

CSCE 5222 Feature Engineering
Project Report

Group 5

Member: Parth R Bhanderi
Darshit J Kajavadra

1. Problem statement

Our project goal is to detect Squat position in the video.

The goal of this research is to create a reliable image processing system that can recognize different positions such as standing and squatting from a video frame. The variation in the background, camera angle, and form of the human characters across multiple videos makes this work extremely difficult. The methodology used in this project has the potential to be applied in a wide range of industries, such as analyzing the position of an athlete in sports.

Cite: [IEEE](#)

2. Data used

There are three mp4 video files with the length of 2 to 3 seconds which show a person doing squats. All the video files have different backgrounds and different camera angles.

3. Method

- 1. Video Loading and Frame Capture:** Opening and reading frames from a video file is achieved by initializing a connection to the video file (`cv2.VideoCapture(video_path)`), reading and retrieving the next frame (`cap.read()`), and converting frames from color to grayscale using `cv2.cvtColor(frame, cv2.COLOR_RGB2GRAY)`
- 2. Structural Similarity Index (SSI) Calculation:** Assessing the similarity between frames is achieved by measuring the structural similarity index between two images using `ssim(image1, image2)`.
- 3. Frame Comparison for Similarity:** Identifying the frame with the lowest similarity to a reference frame involves looping through frames, comparing their structural similarity, and updating variables based on the best match.
- 4. Contour Analysis:** Identifying significant shapes or objects in frames is achieved through image differencing, blurring, and thresholding to highlight changes, finding contours,

filtering contours based on area and bounding rectangle characteristics, and sorting contours to identify the most relevant one.

- 5. Image Masking and Bitwise Operations:** Isolating specific regions of interest in frames involves creating a mask to define those regions, drawing contours on the mask to highlight specific areas, and applying bitwise operations to combine the original frame with the mask.

4. Evaluation

Structural Similarity Index (SSIM)

Name	Size	Type	Accuracy
squat_1677_segmented	480 x 362	JPG	0.70
squat_1667_segmented	480 x 360	JPG	0.89
squat_1668_segmented	480 x 270	JPG	0.76



Figure 1.1



Figure 1.2

Figure 1.1 is the ground truth for video 1667.mp4 and figure 1.2 is the obtained result squat_1667_segmented.jpg.



Figure 2.1

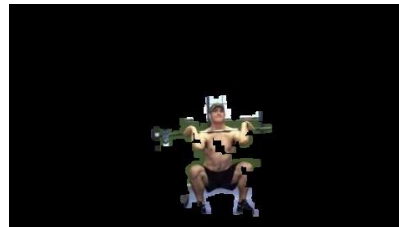


Figure 2.3

Figure 2.1 is the ground truth for video 1668.mp4 and figure 2.2 is the obtained result squat_1668_segmented.jpg.



Figure 3.1

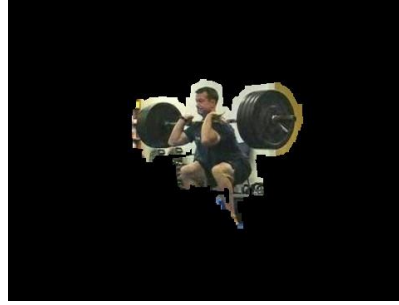


Figure 3.4

Figure 3.1 is the ground truth for video 1677.mp4 and figure 3.2 is the obtained result squat_1677_segmented.jpg.

5. Timeline

Phase	Tasks	Status	Date
1. Project Planning and Setup	Define objectives and requirements	Completed	09/05/2023
2. Data Preparation	Collect or select video data Extract frames from the video	Completed	09/05/2023
3. Frame Comparison and Similarity Analysis	Implement frame comparison using SSIM Identify the frame with the highest similarity	Completed	11/01/2023
4. Contour Analysis	Implement contour detection and analysis	Completed	11/07/2023
5. Image Masking and Processing	Implement image masking and bitwise operations	Completed	11/10/2023
6. Evaluation	Evaluate results using SSIM and other metrics	Completed	11/14/2023
7. Fine-Tuning	Fine-tune parameters for better performance	Pending	12/05/2023
8. Documentation and Code Cleanup	Document code and add comments Clean up code and organize project files	Pending	12/05/2023