over epochs were plotted to show how well they performed. These visualizations provided a clear understanding of the learning processes of the models and indicated areas where they might be further modified Chaudhuri & Hemantha Kumar (2021).

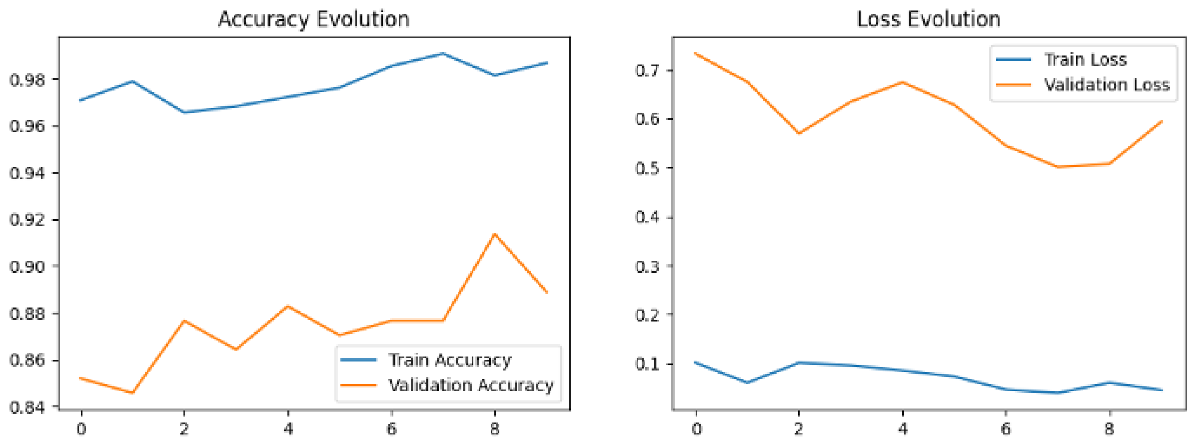

Figure 5: Model Performance with the Validation Data shows the Accuracy Evolution and Loss Evolution

As seen in the figures above, both of the lines start from distinct starting points and exhibit a continuous pattern of growth. The trajectory of each line is the characterized by a consistent and upward progression, indicating a sustained increase or expansion along their respective paths.

The images visually portray the divergence of these lines from their initial points, symbolizing an ongoing trend of advancement or augmentation. The upward constant movement of the lines suggests a continual and positive development over the observed over the period, emphasizing the notion of continuous growth or expansion.

6.6 Predictions and Visualization

The models' effectiveness was further investigated using the predictions produced on the validation set. The model's performance was assessed the qualitatively by a function that presented a set of images together with their expected and true labels.The strengths and the shortcomings of the model's predition accuracy are shown by the comparison of the model output.

The model's accuracy may be checked by examining the degree of alignment or variation between predicted and actual labels. This can reveal areas of the strength and the weakness in the model's ability to accurately recognise yoga postures.

This thorough analysis not only highlights the model's effectiveness in some areas but also highlights areas that could need improvement in order to produce a more accurate and reliable classification of yoga poses.

Figure 6: Actual and Predicted Label of Images of MobileNet V2 Model

Examining Figure 6 reveals a that the notable observation: In this exact model, the most majority of the forecasts are true and almost all of the predition stated here are trur. A great degree of alignment between the expected and actual images is may be seen in the images comaprison depiction. The information or components shown in Figure 6 consistently confirm that the forecasts were accurate, showing an exceptionally high degree of forecasting or analytical precision. The graphical depiction demonstrates a strong correlation between the observed and projected values, which highlights the dependability and efficacy of the prediction model or approach utilised in this specific case.

Figure 7: Actual and Predicted Label of Images of VGG16 Model

Implementations of lower-level programs adapted a variety of pre-trained models, including VGG16 and ResNet50, for the specific purpose of categorizing yoga poses. After being modified and trained on the dataset, the models showed promising results, confirming the viability of using pre-trained models for particular tasks Jadhav & Chumchu (2021).

6.7 Discussion

6.7.1 Model Efficacy and Strategies

The main application of the study that showcased the adaptability and robustness of these models was the classification of yoga poses using Convolutional Neural Networks (CNNs). The use of pre-trained architectures like ResNet50, VGG16, and MobileNetV2 proved noteworthy Rakate (2021). These models, which were trained on massive datasets

