Title: A Proposed Algorithm for Video Summarization for Driver Behaviour Analysis

This algorithm combines keyframe selection, event detection, and RNN-based sequence modeling to create effective video summaries.

Goals and objects:

Video summarization is crucial for driver behavior analysis for several reasons:

- ✓ **Efficiency**
- **Data Reduction:** Video summarization reduces the amount of data that needs to be processed, stored, and analyzed. This is particularly important when dealing with long hours of driving footage.
- **Quick Review:** It allows for quick review and identification of key events, such as sudden braking, lane changes, or potential collisions, without having to watch the entire video.
- ✓ **Focus on Relevant Events**
- **Highlighting Critical Moments:** Summarization techniques can highlight critical moments that are indicative of driver behavior, such as instances of distraction, aggressive driving, or compliance with traffic rules.
- **Event Detection:** It helps in detecting specific events that are relevant for behavior analysis, making it easier to focus on significant occurrences rather than sifting through irrelevant footage.
- ✓ **Improved Analysis**
- **Behavior Patterns:** By summarizing videos, analysts can more easily identify patterns in driver behavior over time, such as frequent speeding or consistent adherence to safe driving practices.
- **Training and Feedback:** Summarized videos can be used for training purposes, providing drivers with feedback on their performance and areas for improvements.
- ✓ **Resource Optimization**
- **Cost-Effective:** Reducing the amount of data to be processed and stored can lead to significant cost savings in terms of storage and computational resources.
- **Scalability**: It makes the analysis process more scalable, allowing for the monitoring of multiple drivers or vehicles simultaneously without overwhelming the system.
- ✓ **Enhanced Safety**
- **Proactive Measures:** By quickly identifying risky behaviors, interventions can be implemented to improve driver safety and reduce the likelihood of accidents.
- **Regulatory Compliance:** Ensures that drivers are adhering to safety regulations and standards, which is crucial for fleet management and regulatory compliance.

In summary, video summarization streamlines the process of analyzing driver behavior, making it more efficient, focused, and effective. This ultimately leads to better insights, improved safety, and optimized resource usage. We proposed an algorithm that combines keyframe selection, event detection, and RNN-based sequence modeling to create effective video summaries.

Novelty Aspects:

- We combine keyframe selection, event detection, and RNN-based sequence modeling in a single framework is a comprehensive approach that leverages the strengths of each technique.

- The focus on real-time summarization of driver behavior is crucial for applications in autonomous vehicles and IoV, providing immediate insights and enhancing safety.
- We apply these techniques specifically to IoV and autonomous vehicles is a relatively unexplored area, offering significant potential for innovation and practical impact.
- The proposed method aims to improve both the safety and efficiency of autonomous driving systems by providing detailed and actionable summaries of driver behavior.
- Further more, we can have Integrating video data with other modalities (e.g., sensor data, audio) to create comprehensive summaries is a cutting-edge approach that can provide a more holistic understanding of driving behavior.

Proposed Algorithm

1. **Keyframe Selection**

Keyframe selection is used to identify and retain the most informative frames from a sequence. These keyframes are essential for tasks such as video summarization, visual localization, and mapping in robotics. By selecting keyframes, we can reduce the amount of data to process while preserving the essential information.

- ✓ **Frame Extraction:** we Extract frames from the video at regular intervals. Capture a sequence of frames from a video or a series of images
- ✓ **Feature Extraction:** we use a pre-trained Convolutional Neural Network (CNN) like ResNet or VGG to extract features from each frame. We extract features from each frame. These features could be edges, corners, or other significant points in the image.
- ✓ **Similarity Measurement:** We measure the similarity between consecutive frames using metrics like Euclidean distance, cosine similarity, or other relevant measures.
- ✓ **Clustering**: We apply clustering algorithms (e.g., K-means) on the extracted features to group similar frames.
- ✓ **Thresholding:** We set a threshold to determine when a frame is significantly different from the previous ones. If the difference exceeds the threshold, the frame is considered a keyframe.
- ✓ **Keyframe Identification:** Select the centroid of each cluster as the keyframe, representing the most significant frames in the video. We select frames that meet the criteria as keyframes. These frames should represent the most significant changes or events in the sequence.

2. Event Detection and Classification

- ