Prototypical Network to solve the problem of Few Shot Learning

Project report submitted in partial fulfillment
of the requirements for the degree of

Bachelor of Technology
in
Computer Science and Engineering

by

Arnav Jain - 21UCS028
Gautam Mittal - 21UCS081
Jeetaksh Gandhi - 21UCS098

Under Guidance of
Dr. Lal Upendra Pratap Singh

LNMIIT
The LNM Institute of
Information Technology

Department of Computer Science and Engineering
The LNM Institute of Information Technology, Jaipur

November 2024

The LNM Institute of Information Technology
Jaipur, India

CERTIFICATE

This is to certify that the project entitled “Few Shot Learning” , submitted by Arnav Jain (21UCS028), Gautam Mittal (21UCS081) and Jeetaksh Gandhi (21UCS098) in partial fulfillment of the requirement of degree in Bachelor of Technology (B. Tech), is a bonafide record of work carried out by them at the Department of Computer Science and Engineering, The LNM Institute of Information Technology, Jaipur, (Rajasthan) India, during the academic session 2024-2025 under my supervision and guidance and the same has not been submitted elsewhere for award of any other degree. In my/our opinion, this report is of standard required for the award of the degree of Bachelor of Technology (B. Tech).

Date

Adviser: Dr. Lal Upendra Pratap Singh

Acknowledgments

It gives us immense pleasure to express our heartfelt gratitude to Dr. Lal Upendra Pratap Singh, whose guidance, encouragement, and unwavering support were vital throughout this project. His profound knowledge and constructive feedback at every stage of the research significantly contributed to the depth and quality of this work. We are deeply indebted to him for his mentorship, which provided the clarity and direction to achieve the project objectives.

We are equally grateful to our friends, Harsh Aggarwal and Hardik Paarakh, for their invaluable collaboration and contributions. Their innovative ideas, thoughtful discussions, and dedicated efforts significantly addressed challenges and refined the project outcomes. Their team spirit and commitment made this journey both productive and memorable.

We would also like to acknowledge the numerous researchers and authors whose scholarly work was the foundation for this project. The insights and data obtained from their research papers were instrumental in shaping the framework and enriching the content of this study. Their contributions to the field inspired and helped navigate complex areas of the subject.

Furthermore, We sincerely thank our family and peers, whose encouragement and moral support motivated us throughout this journey. Their constant belief in our abilities gave us the strength to persevere and achieve the desired goals.

Finally, We express our appreciation to all those who contributed, directly or indirectly, to the successful completion of this project.

Abstract

Few-shot learning presents a significant challenge in deep learning due to the scarcity of labeled data, which traditional supervised learning methods cannot effectively address. We propose Prototypical Networks for the problem of few-shot classification, where a classifier must generalize to new classes not seen in the training set, given only a small number of examples of each new class. Prototypical Networks have emerged as a powerful framework for few-shot learning, leveraging metric-based techniques to classify unseen categories with minimal labeled examples. This project implements a **Prototypical Neural Network (PNN)** for **few-shot classification tasks**, tailored to the Yoga-82 dataset, which contains various yoga poses. The primary objective is to design, train, and evaluate a simple yet effective model capable of classifying novel yoga poses with limited labeled data, emphasizing computational efficiency and interpretability. Prototypical Networks overcome this limitation by employing episodic training, mimicking the few-shot evaluation process during training. Each task consists of support sets for prototype computation and query sets for classification.

Prototypes, computed as the mean embeddings of the support examples for each class, serve as the reference points in the embedding space for classification. Query samples are classified based on their distances to these prototypes using metrics such as the Euclidean distance. This intuitive and interpretable approach makes Prototypical Networks well-suited for scenarios with constrained data availability. The training process employs episodic learning, the hallmark of Prototypical Networks. Each training episode dynamically samples tasks comprising N_c classes, N_s support examples per class, and N_q query examples per class. Class prototypes are calculated as the mean embeddings of the support examples for each class, while query samples are classified based on their distances to these prototypes. The training objective minimizes the negative log-likelihood of the true class labels over the softmax of the computed distances. This work paves the way for continued refinement and adaptation of Prototypical Networks to domain-specific challenges, including fine-grained classification tasks in highly similar datasets.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 The Area of Work

This project focuses on **Few-Shot Learning (FSL)** for **Yoga Posture Recognition** using computer vision techniques and the Yoga-82 dataset. The goal is to develop a model capable of accurately identifying yoga postures from limited labelled examples, addressing the challenges of data scarcity and variability inherent in human posture recognition tasks. Few-shot learning leverages meta-learning or embedding-based approaches to generalize well with minimal training samples. This capability is precious for applications like yoga posture recognition, where collecting and annotating large-scale datasets can be resource-intensive and time-consuming. With 82 classes and various yoga poses, the Yoga-82 dataset provides a thorough standard for training and assessment. In order to extract and evaluate pose-specific characteristics, the research investigates cutting-edge computer vision frameworks, such as transformer-based architectures and convolutional neural networks (CNNs). To improve recognition performance, it also uses pose estimation algorithms.

Important goals consist of:

- Applying methods for Few-Shot Learning to the Yoga-82 dataset.
- Metric learning and feature extraction are used to enhance posture classification.
- Assessing the model's ability to identify hidden yoga poses using a few training instances.

This research aims to contribute to advancements in pose recognition systems with applications in fitness monitoring, virtual training, and wellness technology.

1.2 Problem Addressed

Yoga has gained global popularity as a physical, mental, and spiritual practice. However, mastering yoga postures often requires the guidance of skilled instructors, which might not always be accessible or affordable. Additionally, improper execution of yoga poses can reduce the effectiveness or even cause injuries. This creates a need for an automated system capable of recognizing and evaluating yoga postures to assist practitioners in real time.

The primary use case for this project is to develop an intelligent **Yoga Posture Recognition System** using Few-Shot Learning techniques. Such a system can empower users to:

- **Learn and Improve Postures:** Provide feedback to practitioners on their poses, helping them refine and improve alignment.
- **Assist Instructors:** Enable yoga instructors to monitor multiple practitioners remotely or during large classes.
- **Facilitate Personalized Training:** Help fitness apps provide customized recommendations and track progress effectively.
- **Promote Accessibility:** Allow beginners and those without access to professional trainers to practice yoga confidently and safely.

1.3 Problem Statement

Traditional computer vision models require large amounts of labelled data to perform effectively. However, collecting and annotating datasets with numerous yoga poses is time-consuming, expensive, and labour-intensive. Moreover, yoga postures exhibit high variability due to differences in practitioner body types, angles, and styles. These challenges make relying solely on supervised learning approaches for yoga posture recognition impractical.

The project addresses the following problems:

- **Data Scarcity:** Limited availability of labelled examples for each yoga pose.
- **Pose Variability:** Significant intra-class variation due to differences in body proportions, clothing, camera angles, and execution styles.
- **Class Imbalance:** Certain yoga poses are more common than others, leading to imbalanced datasets.
- **Generalization to New Poses:** The system must recognize new poses with only a few labelled examples or none.
- **Real-Time Performance:** Achieving fast and accurate recognition for real-time feedback to users.

1.4 Why Few-Shot Learning ?

Few-shot learning (FSL) is an ideal solution for this use case as it enables models to generalize effectively from a few training examples. By leveraging meta-learning, prototypical networks, or contrastive learning, FSL models can learn transferable features and adapt quickly to new pose categories with minimal data.

Applying FSL to the Yoga-82 dataset allows for robust recognition of diverse poses while significantly reducing the data annotation effort. This approach is well-suited for real-world scenarios where data collection and labelling are resource-constrained, paving the way for scalable and user-friendly yoga training systems.