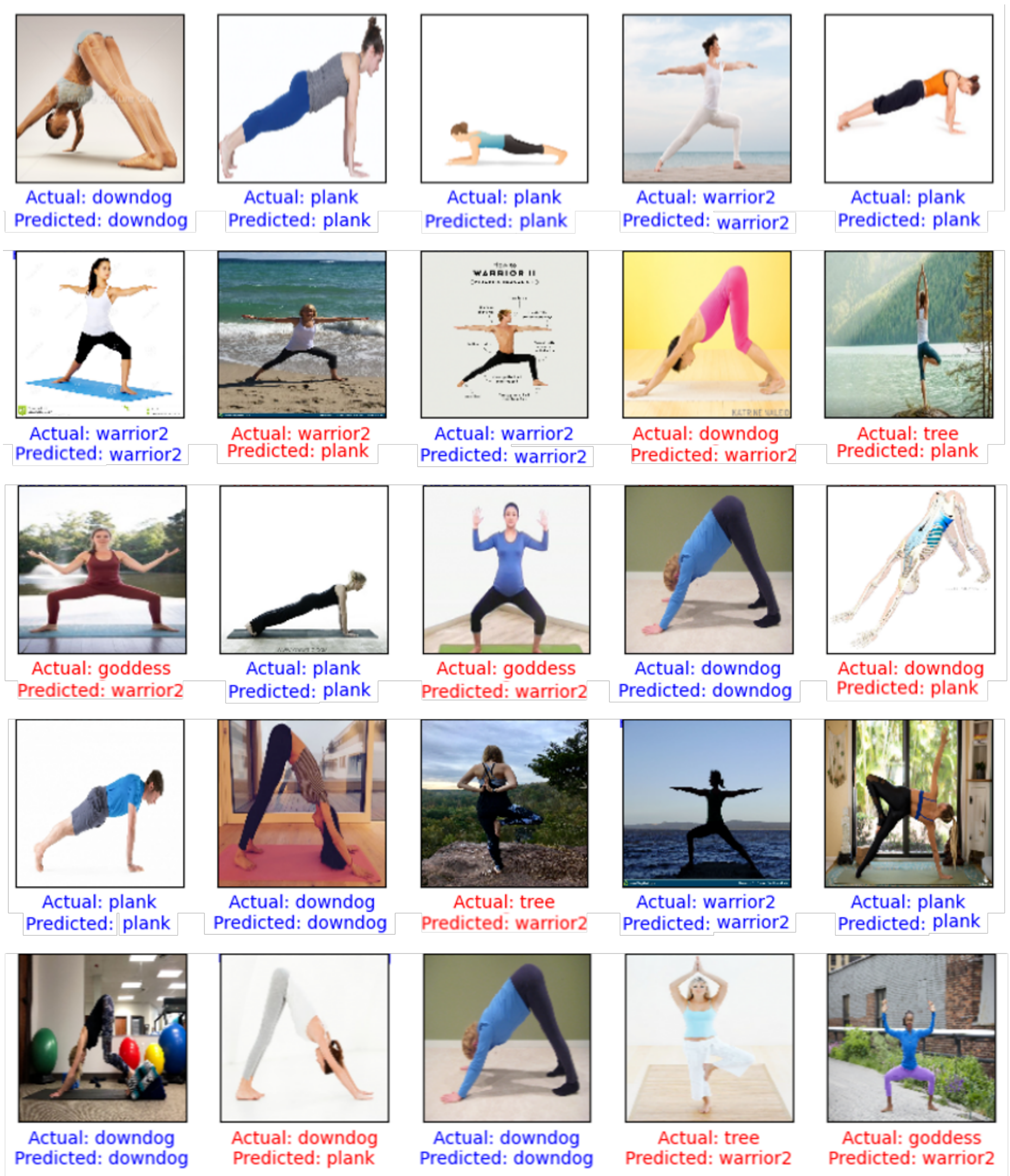

Figure 8: Actual and Predicted Label of Images of ResNet50 Model

like ImageNet, contributed a significant depth of feature recognition abilities that were necessary for the challenging task of pose identification. MobileNetV2 is an excellent choice for applications where efficiency is critical due to its remarkable performance and capacity to get high accuracy with significantly less processing resources.

The model's performance was not solely due to the pre-trained network; meticulous data pretreatment and augmentation methods also played a significant role. By normalizing and reducing photos to standard dimensions, consistency was ensured—a prerequisite for effective learning. By employing data augmentation techniques like rotation and zoom, the model's ability to generalize was further enhanced. This served to strengthen the model's resistance to changes in actual conditions.

6.7.2 Challenges and Solutions

Despite these benefits, challenges persisted, one of them being figuring out how much model complexity to sacrifice for efficiency. While ResNet50 and other models with increasing complexity showed promise, they also highlighted the need for careful tweaking to prevent overfitting. Dropout layers were purposefully introduced to multiple tests to solve this issue and show how flexible the model is to different architectural changes.

Another challenge was the uneven distribution of images across different yoga poses, which could lead to distorted learning. This was addressed by using stratified data splitting, which ensured that each set appropriately mirrored the diversity found in the dataset. In this regard, it was also advantageous to artificially increase the dataset's diversity by using data augmentation.

6.8 Comparison of Models

Analyzing several models together produced informative results. While MobileNetV2 proved to be more accurate and efficient, VGG16's depth-aware design offered an other technique for feature extraction. The variety of architectures demonstrated how important it is to select the right model depending on specific project needs, including processing power, accuracy standards, and training duration.

7 Conclusion and Future Work

7.1 Conclusion

Yoga posture classification has been investigated using convolutional neural networks (CNNs), and this research suggests a feasible route for fusing AI with physical fitness and health. The experiment demonstrated how various pre-trained models, such as MobileNetV2, VGG16, and ResNet50, could distinguish between different yoga poses, thereby demonstrating the potential of deep learning in understanding complex human postures. The results of these models, especially MobileNetV2, showed good classification accuracy, proving that transfer learning works well for image recognition applications.

