Trackings exercise reps using a webcam :

1. Datasets

a. Kinetics Dataset:
Contains around 650,000 clips covering 400–700 human actions. It's widely used for action recognition tasks. The dataset includes various sports and physical activities, which could help pre-train models for recognizing common exercise movements.

b. NTU RGB+D Dataset:
Includes RGB, depth, skeleton data, and infrared video for 60 actions, with 56,000 sample. Contains some exercises like hand-clapping, which could be fine-tuned for rep-counting models.

c. HMDB5:
A human motion database with 6,849 clips divided into 51 action categories, including some related to physical exercises. Can be used to pre-train models for human action recognition.

d. UCF101: This dataset consists of 13,320 videos categorized into 101 action classes. covers more diverse actions, some of them overlap with fitness-related activities.

e. Pose-based Datasets (like COCO or MPII):
These datasets focus on human pose estimation, providing annotations of key body joints. helpful in tracking key joint movements, assisting in the detection of rep-like motions in exercises.

f. InfiniteRep: (https://paperswithcode.com/dataset/infiniterep)
open-source dataset for fitness and physical therapy (PT) applications. It includes 1k videos of diverse avatars performing multiple repetitions of common exercises.
100 videos per exercise, spanning 5 to 10 repetitions each (1,000 videos total)
7 unique indoor scenes
Realistic environmental occlusion (+ corresponding labels)
Diverse lighting conditions
Varied body shape, skin tones, and clothing
Rich annotations for 2D and 3D supervision

g. **Workout/Exercises Video:**
very large sets of gym exercises videos
https://www.kaggle.com/datasets/hasyimabdillah/workoutfitness-video/data
Mostly sourced from YouTube

h : MEx - Multi-modal Exercise Dataset for Human Activity Recognition
- https://data.mendeley.com/datasets/p89fwbzmkd/2
- The MEx Multi-modal Exercise dataset contains data of 7 different physiotherapy exercises, performed by 30 subjects recorded with 2 accelerometers, a pressure mat and a depth camera.

2. Models

a. Optical Flow-Based CNNs

Optical flow captures the motion between two frames, helping models identify movement patterns. Two-Stream CNNs, where one stream takes optical flow input and another takes raw image input, have been successful in action recognition.
Optical flow can be leveraged to detect movements during exercises such as squats or push-ups, while the CNN will help to extract spatiotemporal features to count reps.

b. 3D Convolutional Neural Networks (3D CNNs)

3D CNNs extend 2D convolutions to the time dimension, making them well-suited for video-based tasks.
Can be used for extracting features from video sequences for recognizing different phases of an exercise (e.g., the up and down motion of a squat).

c. Pose Estimation Models (OpenPose, BlazePose)

Pose estimation models detect human key points and track skeletal movements, which can be used for tracking exercise form and repetitions.
For counting reps, pose estimation could offer precise tracking of joint movements during exercises like lunges or bicep curls.

d. Recurrent Neural Networks (RNNs) and Long Short-Term Memory Networks (LSTMs)

These models can be combined with CNNs to capture temporal dependencies in videos. RNNs or LSTMs can be used to identify repetitive motion patterns, such as the repetitive bending of a push-up or squat.

e. Transformers (Video Transformer Networks)

Transformers have recently been applied to video tasks, utilizing attention mechanisms to model temporal dependencies in video sequences.
You can combine optical flow with a Transformer network to accurately detect the beginning and end of each exercise repetition.

reason of my model selection:

blog - https://towardsdatascience.com/vision-based-rep-counting-in-the-wild-cb9a4d1bdb7e
I read this blog and realized it better to use optical flow based method with CNN becuase it is prone to non-generic exercise and it can cover larger corpus of exercise.

3. Evaluation Metrics

a. Accuracy

Measures the percentage of correctly counted reps versus the ground truth. Basic metric to assess how well the system counts the number of repetitions.

b. Precision and Recall

Precision measures how many of the detected reps are correct, while recall measures how many true reps were detected. Important for understanding the model's performance in detecting reps while minimizing false positives and negatives.

c. F1 Score:

The harmonic mean of precision and recall, which balances false positives and false negatives. Provides a balanced view of the model's rep-counting ability.