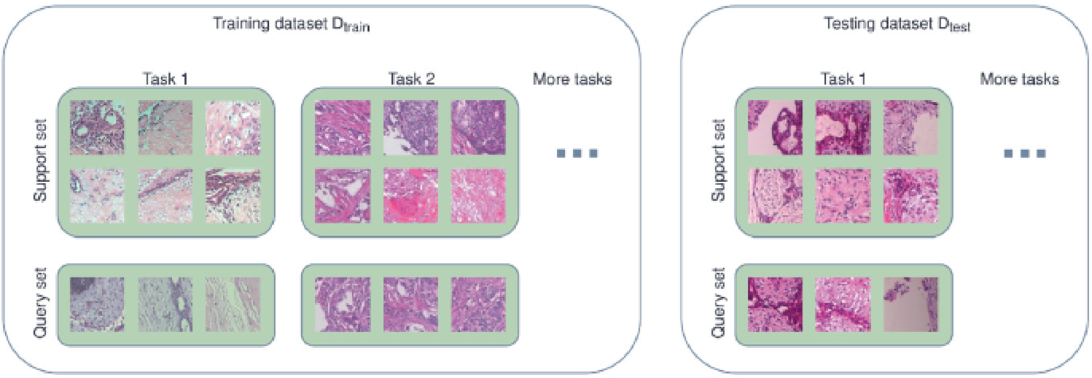

FIGURE 1.1: The division of Dataset in Few shot Learning Approach.[3]

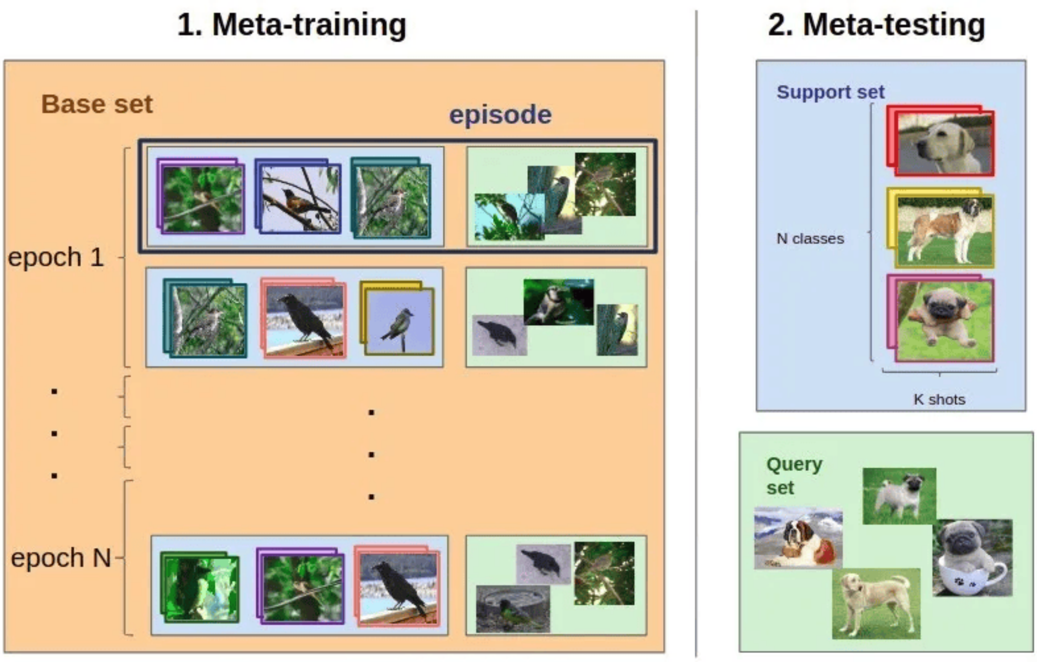

FIGURE 1.2: Few Shot learning - An Episodic Learning approach.

Chapter 2

Literature Review

2.1 Introduction

The emergence of few-shot learning in machine learning has shown why it is currently one of the most fascinating research areas, as it has some solutions for tasks with very few labelled data. A quality review of previous research strengthened the basis of this research as it helped to find appropriate models and datasets in line with the expectations of this research.

While studying past research, we understood the best techniques, advantages, and shortcomings, if any. This helped us choose the most efficient methods and architectures in similar cases preceding ours, thus helping our research. Additionally, the papers identified typical issues and obstacles, such as overfitting, lack of potentiation and biases in the dataset, which we sought to mitigate in our approach.

The efficiency of our experimental scheme was increased because of the complexity of the literature used. We considerably shortened the period of adjustments to the process and thus tended to a more constructive line of research.

2.2 Representation Learning by Detecting Incorrect Location Embeddings [1]

2.2.1 Summary:

With a significant emphasis on form understanding through positional embedding mismatches and sparse inputs, DILEMMA is a reliable and efficient method for self-supervised representation learning. It reduces the computational load, offers a general performance gain across various tasks, and forms the foundation for developing efficient and successful vision models.

2.2.2 Key Idea:

- **Central Idea:** For ViT, the image is broken into patches and treated as sequences of tokens. Of these tokens, only some are purposefully mis-embedded regarding their positions. The model is thus trained to know the difference between "correct" and "incorrect" embeddings, enabling the model to learn about spatial and structural relationships instead of just textures.
- **Sparsity Mechanism:** To enhance efficiency, the input is trained to be sparse (i.e., random patches are removed) to make the model more robust to missing information and less expensive in computation.

2.2.3 Key Contributions:

- **DILEMMA Loss:** Incorporates a binary classification loss (for pinpointing positional discrepancies) along with existing contrastive learning methods. Promotes the model to give less weight to texture-based characteristics and more to shape-based ones, which aligns with the finding that shape bias is beneficial for generalization.
- **Integration with SSL Methods:** Implemented in widely used SSL platforms like MoCoV3, SimCLR, and DINO, providing the same performance lift in various tasks within the same computational budgets.
- **Efficiency:** Flexible sparsity offers less memory and computation overhead, resulting in larger batch sizes and quicker training.

2.2.4 Key Takeaways:

- **Shape Bias over texture:** One of the advantages of DILEMMA is that it glorifies the use of shapes in object retrieval, which assists in effective generalization in machine tasks such as pose classification and segmentation of objects.
- **Sparsity as a Feature:** Randomized token dropping enhances the model's efficiency and allows for better suppression of occlusions, thus minimizing the gap between the training and testing domains.
- **Compatibility:** DILEMMA perfectly operates with its fellow techniques (e.g., MoCoV3, DINO, and MAE mask autoencoders), ideally without altering the underlying structures.

2.3 An Efficient Deep Convolutional Neural Network Model For Yoga Pose Recognition Using Single Images [2]

2.3.1 Summary:

This work looks to improve complex yoga pose estimation performance from a single RGB image using the recently proposed Convolutional Neural Network (CNN) model called YPose.The work aims at occlusion, inter-class similarities, intra-class differences, and different angles of view.

2.3.2 Methodology:

- **ROI Segmentation:** This involves the adoption of mask R-CNN techniques to delete everything but the ROI segmentation for more focus on the image input.
- **EfficientNet Backbone:** This uses the EfficientNet B4 architecture for feature extraction.
- **Dense Refinement Blocks:** Better use of feature representation in complicated poses by adding the dense refinement layers aimed at DenseNet.
- **Classification:** Here, global average pooling and fully connected layers are used, allowing for the classification of three levels of the hierarchical eighty-two yoga pose (6, 20, and 82 classes).

2.3.3 Results and Takeaways:

- State-of-the-art accuracy of 93.28% (Top-1) and 98.04% (Top-5) as recorded on the Yoga-82 dataset, surpassing previous methods by nearly 13.9%.
- She exhibited tolerance to background noise, occlusions, and high levels of complexity in posture while remaining competitive in parameter efficiency (22.68M parameters).

2.3.4 Advantages:

- A practical and lightweight model suitable for use in low-resource environments such as Google Colaboratory.
- Demonstrates tolerance to changes in the complexity level of the intended pose, background scenes and the use of artificial images.