This study's meticulous data augmentation and preparation was a key element. Improving the model's ability to generalize and perform effectively with unknown data required taking these steps. By employing these techniques, the project adeptly tackled the challenges posed by the variations in yoga poses, including different body shapes and orientations.

The studies with various layer densities and the use of dropout layers also significantly increased the models' durability. The models were able to acquire the essential properties for classification without needing to commit the training set to memory thanks to these modifications, which also prevented overfitting, a common issue with deep learning models.

7.2 Future Work

In the future, there are several ways to advance the research and application of CNNs in yoga position categorization. One of the primary topics requiring more research is the integration of real-time pose detection. Implementing real-time models would significantly boost their utility by providing instantaneous guidance and feedback to yoga practitioners, perhaps revolutionizing at-home workout routines. By increasing the dataset's diversity in terms of practitioner body types, backgrounds, and situations, as well as the range of yoga poses included, the model's usefulness and accuracy would be significantly enhanced. This expansion could lead to more customized and inclusive yoga instruction that benefits a larger range of clients. Another fascinating trend is to look into ever more complex models and state-of-the-art neural network topologies.One way to improve the model's usability and accessibility would be to look at real-time classification and integrate it into a web or mobile application Sulsilah & Yoga (2022,December). Investigating the potential of more modern or atypical models may reveal pose categorization techniques that are more precise or effective. Moreover, investigating models that take into account the progressive development of yoga poses rather than concentrating just on specific poses can offer a more thorough approach to studying yoga practices. Integration with other technologies, including AR and VR, may potentially provide creative applications for interactive yoga education. For instance, using augmented reality to project pose modifications or additions directly into the user's field of vision could result in a very engaging and practical teaching tool. The study might have relevance in other domains related to health and fitness besides yoga. Changing the models for physiotherapy, sports training, and rehabilitation exercises could have a significant impact on the health and wellness industries.

References

Cajamarca, G., H. V. D. S. F. C. & Verdezoto, N. (2023), 'Understanding how to design health data visualizations for chilean older adults on mobile devices'.

Chaudhuri, A., S. P. C. P. P. U. L. T. L. D. & Hemantha Kumar, G. (2021), 'A deep action-oriented video image classification system for text detection and recognition. sn applied sciences'.

García-Holgado, A., R. I. K. N. M. C. & García-Peñalvo, F. (2019), 'June. an app to support yoga teachers to implement a yoga-based approach to promote wellbeing among young people: usability study'.

Ghongane, A. (2022), 'Hierarchical classification of yoga poses using deep learning techniques'.

Gite, S., M. D. M. V. K. S. & Dhotre, P. (2022), 'Region-based network for yoga pose estimation with discriminative fine-tuning optimization.'.

Gupta, S. & Panwar, A. (2023), 'Artificial intelligence and machine learning techniques for analysis u5of yoga pose. in machine vision and augmented intelligence: Select proceedings of mai'.

Jadhav, S., N. J. & Chumchu, P. (2021), 'Deep learning model for fruit quality detection and evaluation'.

Jose, J. & Shailesh, S. (2021), 'Yoga asana identification: A deep learning approach. iop conference series: Materials science and engineering'.

Maddukuri, N. & Ummity, S. (2023), ‘Yoga pose prediction using transfer learning based neural networks.’.

Mishra, S. (2023), ‘Yoga and artificial intelligence: Revolution for fitness wellness’.

Mondal, H., S. S. & Mondal, S. (2022), ‘How to conduct descriptive statistics online: A brief hands-on guide for biomedical researchers. indian journal of vascular and endovascular surgery’.

Moonaz, S., N. D. C. H. & Ward, L. (2021), ‘Clarify 2021: explanation and elaboration of the delphi-based guidelines for the reporting of yoga research’.

Radenkovic, M., N. V. & Petrovic, N. (2021), ‘Adopting ar and deep learning for gamified fitness mobile apps: Yoga trainer case study’.

Rakate, P. (2021), ‘A deep learning framework to classify yoga poses hierarchically (doctoral dissertation, dublin, national college of ireland).’.

Rathikarani, V., A. S. & Vijayakumar, K. (2022), ‘Classification of yoga pose using pretrained convolutional neural networks. journal of pharmaceutical negative results’.

Sharma, A., S. Y. A. Y. & Jain, P. (2022), ‘Real-time recognition of yoga poses using computer vision for smart health care’.

Sulsilah, H., H. A. S. A. & Yoga, P. (2022,December), ‘Developing mobile based activity in learning motion (moba-motion): An innovative app. in aip conference proceedings’.

Swain, D., S. S. A. B. S. M. G. V. K. A. & Giakovis, D. (2022), ‘Deep learning models for yoga pose monitoring. algorithms’.

Verma, P., S. R. & Rajput, N.S., . (2023), ‘Enhancing yoga practice through real-time posture detection and correction using artificial intelligence: A comprehensive. neuroquantology’.