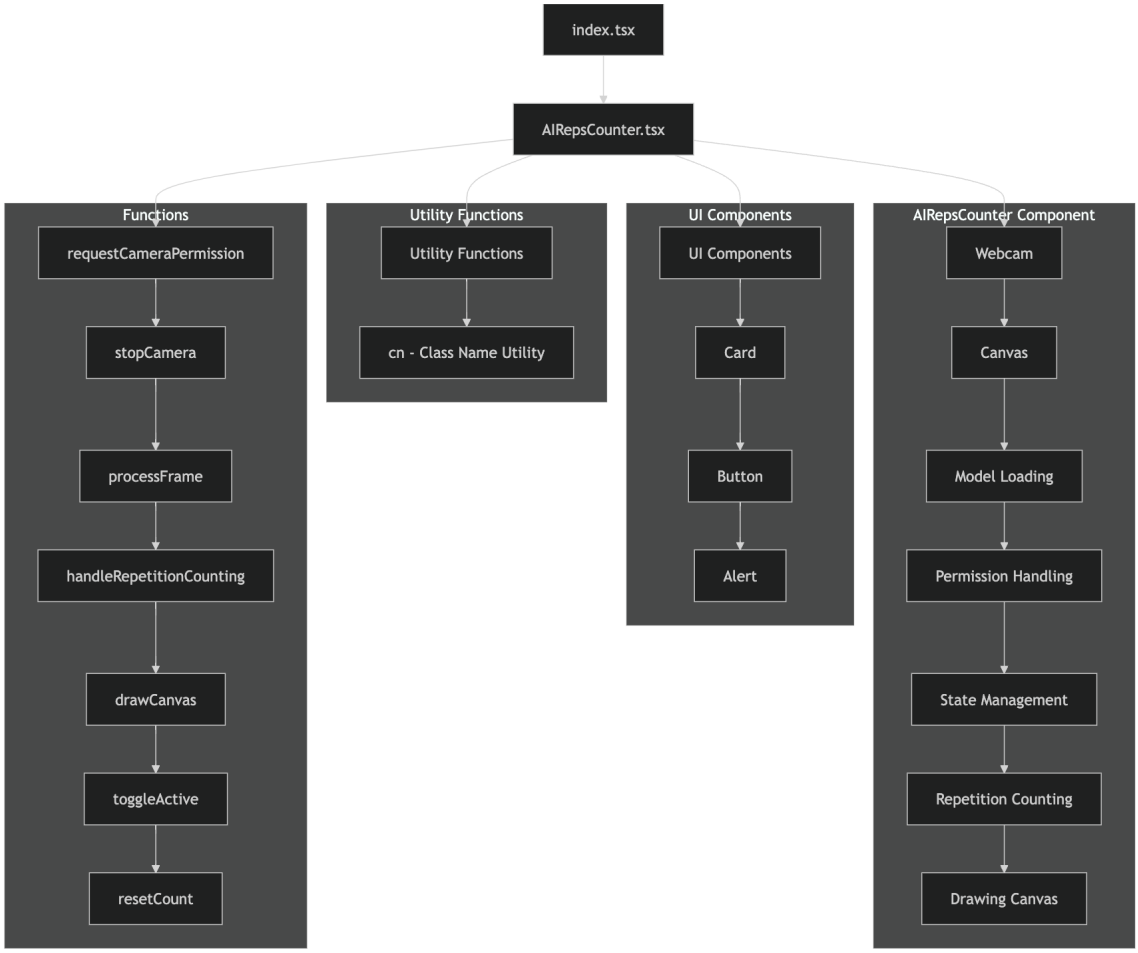
d. IoU (Intersection over Union)

Measures the overlap between the predicted and actual bounding boxes or key points in pose estimation tasks. For pose-based methods, this metric can help evaluate how well the model tracks the user's form during exercises.

e. MAE (Mean Absolute Error)
Measures the difference between the true rep count and the predicted count. Provides insight into the average number of reps the model gets wrong, either by over- or under-counting.

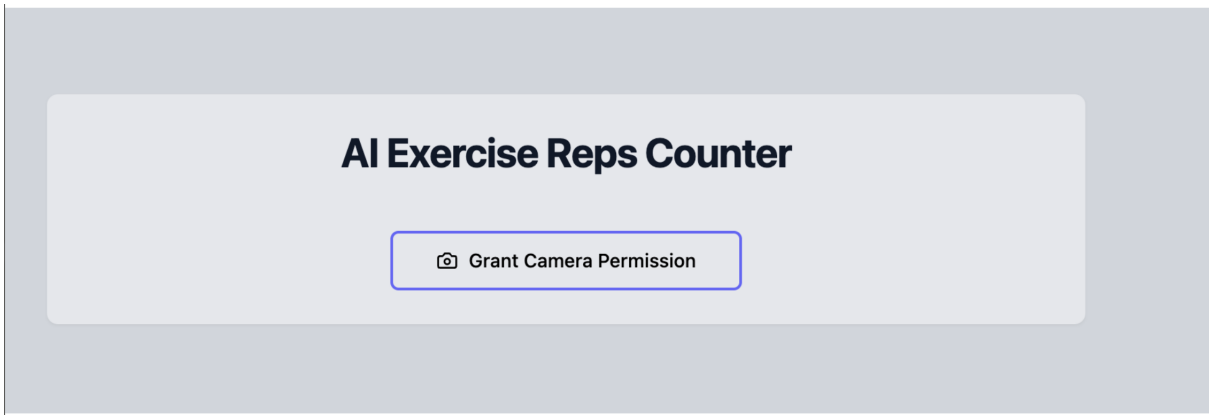
f. Time-to-Detect
Measures how quickly the model can recognize the beginning and end of a repetition. Important for real-time applications where lag in rep detection can lead to poor user experience.