2.4 The Summary Table for Literature Survey:

Serial No	Paper Title and year	Description	Outcome	Limitations
1	CAM based fine-grained spatial feature supervision for hierarchical yoga pose classification using multi-stage transfer learning (2024) [4]	Hierarchical yoga pose classification (YPC) technique within a multi-stage, multi-tasking framework using transfer learning. Approach involves three stages of training: Stage 1: Uses a special combination of three loss functions—cross-entropy, self-supervised contrastive loss, and supervised contrastive loss. This helps the model learn to distinguish between different yoga poses better. Stage 2: Fine-tunes the model using the cross-entropy loss function to further improve accuracy.Stage 3: Uses an Encoder-Decoder network with an attention mechanism and integrates detailed spatial features from HiResCAM and XGrad-CAM. This helps the model understand finer details of the yoga poses.	Achieves state-of-the-art performance in hierarchical yoga pose classification. Peak Top-1 accuracy reported is 95.89% for Yoga-6, 93.85% for Yoga-20, and 90.0% for Yoga-82 classes, giving an improvement of 6.1%, 9.3%, and 10.9% respectively over previous state-of-the-art methods. Use of multi-stage training approach with novel loss functions and attention mechanisms is effective in enhancing the model's performance on complex yoga pose datasets.	Relies heavily on the availability and quality of hierarchical annotations in the dataset, which might not be available for all pose datasets. Computational complexity and resource requirements for training the multi-stage framework can be high leading to difficulties in real-time deployment on edge devices.
2	Deep-Learning-Based Yoga Pose Classification Using Wavelet Transform (2024) [5]	A deep-learning model for yoga pose classification, aiming to build a quality assessment and personalized feedback system for yoga practitioners. The model employs wavelet transform on input images to extract subbands (approximation, horizontal, vertical, and diagonal coefficients). Each subband is then processed through separate convolutional neural networks (CNNs). The results from these networks are fused to predict the final yoga pose class. The proposed model's performance is compared with a standard CNN model and models using individual subbands.	The proposed wavelet-based model significantly outperforms the regular CNN model. The accuracy for the proposed model is 91% for training data and 80% for test data, compared to 61% and 50%, respectively, for the standard CNN model. The results indicate that wavelet-based decomposition improves the classification accuracy by effectively capturing detailed information in the images.	The model relies on the availability of sufficient data, which is mitigated by data augmentation in this study. The runtime is higher compared to a regular CNN due to separate processing of each subband, although parallelization can potentially address this issue. The computational requirements may pose challenges for real-time applications.

3	Robust Classification of Similar Yoga Poses based on Grouping and Hierarchical Inference (2023) [6]	A method to improve the classification of similar-looking yoga poses using a hierarchical inference and grouping technique. The approach enhances a pre-trained DenseNet121 model and incorporates two methods - XGBoost and YOLOv5 - for refining predictions. The DenseNet121 model which is pre-trained on ImageNet, is fine-tuned on the Yoga-82 dataset. The hierarchical classification framework uses DenseNet121 for parent labels, followed by local classification within each parent label group using XGBoost and YOLOv5.	The proposed hierarchical method achieved an accuracy of 87.33%, which is an improvement over the method by Verma et al. (CVPR 2020) which achieved 79.31%. Additionally, the fine-tuned DenseNet121 model achieved an accuracy of 88.89%, demonstrating significant enhancement over previous models.	The model's performance heavily relies on the quality and completeness of the Yoga-82 dataset, which may contain broken URLs or low-quality images. The computational complexity, particularly in training and fine-tuning multiple models, poses challenges for real-time deployment. Additionally, the hierarchical approach may not be as effective for datasets with less structured hierarchical relationships.
4	Yoga Posture Analysis using Deep Learning (2024) [7]	This paper uses deep learning techniques, particularly Convolutional Neural Networks (CNNs), for the analysis and recognition of yoga poses. The research focuses on three main algorithms: YOLO (You Only Look Once), MediaPipe, and OpenPose. Each of these algorithms is leveraged to identify key alignments in yoga postures and measure the correctness of these poses. The YOLO algorithm, traditionally used for object detection, has been adapted to detect joints and measure angles for various yoga poses. MediaPipe and OpenPose provide additional frameworks for detecting and analyzing yoga postures using pre-trained models and real-time data processing.	The study demonstrates that deep learning models, specifically YOLO, MediaPipe, and OpenPose, are effective in accurately recognizing and analyzing yoga poses. Initial results indicate that YOLO achieves an average precision of 92.5% on the dataset, outperforming MediaPipe with 91.80% precision and OpenPose with 85.94% precision. These models are capable of identifying key body joints and calculating angles between them and can provide real-time feedback on posture correctness.	The effectiveness of the models relies heavily on the availability and quality of labeled datasets, such as Yoga-82 and Kaggle datasets. Inadequate or low-quality data can significantly impact model performance. The current models are trained to detect only a few yoga poses, which limits their applicability. Expanding the pose detection capabilities would require additional data and computational resources.

5	Yoga Pose Estimation with Machine Learning (2023) [8]	This paper focuses on enhancing pose estimation systems using OpenPose for keypoint detection, SMOTE for addressing class imbalance, and LightGBM for classification. OpenPose detects specific body points for precise pose estimation, SMOTE oversamples minority classes to balance the dataset, and LightGBM provides high accuracy and faster training speeds on large datasets.	The application of OpenPose, SMOTE, and LightGBM lead to improved accuracy and performance in pose estimation tasks, with a significant increase in classification accuracy by 15% and a reduction in training time by 30% when compared to traditional CNNs and Deep Learning models.	The method may face challenges related to computational complexity and resource requirements, particularly during the integration and training of the three methodologies. Additionally, the effectiveness of SMOTE in oversampling might not generalize well across all types of pose datasets.
6	Yog-Master: Detection and Correction of Yoga Postures using Augmented Reality (2023) [9]	The paper presents a web-based application aimed at detecting and correcting yoga postures using Augmented Reality (AR). The system utilizes a camera to capture the user's yoga poses, TensorFlow MoveNet to identify 17 key body points, and a neural network for classification and correction. Users receive visual and audio cues for corrections. The application integrates React for the front-end and Keras with Node.js for the backend, leveraging datasets like Yoga-82 for training.	The model has achieved an accuracy of 99.47% in detecting and classifying yoga poses. It provides real-time feedback and corrections, enhancing the quality and safety of yoga practice.	The system's accuracy depends on the quality of the captured images and the diversity of the training dataset. Computational complexity and the requirement for real-time processing may pose challenges for deployment on devices with limited resources.
7	Yoga Pose Recognition Based On Convolutional Neural Networks (2023) [10]	This study investigates the application of Convolutional Neural Networks (CNNs) for yoga pose recognition, aiming to facilitate automated yoga self-training systems. The research employs various CNN architectures, excluding powerful models like VGG-16, AlexNet, and GoogleNet initially, before conducting a detailed analysis of these models via transfer learning for classifying 82 yoga classes. The experiment focuses on evaluating sensitivity, precision, F1-score, specificity, Top-1 accuracy, and Top-5 accuracy to determine the most suitable model for yoga pose classification.	The study identifies GoogleNet as the most appropriate model for classifying 82 yoga classes, achieving superior performance metrics compared to VGG-16 and AlexNet. GoogleNet exhibits high percentages across key evaluation parameters, including specificity (99.94%), sensitivity (80.14%), precision (92.97%), F1-score (85.93%), Top-1 accuracy (87.89%), and Top-5 accuracy (97.08%). Transfer learning emerges as a valuable technique for enhancing classification performance, particularly with smaller datasets.	The pose variety dataset may not encompass all possible yoga poses, potentially limiting the generalizability of the findings. Additionally, the exclusion of certain powerful CNN models in the initial experiment phase may influence the comprehensiveness of the comparison. The training process requires substantial computational resources, posing challenges for real-time deployment on edge devices.

8	Prediction of Yoga Pose from YouTube Dataset using Skeleton Feature Extraction Based ISDL Model (2024) [11]	An automated system for yoga pose detection and classification based on deep learning techniques. Advancements in computational probing to facilitate precise yoga practice by providing real-time feedback on posture accuracy. The methodology involves a multi-stage process comprising preprocessing, feature extraction using skeleton-based models, classification utilizing an improved synergic deep learning (ISDL) model, and further refinement using the Archimedes optimization algorithm (AOA). Preprocessing involves noise removal and histogram equalization to enhance image quality. Skeleton-based models extract spatial features, which are then input into the ISDL model for classification. The AOA is employed to fine-tune model parameters, improving overall performance.	The proposed ISDL model demonstrates superior performance in yoga pose classification compared to state-of-the-art methods, achieving high accuracy levels. Experimental results validate the effectiveness of the approach, with classification accuracy reaching 96%, sensitivity at 98%, and specificity at 91%. The system's ability to provide real-time feedback on posture accuracy can significantly enhance yoga practice, promoting physical and mental well-being.	The study relies on data extracted from YouTube videos, potentially limiting dataset diversity and generalizability. Additionally, the preprocessing techniques may not fully address variations in image scale, resolution, and illumination, impacting classification accuracy. Additionally, considerations for optimizing computational resources and real-time deployment on practical devices are necessary for widespread adoption.
9	Segmentation quality assessment network-based object detection and optimized CNN with transfer learning for yoga pose classification for health care (2023) [12]	This paper presents an approach for yoga pose classification utilizing a Remora Lion Algorithm (RLA)-based Convolutional Neural Network (CNN) with Transfer Learning (TL). It addresses challenges such as illumination variation, image resolution, and clothing deviations. Trained models include AlexNet, VGG-16, ResNet-50, DenseNet-201, Xception, MobileNet-v2, ShuffleNet, EfficientNet-b0, SqueezeNet, and GoogleNet. The RLA integrates Remora Optimization Algorithm (ROA) and Lion Algorithm (LA).	The proposed RLA-based CNN with TL achieves significant performance improvements in yoga pose classification, providing high accuracy and robustness against various challenges. Experimental results demonstrate high accuracy (0.944), sensitivity (0.957), specificity (0.937), low CV ratio (4.554), and high ROC value (0.9943) on the Yoga-82 dataset.	Limitations include dependency on specific algorithms and datasets, potentially limiting generalizability. Additionally, the computational complexity of training and deployment could hinder real-time application, especially on resource-constrained devices.

