YogaFeed: Real time pose assessment and improvement

YogaVision
Diptesh Saha, Parimal Bera & Samapan Kar

August 27, 2024

Abstract

In this project, we propose a comprehensive computer vision-based approach to assess and improve yoga poses in real time. The core objective is to quantitatively evaluate a practitioner's yoga pose by comparing it with a standard pose, providing a precise grade or feedback. The system extracts human body skeleton keypoints from the input image and encodes these features using contrastive learning techniques. Specifically, we utilize both coarse and fine triplet examples to enhance pose feature encoding, ensuring accurate comparison and assessment.

To address the challenge of limited diversity in existing pose datasets, our work introduces the Yoga-82 dataset, which offers fine-grained hierarchical labels for 82 distinct yoga poses. This large-scale dataset is designed to tackle complex poses that require detailed annotations, supporting the classification of poses across multiple levels of difficulty.

Furthermore, the system incorporates a self-training component aimed at guiding practitioners to perform yoga poses correctly, minimizing the risk of injury. By analyzing posture from different angles and providing visualized feedback for rectification, the system assists in the correct alignment and execution of poses.

Additionally, we introduce the YogaTube dataset, a video-based benchmark for yoga action recognition, which includes over 5,000 videos categorized into 82 yoga poses. Our model leverages a three-stream architecture, combining pose estimation, optical flow, and raw video features for enhanced accuracy in pose recognition.

Through extensive experiments on benchmark datasets, our approach demonstrates superior performance, contributing to the development of advanced computer-assisted yoga training systems.

1 Introduction

1.1 Description

Design a personalized yoga pose grading system using deep learning to adapt to individual physical capabilities, enhancing pose accuracy and user experience.

1.2 What?

This project aims to develop a computer vision-based yoga pose grading system that leverages deep learning techniques to evaluate and improve yoga postures by comparing them against standard poses. The system will utilize contrastive learning, incorporating both coarse and fine triplet examples, to enhance the encoding of pose features, ensuring accurate assessment. By providing precise grading and actionable feedback, the system guides users to correct their poses, reducing injury risks and offering visual cues for better alignment. The effectiveness of this approach will be validated through extensive experiments on benchmark datasets, advancing the capabilities of computer-assisted yoga training systems.

1.3 Why?

This project is crucial as it promotes correct yoga practice, enhancing users' physical and mental well-being by providing real-time feedback and minimizing the risk of injuries. By combining computer vision and contrastive learning in a practical, real-world application, it allows exploration of cutting-edge deep learning techniques within a growing field. Yoga pose assessment presents a complex challenge due to the lack of diverse datasets; this project addresses this gap by creating a specialized dataset and employing advanced pose recognition methods. Moreover, developing a system that supports self-guided yoga training empowers individuals without access to professional instructors, broadening the accessibility of high-quality yoga practice. The project's potential extends beyond yoga into fitness, rehabilitation, and sports training applications, making it a valuable addition to your portfolio and a stepping stone for future innovations.

1.4 How?

- **Skeleton Feature Extraction:**
  - Utilize pose estimation models (e.g., OpenPose or Mediapipe) to extract key points from yoga poses.
  - These key points will form the foundation for pose assessment.
- **Contrastive Learning:**
  - Implement contrastive learning with a triplet loss function.

- Use coarse triplet examples (different categories) and fine triplet examples (same category, different qualities) to enhance feature encoding.

- **CNN-Based Pose Classification:**
  - Employ CNN architectures (e.g., ResNet or DenseNet) to classify poses based on extracted skeleton features.
  - Incorporate hierarchical labels for detailed classification.

- **Deviations Analysis and Corrective Instructions:**
  - Design a system to analyze deviations from standard poses by comparing them with benchmarks.
  - Integrate front and side view analysis for improved accuracy.
  - Provide corrective instructions to guide users toward proper pose execution.

- **Incorporation of 3D Pose Estimation:**
  - Explore the integration of 3D pose estimation techniques for a comprehensive analysis of posture and movement.

- **Real-Time Feedback System:**
  - Integrate a real-time feedback system that offers immediate corrective guidance during practice.
  - Utilize continuous data flow from pose estimation models for dynamic feedback.

- **Personalized Pose Difficulty Calibration:**
  - Include a feature for personalized pose difficulty calibration, adapting the system to the user's capabilities.