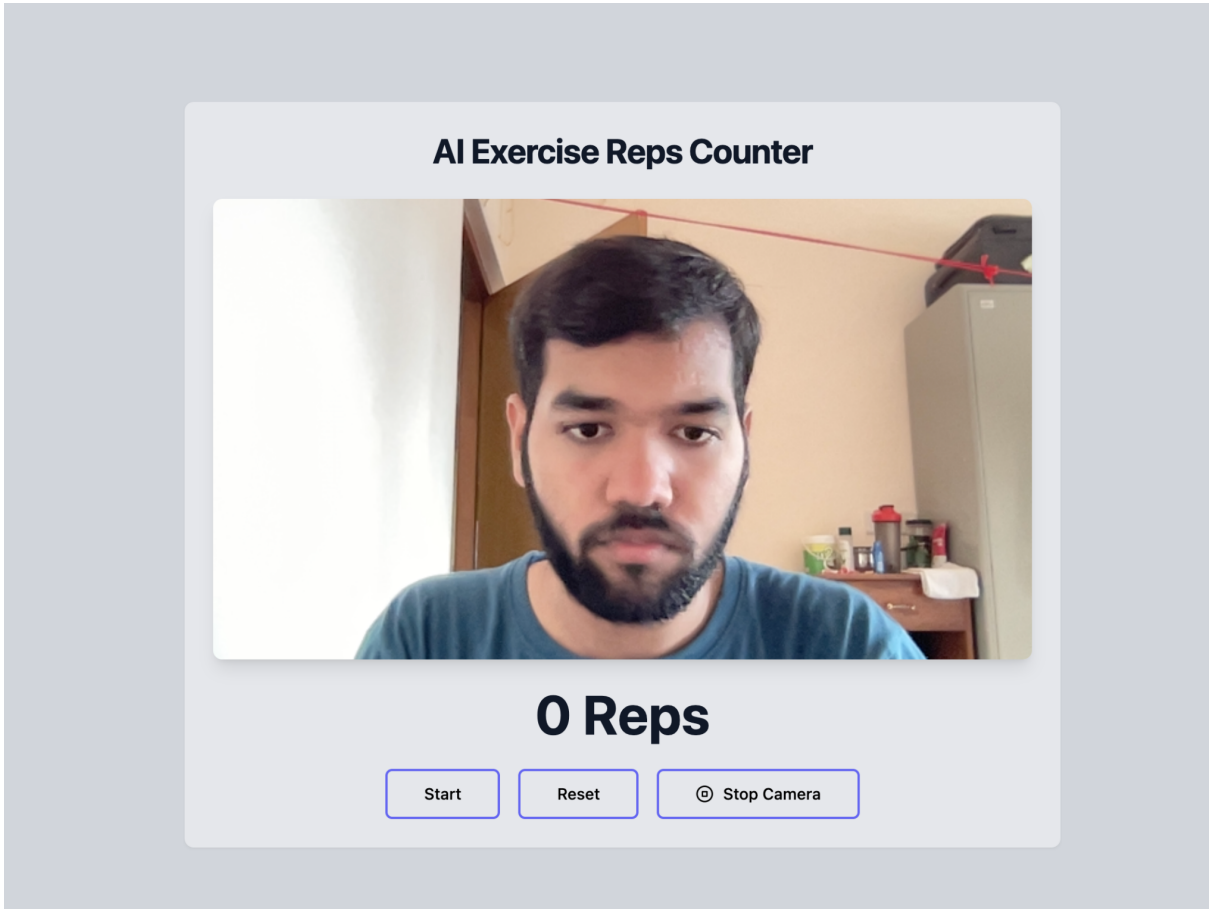
flow diagram of Web:



demo -

UI

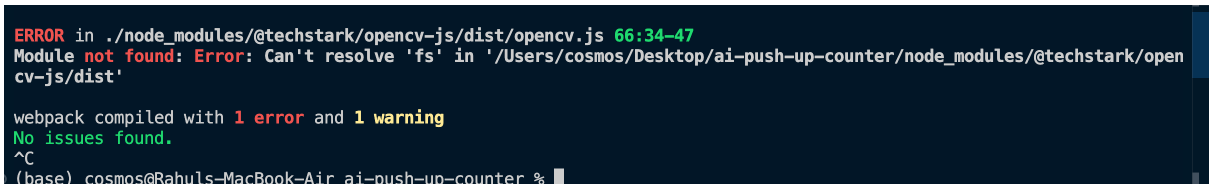

AI Exercise Reps Counter
Grant Camera Permission


AI Exercise Reps Counter
0 Reps
Start
Reset
Stop Camera

video :

file : demo_video.mp4

error -



I am using opencv-js library in that it requires the fs(file) module which is not supported by browser (client side) .

I try to set false to it in fallback resolve of webpack but still error is there.

thing are completed except this error.