10	Dense-PoseCompare: A Comparative Study of DenseNet Models in Yoga Pose Classification (2024) [13]	This paper uses DensePoseCompare framework, which performs an extensive investigation of DenseNet models for precise and effective yoga pose detection. It compares three DenseNet architectures - DenseNet121, DenseNet169, and DenseNet201 using a dataset of 1551 yoga posture images, including downdog, goddess, plank, tree, and warrior2. The study involves fine-tuning these models on 1081 training images and evaluating their performance on 470 testing images. The methodology incorporates deep transfer learning, using pre-trained DenseNet models on ImageNet, followed by fine-tuning with the yoga posture dataset. The study also utilizes the explainable AI method LIME to provide insights into the decision-making process of the models.	DenseNet201 outperforms the other models, achieving testing accuracy of 99.07%. This indicates the effectiveness of DenseNet architectures in capturing detailed aspects of yoga postures and highlights the potential of fine-tuning deep transfer learning approaches for yoga pose classification.	The approach relies on a small dataset, which may limit the generalizability of the results. The computational complexity and resource requirements for training and deploying these models can be significant. The need for high-quality annotated data is crucial for achieving optimal performance, which might not always be available.
11	Handcrafted Feature Assisted Lightweight Encoder Decoder based Classifier for Yoga Posture Recognition (2024) [14]	The proposed approach utilizes a lightweight neural network architecture based on an encoder-decoder framework, incorporating keypoint-based pose estimation from Google's MoveNet tool. The classifier operates in two phases: feature set formation and classification. In the first phase, keypoints are extracted from the Yoga-82 dataset using MoveNet, and handcrafted features such as angular and distance measurements are calculated and merged with these keypoints. In the second phase, the combined features are fed into an encoder-decoder network with an output softmax layer to classify yoga postures.	The proposed model performs well on the Yoga-82 hierarchical dataset, achieving improved accuracy when incorporating keypoints extracted using MoveNet. The results indicate that the Adam optimizer offers better consistency compared to RMSprop in training and validation phases. The encoder-decoder architecture is effective in classifying yoga postures with high accuracy, demonstrating the potential of lightweight models in this domain.	The model relies heavily on the quality of keypoint extraction and handcrafted features, which may introduce errors in different environments. Additionally, challenges such as occlusion, camera invariance, and multi-person pose recognition need to be addressed for broader applicability.

12	Transfer Learning-Based Method for Classifying Yoga Poses using Deep Convolutional Neural Networks (2023) [15]	A technique for automatically identifying and classifying yoga positions in photos is presented in the study article. The paper builds an adequate model, TLMYPC-DCNN, for precisely categorizing yoga poses using deep CNNs and transfer learning. The suggested model performs well in precisely and consistently classifying yoga positions. It uses the Yoga Pose Dataset and MobileNetV2 for picture classification. The difficulties and possible uses of automated yoga position recognition are also covered in the article.	The method described in this research, the Transfer Learning-Based Method for Classifying Yoga Poses using Deep Convolutional Neural Networks (TLMYPC-DCNN), uses transfer learning and deep CNNs to identify yoga poses accurately. The Yoga Pose Dataset assesses the model's performance, revealing an efficient classification of yoga positions. Future developments could involve investigating the use of various pre-trained models to increase classification accuracy and investigating the use of the suggested technique for real-time yoga posture recognition.	The difficulty of precisely identifying yoga postures because of a variety of conditions, including low illumination, erratic video quality, poor visibility, closeness, background noise, and camera angle, is the research paper's method's weakness. These factors may affect how well the suggested method correctly categorises yoga poses. Furthermore, the results may only apply to a narrow range of yoga poses and practitioners due to the dependence on Electromyography for examining leg muscles and using a particular group of participants and yoga positions.

13	Yoga Pose Detection and Identification Using MediaPipe and Open-Pose Mode (2023) [16]	The paper discusses the development of a yoga pose estimation system using the MediaPipe and OpenCV technologies. It emphasizes the importance of maintaining proper posture during yoga to prevent injury and discomfort. The proposed system combines the benefits of both MediaPipe and OpenCV to provide real-time feedback on the user's posture, enabling fitness enthusiasts to improve their technique and form. The document also explores studies on human pose estimation using deep neural network-based landmark detection, computer vision libraries such as OpenCV, and machine learning frameworks like MediaPipe.	The research paper's outcome is the development of a deep neural network-based yoga posture estimation system using MediaPipe and OpenCV technologies. The system provides real-time feedback on posture and proposes corrective actions to improve technique and form during yoga practice. Future aspects of the research include the potential integration of additional sensors, such as pressure sensors or inertial measurement units, and the exploration of more complex deep-learning models for enhanced posture estimation. Additionally, the system's applications in physical therapy, athletic training, and business ergonomics are highlighted as potential areas for further improvement and implementation.	The limitations of the method in the research paper include the challenge of limited publicly accessible datasets for yoga poses, which hinders research in yoga posture categorization. Additionally, using real-world data in the proposed system may pose challenges in maintaining consistent backdrop and illumination, potentially affecting the system's performance. Furthermore, the study acknowledges the need for future research to address the complexity of various yoga positions and to consider user traits such as degree of fitness, injury history, and body type for more personalized posture corrections.
14	Yoga Pose Detection Using Long-Term Recurrent Convolutional Network (2023) [17]	The paper presents a novel approach for yoga pose detection using a Long-Term Recurrent Convolutional Network (LRCN), which combines Convolutional Neural Network (CNN) for spatial feature extraction and Long Short-Term Memory (LSTM) for temporal predictions. The model, trained on a dataset of 88 videos featuring six different yoga asanas, achieves an accuracy of 81%. The research, conducted by authors affiliated with KLE Technological University, Hubballi, India, and presented at the 2023 IEEE 8th International Conference for Convergence in Technology (I2CT), aims to contribute to the field of Human Activity Recognition (HAR) and assist in the correct performance of yoga poses for health and wellness.	The paper reports creating a model capable of recognizing six yoga poses with an accuracy of 81%. This model utilizes a hybrid approach of CNN and LSTM to analyze and identify poses from video frames, independent of open pose or pose net methods. The authors propose integrating the model into mobile applications to provide real-time guidance for yoga practitioners. They also plan to enhance the model by incorporating more intricate yoga poses, which could broaden its applicability and make it a more comprehensive fitness and health technology tool.	The paper does not explicitly state the limitations but inferred limitations include the dataset's potential lack of diversity, the model's generalization capability to unseen poses and different individuals, real-time performance on devices with limited processing power, robustness to environmental changes, and scalability to a broader range of poses or exercises. The authors' future work addresses these by expanding the model's pose recognition capabilities and enabling its use on mobile devices.