- **Integration with Wearable Sensors:**
  - Explore the integration of wearable sensors to provide additional data points (e.g., joint angles, movement velocity).
  - Use this data to refine pose grading and feedback mechanisms.

- **Pose Sequence Grading:**
  - Develop a feature to grade sequences of poses, assessing transitions, alignment, and flow.
  - Offer feedback on overall practice flow, crucial for advanced routines.

1.5 Literature Survey

The paper **"A Computer Vision-Based Yoga Pose Grading Approach Using Contrastive Skeleton Feature Representations** introduces an innovative approach by integrating contrastive learning with CNN-based pose classification, aiming to improve the precision of yoga pose grading. This project builds on these advancements, leveraging real-time feedback and pose estimation techniques to create a robust self-guided yoga training system.

2 Proposed methodology

- **Data Collection and Preparation:**
  - Gather and augment a diverse yoga pose dataset with front and side views to ensure robust training.
- **Skeleton Feature Extraction:**
  - Use pre-trained pose estimation models (like OpenPose) to extract and normalize key skeletal points from the dataset.:
- **Contrastive Learning:**
  - Implement a triplet loss-based contrastive learning framework with coarse and fine examples to refine pose feature encoding.
- **CNN-Based Classification:**
  - Develop a CNN model (e.g., ResNet) to classify poses, using hierarchical labeling to capture both overall and detailed pose accuracy.
- **System Development:**
  - Create a real-time feedback system that provides corrective guidance based on pose deviations, incorporating 3D pose estimation for enhanced accuracy.
- **Evaluation:**
  - Assess system performance using standard metrics and user studies, iterating on feedback to refine the model.
- **Deployment:**
  - Design a user-friendly interface and integrate the system with mobile devices or wearables for practical use.

- **Future Work:**
  - Extend the system to grade pose sequences and explore applications in other domains like rehabilitation and sports training.

3 Work plan and time line

Task	Start Date	End Date
Dataset Preparation	31-08-2024	13-08-2024
Skeleton Feature Extraction	14-08-2024	28-08-2024
Model Developement, Contrastive Learning & CNN	29-08-2024	30-09-2024
Model Evaluation & Tuning	01-10-2024	4-10-2024
Additional Work	5-10-2024	19-10-2024
Testing & Validation	20-10-2024	25-10-2024
Final Deployment & Documentation	26-10-2024	10-11-2024

4 Work plan division in your group

- Dataset Preparation : Samapan Kar, Diptesh Saha
- Skeleton Feature Extraction : Diptesh Saha, Parimal Bera
- Model Developement, Contrastive Learning : Everyone
- Model Evaluation & Tuning : Diptesh Saha
- Additional Work : Everyone
- Testing & Validation : Everyone
- Deployment & Documentation : Parimal Bera

5 References

Some references related to our project below:

References

[1] Hua-Tsung Chen, Yu-Zhen He, and Chun-Chieh Hsu. Computer-assisted yoga training system. *Multimedia Tools and Applications*, 77:23969–23991, 2018.

[2] Geetanjali Kale, Varsha Patil, and Mousami Munot. A novel and intelligent vision-based tutor for yogāsana: e-yogaguru. *Machine Vision and Applications*, 32, 01 2021.

[3] Manisha Verma, Sudhakar Kumawat, Yuta Nakashima, and Shanmuganathan Raman. Yoga-82: A new dataset for fine-grained classification of human poses. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 4472–4479, 2020.

[4] Yubin Wu, Qianqian Lin, Mingrun Yang, Jing Liu, Jing Tian, Dev Kapil, and Laura Vanderbloemen. A computer vision-based yoga pose grading approach using contrastive skeleton feature representations. *Healthcare*, 10(1), 2022.

[5] Santosh Kumar Yadav, Amitojdeep Singh, Abhishek Gupta, and Jagdish Lal Raheja. Real-time yoga recognition using deep learning. *Neural computing and applications*, 31:9349–9361, 2019.

[6] Santosh Kumar Yadav, Guntaas Singh, Manisha Verma, Kamlesh Tiwari, Hari Mohan Pandey, Shaik Ali Akbar, and Peter Corcoran. Yogatube: a video benchmark for yoga action recognition. In *2022 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2022.