15	HARNet: design and evaluation of a deep genetic algorithm for recognizing yoga postures (2024) [18]	The paper introduces HARNet, an interactive toy system designed to help autistic children learn yoga through deep learning and IoT technologies. It utilizes a deep genetic algorithm to recognize yoga postures with high accuracy and incorporates touch sensors and a camera for real-time feedback. The toy aims to improve mental abilities, social interaction, and language development in autistic children, leveraging the therapeutic benefits of yoga. The research includes a survey on the toy's effectiveness, a detailed methodology, and suggestions for future enhancements, highlighting the potential of HARNet for other activity recognition applications.	The research outcomes include the successful development of the HARNet system, which demonstrates a high accuracy of 98.52% in recognizing yoga postures from the Yoga-82 dataset. The system's ability to provide interactive and engaging learning experiences for autistic children shows promise in improving their mental abilities and social skills. Future aspects of the research include enhancing the toy with additional resources, improving its sustainability, and exploring the generalization of the HARNet algorithm for broader applications in human activity recognition.	The potential limitations of the method could include the need for a controlled environment to ensure accurate yoga pose recognition, the limited number of yoga postures the system can recognize, and the reliance on specific hardware components that may require maintenance or upgrades. Additionally, the system's effectiveness may vary among autistic children, and further research is needed to validate its long-term benefits and adaptability to different settings and needs.

16	Exploration of deep learning architectures for real-time yoga pose recognition (2023-24) [19]	The paper ”Exploration of Deep Learning Architectures for Real-time Yoga Pose Recognition” by Sumeet Saurav, Prashant Gidde, and Sanjay Singh explores the development and evaluation of deep learning models for real-time yoga pose recognition. It discusses the significance of vision-based assistive technologies and the potential applications of yoga pose recognition systems in self-training, monitoring, and assistance for practitioners. The paper presents various deep learning architectures, including a hybrid CNN and LSTM model and different fine-tuned variants, for accurate and efficient recognition of yoga poses in video sequences.	The paper presents the development and evaluation of deep learning models for real-time yoga pose recognition, achieving high recognition accuracy and computational efficiency. The prospects include the potential application of the developed models in self-training, monitoring, and assistance for yoga practitioners and their deployment on resource-constrained embedded platforms. The outcomes demonstrate the practical applicability of the models for promoting safe and effective yoga practices, with scope for further advancements in vision-based assistive technologies.	The limitations of the method described in the paper include the reliance on a dataset consisting of video sequences of only six yoga poses, potentially limiting the generalizability of the models to a broader range of yoga postures. Additionally, while capable of classifying yoga poses in real-time on resource-constrained embedded platforms, the proposed techniques still need to be suitable as a base to build a tutorial, indicating a need for further development to support comprehensive self-training applications. Furthermore, the use of IMUs in some pose recognition systems may not be widely accepted by yoga practitioners, potentially impacting the practicality and adoption of such technologies.

17	Novel deep learning models for yoga pose estimator (2023) [20]	The paper addresses the creation of innovative deep-learning models for position estimation. It presents the LGDeep model, an ensemble-based methodology that integrates feature extraction techniques like LDA and GDA with Xception, VGGNet, and SqueezeNet. The trial results illustrate how precise and accurate the model is, highlighting its potential influence on enhancing the efficacy and safety of yoga practice. Along with a thorough literature review, an in-depth methodology, and experimental results, the research highlights the advancement and significance of the proposed LGDeep model in the field of yoga position classification. Additionally, the publication highlights the LGDeep model's potential uses outside of yoga posture recognition and recommends future studies to improve the model's generalizability.	The paper's outcomes include introducing four novel ensemble-based yoga pose estimation models, with the LGDeep model achieving 100% classification accuracy, surpassing previous approaches. The future aspects involve the potential application of the LGDeep model beyond yoga pose recognition and the need for further research to enhance the model's generalizability and user accessibility. Additionally, the paper emphasizes the significance of computerized reasoning-based programs in recognizing yoga poses and offering personalized advice to improve practitioners' posture.	The research paper introduces four novel ensemble-based deep learning models for yoga pose estimation, showcasing their superior performance compared to individual DL techniques. The proposed LGDeep model achieves high accuracy and precision, outperforming previous state-of-the-art approaches. The study emphasizes the significance of accurate yoga pose recognition in improving practitioners' safety and efficacy, highlighting the potential applications of the LGDeep model beyond yoga pose recognition. Additionally, the paper suggests future research directions, including dataset expansion and the development of user-friendly interfaces for the yoga posture recognition system to enhance its generalizability and accessibility.

18	Enhanced Yoga Posture Detection using Deep Learning and Ensemble Modeling (2023) [21]	The researchers aim to develop a deep learning-based solution to detect yoga poses accurately. The goal is to outperform all the previous works that used the Yoga-82 dataset in terms of accuracy. To classify the yoga poses, the researchers experimented with several pre-trained CNN models, including DenseNet, Xception, and MobileNet. They slightly altered the models to improve their performance. Additionally, they used an ensemble modelling approach to combine the predictions of the individual models and further boost the accuracy.	The researchers achieved significant improvements in accuracy compared to previous works that used the Yoga-82 dataset. For the 82 yoga poses, the researchers reported a top-5 accuracy of 94.54% and a top-10 accuracy of 96.81% using the Xception model. For the 6 major yoga classes, the researchers achieved a Top-5 accuracy of 95.34% and a Top-10 accuracy of 97.58% using the Xception model.	The current system has encountered overlapping issues with the yoga poses, which need to be addressed through further studies. The researchers mentioned plans to develop a smartphone application for pose detection, but this was not included in the current research, limiting the practical application of the proposed approach.
19	iSmartYog : A real time Yoga Pose recognition and Correction Feedback Model using Deep Learning for Smart Healthcare (2023) [22]	The research proposes an AI-based " iSmartYog " system for real-time yoga pose recognition and correction feedback using Deep Learning Techniques. The researchers created a new video dataset called "SmartYog" consisting of 2700 videos of 24 different yoga poses. The researchers developed a hybrid deep learning model that combines a Convolutional Neural Network (CNN) and a Gated Recurrent unit (GRU). The CNN extracts spatial features from the body key points, while the GRU captures the temporal dependencies in the yoga pose sequences.	The iSmartYog model achieved high accuracy levels with a training accuracy of 98.17%, validation accuracy of 93.80%, and testing accuracy of 93.28% on the diverse SmartYog dataset. It also outperformed state-of-the-art models on other public yoga datasets. The iSmartYog model is designed to work with standard RGB cameras, making it cost-effective and suitable for integration into everyday devices and mobile applications for innovative healthcare. The real-time feedback mechanism of this model can detect errors in yoga poses and provide tailored feedback to users, helping them correct their poses and prevent injuries.	A larger dataset containing more asanas could be gathered to improve the accuracy of the proposed model, given the numerous asanas in yoga science. Future research could focus on integrating the model into smart healthcare applications and creating new datasets for specific yoga kriyas like Sudarshan Kriya, cyclic meditation, Hast Mudras, Surya Namaskar and Surya Kriya.

20	A Novel Fuzzy Logic-based Method for Modeling and Recognizing Yoga Pose (2023) [23]	The research proposes a novel method for modelling and recognizing yoga poses using fuzzy logic and a multi-layer perceptron (MLP) neural network. Using the Multi-Layer perceptron for Yoga pose recognition. Due to the degree of uncertainty in human poses, we apply fuzzy logic to improve posture modelling and recognition of human poses. The OpenPose library is used to extract key points from human poses in images, capturing the 3D coordinates of 32-key points on the body.	The proposed method for recognizing yoga movements using a neural network classification and fuzzy logic was tested on a yoga dataset. An average accuracy of 87.48% can be attained with neural networks. Ten rules and seven membership functions are defined in the fuzzy model. The fuzzy model detected the plank and the goddess poses with 91.3% and 65% accuracy, respectively.	The proposed method was tested on a dataset with only five yoga poses (goddess, plank, down-dog,warriors2 and tree). The accuracy of the method in a wider range of yoga poses needs to be discussed. The article compares the proposed method with only a few other methods, such as Yoga-82 and Blaze Pose. A more comprehensive comparison with other state-of-the-art yoga pose recognition techniques would provide a better understanding of the method's performance.
21	An Analytical Comparison of Deep Learning Frameworks for 2D Image-Based Hatha-Yoga Pose Identification (2024) [24]	The study aims to create a model that can classify yoga poses accurately to help practitioners maintain correct form and avoid injuries associated with improper postures. The researchers collected a dataset of 1248 images across 20 different hatha yoga poses and preprocessed the data using techniques such as resizing, normalization, and augmentation. The paper discusses the use of transfer learning with pre-trained models like ResNet50, VGG16, VGG19, EfficientNetB0, MobileNetV2, InceptionV3, Xception, DenseNet121, and DenseNet201, which were fine-tuned for the task of yoga pose classification.	DenseNet201 model achieved a top-1 classification accuracy of 90.4%, outperforming other models like ResNet50, VGG16, VGG19, EfficientNetB0, MobileNetV2, InceptionV3, Xception, and DenseNet121. Among the best performers, EfficientNetB0 stood out with a 98.4% accuracy rate for subsequent investigation concentrated on DenseNet201.The findings suggest that the DenseNet201 model could be a valuable tool for developing automated yoga training systems that offer real-time feedback and support self-guided practice.	The research does not address the application of the model in real-world settings, such as its performance in varying lighting conditions, with different camera angles, or with occlusions (e.g., when body parts are hidden from view). The model's robustness to common image artefacts such as noise, blur, or compression is not evaluated, which could affect its practical use. The paper focuses on pose classification but does not discuss how the model would provide corrective feedback to users, which is an essential aspect of an interactive yoga training system.

22	Fine-Grained Sports, Yoga, and Dance Postures Recognition: A Benchmark Analysis (2023) [25]	The research paper presents the SYD-Net approach for fine-grained human posture recognition, focusing on sports, yoga, and dance postures. It introduces new image datasets for dance styles and sports actions and proposes a patch-based attention (PbA) module to improve posture recognition accuracy. The SYD-Net model integrates spatial attention and channel attention on standard backbone convolutional neural networks (CNNs). Additionally, random region erasing data augmentation is applied to improve accuracy.	The paper discusses the performances of SYD-Net on various datasets, showcasing superior accuracy compared to existing works and benchmarking the proposed approach against prior research. The results demonstrate the efficacy of SYD-Net in achieving state-of-the-art accuracy in fine-grained human posture recognition tasks. Furthermore, the paper presents an ablation study that evaluates the effectiveness of different components of SYD-Net, providing insights into the significance of each module in improving the system's accuracy.	There is a need for a benchmark public image dataset with sufficient interclass and intraclass variations to address sports and dance posture classification. Additionally, the paper acknowledges the need for publicly available image-based datasets with diverse variations of dance styles. Furthermore, the research focuses on specific categories of sports, yoga, and dance postures, potentially limiting the generalizability of the proposed SYD-Net model to other fine-grained human posture recognition tasks.

Chapter 3

Proposed Work

3.1 Prototypical Network:

Prototypical Networks are designed to address few-shot learning by representing each class with a single prototype, computed as the mean of embedded examples for that class. These prototypes are used for classification by measuring the distance between a query point and class prototypes in an embedding space.

3.2 Advantages:

The approach is computationally efficient, reduces model complexity, and achieves superior performance by leveraging simple yet effective inductive biases.

3.3 Proposed Methodology:

3.3.1 Architecture and Embedding:

The network will use the following architecture to transform inputs into a learned embedding space:

- **Input:** Accepts RGB images of dimensions (128 X 128 X 3).
- **Convolutional Layers:** The first convolutional layer applies *32 filters* of size (3 X 3) with the **ReLU activation function**:

  ReLU(x) = max(0,x)

  where (x) is the input value. ReLU introduces non-linearity and avoids the vanishing gradient problem.

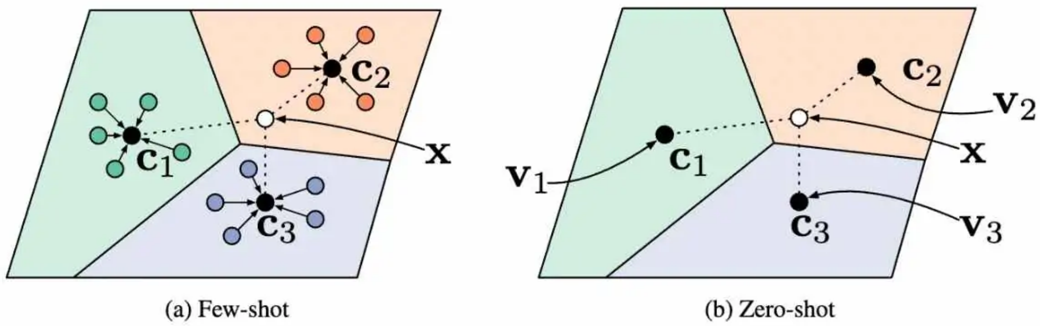

FIGURE 3.1: A prototypical Network implemented With few shot Learning.

- Followed by a (2 X 2) MaxPooling operation to down-sample feature maps.
- The second convolutional layer applies **64 filters** of size (3 X 3) with ReLU, followed by another (2 X 2) MaxPooling layer.
- **Flatten Layer:** Converts the multi-dimensional feature maps into a 1D feature vector.
- **Dense Layer:**
  - A dense layer with **64 neurons** applies ReLU for intermediate feature extraction.
  - The final dense layer (embedding layer) outputs a **1024-dimensional embedding vector** for each input image, representing its position in the metric space.

3.4 Prototype Computation:

- Compute the prototype (C_k) for each class (k) as the mean of the embeddings of support set examples for that class:

  where $(f_\phi(x_i))$ is the embedding of input (x_i) and (S_k) is the set of support examples for class (k).
- **Classification Process:**
  - For a query point (x), compute the distance to each class prototype using squared Euclidean distance:
  - Generate a probability distribution over classes using the softmax function.
  - Assign the query point to the class with the highest probability.

- **Training Procedure:**
  - **Episodic Training:** Simulate few-shot learning tasks during training by constructing episodes. Each episode consists of:
  - A **support set** to compute prototypes.
  - A **query set** to evaluate classification and compute the loss.
  - **Loss Function:** Minimize the negative log-probability of the true class y
- **Optimizations:** Use the **Adam optimizer** for gradient-based optimization, combining momentum and adaptive learning rates:

3.5 Zero Shot Learning:

- Replace class prototypes with embeddings of class meta-data (e.g., textual descriptions or attributes):

  where (v_k) is the meta-data vector for class (k) and (g_θ) is a learnable embedding function.

3.6 Implementation Plan:

- **Dataset Preparation:** Resize images to (128 X 128) and normalize pixel values.
- **Model Development:** Implement the network architecture in PyTorch.
- **Optimizer:** Use Adam optimizer with learning rate scheduling.
- **Experimentation:** Conduct few-shot experiments for various N and C.

3.7 Expected Outcomes:

- Improved accuracy in few-shot tasks by leveraging the efficient prototype-based approach.
- Simplified and interpretable classification framework.

Chapter 4

Simulation and Results

4.1 Evaluation Parameters:

When evaluating the performance of a model, particularly for classification tasks, **Top-1 accuracy** and **Top-5 accuracy** are two primary metrics used to measure its effectiveness:

4.2 Top-1 Accuracy:

This measure assesses the proportion of test situations in which the top prediction made by the model corresponds to the correct class. The model's capacity to produce accurate and specific predictions for the task at hand is demonstrated by a high Top-1 accuracy.

4.3 Top-5 Accuracy:

Top-5 accuracy, on the other hand, calculates the proportion of test cases in which the suitable class appears among the model's top five predictions. This measure offers a forgiving assessment, making it appropriate for activities that allow for several conceivable results or where it is difficult to make precise class distinctions.

4.4 Impact of N and C on Model Evaluation:

4.4.1 N (Number of Classes):

Because there is a greater likelihood of confusion between comparable classes, the process gets more difficult as the number of classes increases. For example, yoga position recognition frequently involves minor variances; therefore, Top-5 accuracy becomes exceptionally informative when N is large.

4.4.2 C (Number of Samples per Class):

The performance heavily depends on the amount of data per class (C). If insufficient samples per class are used to train the model, it may not be able to generalize as well and will, therefore, score low in terms of accuracy. A balanced dataset and sufficiently spaced C values provide better values for Top-1 and Top-5 accuracy.

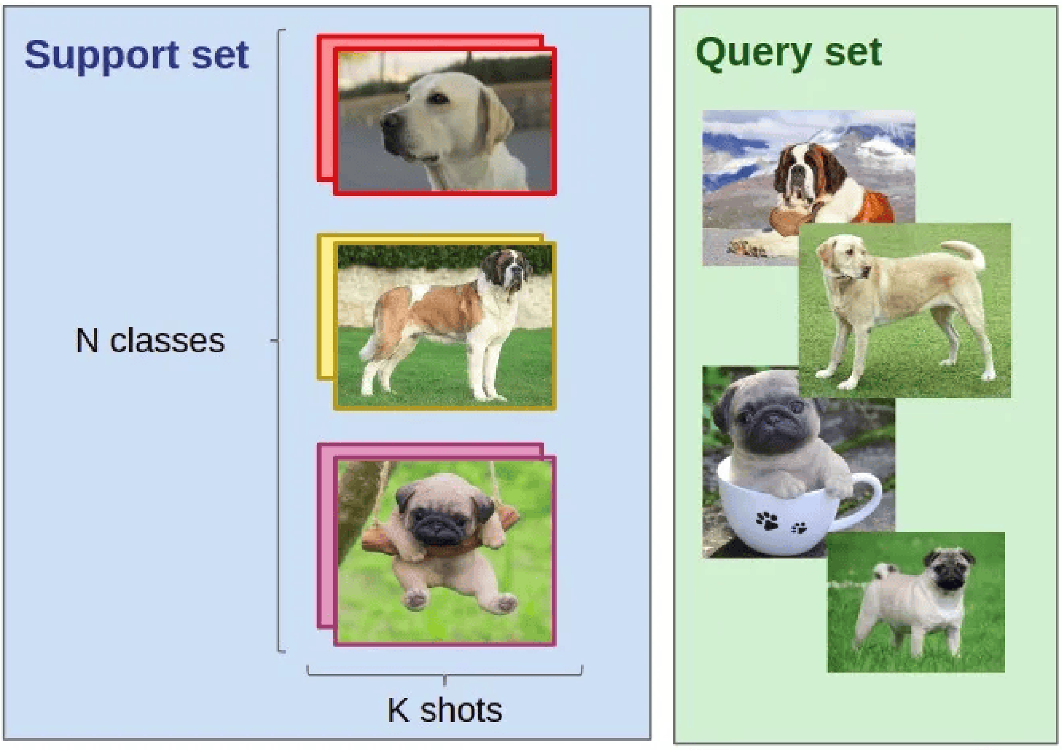

FIGURE 4.1: N classes C shots Episodic Learning.

4.5 Comparing Across Models:

Looking at both Top-1 and Top-5 accuracy for varied values of N and C provides a holistic view of a model's performance. A high Top-1 accuracy is indicative of precision. A large Top-1-to-Top-5 accuracy gap could indicate that there is still room for fine-grained classification improvement. These metrics implicitly advise selecting or improving architectures suited to specific datasets or applications.

TABLE 4.1: Comparison of Models Based on Top-1 Accuracy Scores

No.	Paper Title (Abbreviated)	Accuracy
1	Efficient CNN for Yoga Pose Recognition [2]	93.28%
2	Fine-Grained Sports and Yoga Posture Analysis [25]	79.35%
3	Representation Learning by Detecting Incorrect Embeddings [1]	65.1% (MoCoV), 77% (DINO)
4	CAM-Based Multi-Stage Transfer Learning [4]	90.00%
5	Robust Classification of Similar Yoga Poses [6]	87.33%
6	Yoga Posture Analysis with Deep Learning [7]	92.50%
7	CNN-Based Yoga Pose Recognition [10]	87.89%
8	YogMaster: Yoga Posture Detection with AR [9]	99.47%

4.6 Values Of Our Model:

C values	N values				
	N = 2	**N = 5**	**N = 10**	**N = 15**	**N = 20**
C = 1	0.52	0.182	0.123	0.088	0.065
C = 2	0.5175	0.236	0.126	0.088	0.071
C = 5	0.234	0.244	0.142	0.098	0.081
C = 10	0.135	0.259	0.16	0.104	0.085

TABLE 4.2: Table comparing results for different values of C and N.

4.7 Conclusion from the Table:

The table illustrates the average accuracy for varying values of C (number of samples per class) and N (number of classes). Based on the data of the above table:

4.7.1 Impact of Increasing C:

Increasing C, with more samples per class, continuously improves average accuracy for all N values. This shows how much it matters to have a sufficient number of samples per class if we want a model that generalizes well and performs effectively.

4.7.2 Impact of Increasing N:

With larger N (more classes), average accuracy decreases overall C values- something to be expected since the classification task is going to be more difficult given increasing class numbers, especially with lower C values.

4.7.3 Average Accuracy Across the Table:

The overall average accuracy of the table can be calculated as the mean of all entries:

Average Accuracy = $\sum \frac{all\ values\ in\ the\ table}{Total\ Entries} = \frac{(0.52+0.182+...+0.085)}{20} = 0.160.$

This indicates that the model performs modestly overall, with higher C and lower N yielding the best results.

4.7.4 Key Observations:

- The highest accuracy (0.52) is achieved with C=1 and N=2, demonstrating that classification tasks with fewer classes are easier even with minimal samples.
- For larger values of N, increasing C becomes crucial for maintaining reasonable accuracy. For instance, at N=20, accuracy improves steadily from 0.065 (C=1) to 0.085 (C=10).

In summary, increasing C (sample size) mitigates the negative impact of larger N (number of classes), emphasizing the need for balanced and sufficient datasets for complex classification tasks.

Chapter 5

Conclusions and Future Work

5.1 Future Work Directions:

5.1.1 Cross-Domain Adaptability:

Enhancing the model's capacity to cater to other physical exercises, such as Pilates, Tai Chi or martial stylistics, increases its utility in the fitness and sports services sectors.

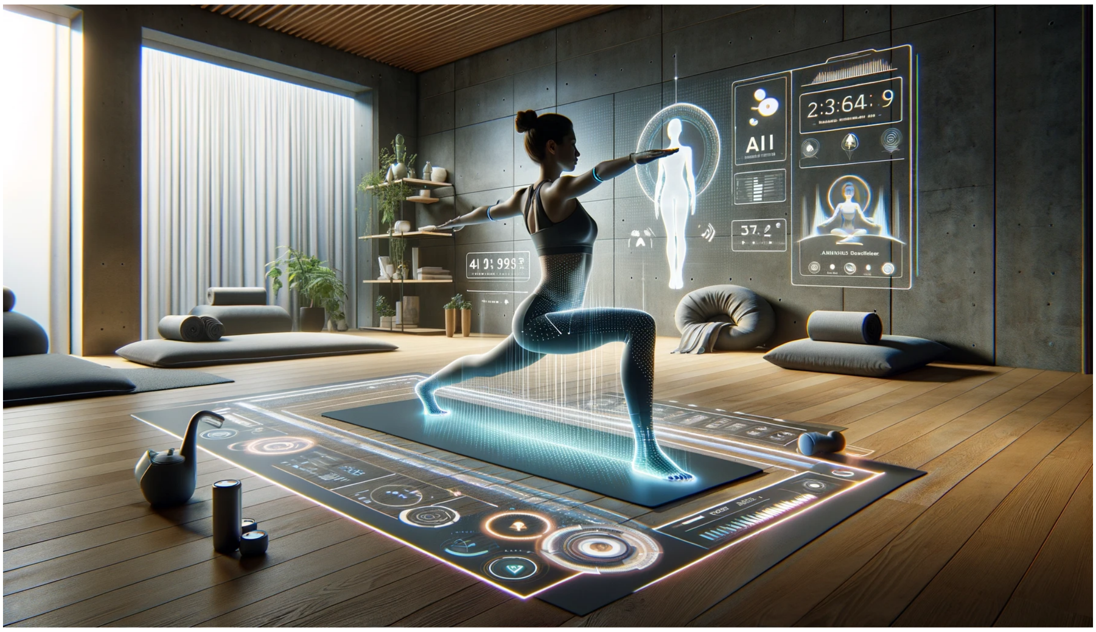

FIGURE 5.1: The future of Yoga Practitioner.

5.1.2 Integration with IoT Devices:

Embed the models in IoT devices, such as intelligent yoga mats or wearables, to improve the efficiency of tracking the pose and provide more biomechanical data.

5.1.3 Multi-language Accessibility:

Integrating computer vision for yoga pose detection with voice navigation systems supports the creation of a training course that is usable throughout the world regardless of the language.

5.1.4 Advanced Analytics:

Deploying pose recognition data helps make inferences on yoga practice to identify such patterns as the popularity of specific poses, statistical injury prevention claims, or efficiency of a workout concerning time analysis.

5.1.5 Healthcare Diagnosis:

In addition to fit club activities, the system could readily be adjusted to endorse medicine, such as postural analysis, to assist in diagnosing disabilities and disorders.

5.1.6 Low-Resource Settings:

Create lightweight and robust models intended for use in low energy-consuming settings, thus ensuring technology deployment in the hinterland and hard-to-reach areas.

5.1.7 Collaborative Learning Systems:

Few-shot models can act as components of CNS learning systems where users and devices contribute to Common Pose Recognition, improving the accuracy and generalization of the recognition over time.

5.1.8 Gamification of Yoga:

Give rise to a gamified yoga mode where synchronization of movements and engagements over a period offers a score for fitness, thus making it attractive to more people.

5.2 Applications of the model:

5.2.1 Custom Made Yoga Classes:

In few-shot learning architectural frameworks, the model could be introduced in mobile phone applications or smart devices, and it will be able to give real-time correction of advanced

FIGURE 5.2: The Application Of few Shots Learning.

yoga postures for each user where no user copies of training data are required. This makes it possible to warm individual users with this novel application without the burden of large-purpose databases for every user.

5.2.2 Health and Fitness Tracking:

The system could help keep track of yoga and make corrections where necessary to yoga positions during practice to avoid injuries owing to wrong postures. The system could also be deployed on smart devices or even intelligent gyms in fitness centres.

5.2.3 Yoga Teachers in the Bag:

Using such models within applications of yoga practice was indeed available. AI models of yoga teach-yourself books incorporated within other fitness apparatus or software can teach yoga to people without professional trainers around.

5.2.4 Rehabilitation and Therapy:

Although a sedentary practice, yoga finds its place in physiotherapy and rehabilitation. Modelling the pose of a subject using shot-learning techniques would play a critical role in assessing the stage of therapy.

5.2.5 Graduate and Post-Graduate Programs:

The said technology may also be implemented in ITARENA centres of excellence, for example, in yoga teachers' training courses when attending students' progress in different postures is assessed from a distance.

5.2.6 Augmented And Virtual Reality:

A few shot frameworks will also be helpful in AR/VR as they will enable yoga pose recognition and engagement in real-time, which will be fun for those who wish to practice yoga.

5.2.7 Content Organisation and Search:

Owners of yoga degradation studios and websites can utilize the models in yoga videos or tutorials about doing the yoga poses where the user is placed and its cataloguing.

Bibliography

[1] S. Sameni, S. Jenni, and P. Favaro, "Representation learning by detecting incorrect location embeddings," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, pp. 9704–9713, 2023.

[2] S. K. Yadav, A. Shukla, K. Tiwari, H. M. Pandey, and S. A. Akbar, "An efficient deep convolutional neural network model for yoga pose recognition using single images," *arXiv preprint arXiv:2306.15768*, 2023.

[3] J. Szołomicka and U. Markowska-Kaczmar, *An Overview of Few-Shot Learning Methods in Analysis of Histopathological Images*, pp. 87–113. Cham: Springer Nature Switzerland, 2023.

[4] C. S. Pradeep and N. Sinha, "Cam based fine-grained spatial feature supervision for hierarchical yoga pose classification using multi-stage transfer learning," *Expert Systems with Applications*, vol. 249, p. 123625, 2024.

[5] V. Atalay Aydın, "Comparison of cnn-based methods for yoga pose classification," *Turkish Journal of Engineering*, vol. 8, no. 1, p. 65–75, 2024.

[6] H.-P. Huynh, D.-H. Nguyen, T. T.-L. Nguyen, M.-V. Nguyen, and Q.-T. Nguyen, "Robust classification of similar yoga poses based on grouping and hierarchical inference," in *2023 International Conference on Multimedia Analysis and Pattern Recognition (MAPR)*, pp. 1–6, IEEE, 2023.

[7] S. Wadhwa, A. Garg, and G. Munjal, "Yoga posture analysis using deep learning," in *2024 11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, pp. 1–6, 2024.

[8] J. Z. Tan, C. P. Lee, K. M. Lim, and J. Y. Lim, "Yoga pose estimation with machine learning," in *2023 11th International Conference on Information and Communication Technology (ICoICT)*, pp. 260–265, IEEE, 2023.

[9] O. Parekh, S. Kondaskar, Y. Shaikh, N. Shende, and N. Deotale, "Yogmaster: Detection and correction of yoga postures using augmented reality," in *2023 3rd International Conference on Pervasive Computing and Social Networking (ICPCSN)*, pp. 103–108, IEEE, 2023.

[10] T. W. Jian, N. A. J. Sufri, and M. A. As'Ari, "Yoga pose recognition based on convolutional neural networks," in *2023 3rd International Conference on Intelligent Cybernetics Technology & Applications (ICICyTA)*, pp. 511–516, IEEE, 2023.

[11] C. Vishwanatha, S. Dhareshwar, *et al.*, "Prediction of yoga pose from youtube dataset using skeleton feature extraction based isdl model," in *2023 International Conference on Evolutionary Algorithms and Soft Computing Techniques (EASCT)*, pp. 1–8, IEEE, 2023.

[12] N. Vallabhaneni and P. Prabhavathy, “Segmentation quality assessment network-based object detection and optimized cnn with transfer learning for yoga pose classification for health care,” *Soft Computing*, pp. 1–23, 2023.

[13] A. K. Rajendran and S. C. Sethuraman, “Denseposecompare: A comparative study of densenet models in yoga pose classification,” in *2024 International Conference on Cognitive Robotics and Intelligent Systems (ICC-ROBINS)*, pp. 471–480, IEEE, 2024.

[14] P. Sarma and S. I. Singh, “Handcrafted feature assisted light-weight encoder decoder based classifier for yoga posture recognition,” in *2024 2nd International Conference on Device Intelligence, Computing and Communication Technologies (DICCT)*, pp. 681–686, IEEE, 2024.

[15] N. Vallabhaneni, P. Panneer, and V. M. Sundaram, “Transfer learning-based method for classifying yoga poses using deep convolutional neural networks,” in *2023 Innovations in Power and Advanced Computing Technologies (i-PACT)*, pp. 1–6, IEEE, 2023.

[16] K. Aarthy and A. A. Nithys, “Yoga pose detection and identification using mediapipe and openpose model,” in *2023 International Conference on Computer Science and Emerging Technologies (CSET)*, pp. 1–7, IEEE, 2023.

[17] U. Kulkarni, Y. Diwan, P. Hegde, P. Mutnale, B. Jain, S. Meena, and S. Gurlahosur, “Yoga pose detection using long-term recurrent convolutional network,” in *2023 IEEE 8th International Conference for Convergence in Technology (I2CT)*, pp. 1–7, IEEE, 2023.

[18] R. R. Subramanian and V. Govindaraj, “Harnet: design and evaluation of a deep genetic algorithm for recognizing yoga postures,” *Signal, Image and Video Processing*, pp. 1–12, 2024.

[19] S. Saurav, P. Gidde, and S. Singh, “Exploration of deep learning architectures for real-time yoga pose recognition,” *Multimedia Tools and Applications*, pp. 1–43, 2024.

[20] A. S. Talaat, “Novel deep learning models for yoga pose estimator,” *SN Applied Sciences*, vol. 5, no. 12, p. 341, 2023.

[21] S. Imran, Z. Sadman, A. Islam, and D. Z. Karim, “Enhanced yoga posture detection using deep learning and ensemble modeling,” in *2023 3rd International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*, pp. 1–6, IEEE, 2023.

[22] H. C. Saini, “ismartyog: A real time yoga pose recognition and correction feedback model using deep learning for smart healthcare,” in *2023 International Conference on Smart Systems for applications in Electrical Sciences (ICSSES)*, pp. 1–6, IEEE, 2023.

[23] M. G. Fard, G. A. Montazer, and D. Giveki, “A novel fuzzy logic-based method for modeling and recognizing yoga pose,” in *2023 9th International Conference on Web Research (ICWR)*, pp. 1–6, IEEE, 2023.

[24] D. Debalaxmi, V. Ranga, and D. K. Vishwakarma, “An analytical comparison of deep learning frameworks for 2d image-based hatha-yoga pose identification,” in *2024 IEEE International Conference for Women in Innovation, Technology & Entrepreneurship (ICWITE)*, pp. 271–276, IEEE, 2024.

[25] A. Bera, M. Nasipuri, O. Krejcar, and D. Bhattacharjee, “Fine-grained sports, yoga, and dance postures recognition: A benchmark analysis,” *IEEE Transactions on Instrumentation and Measurement*, 2023.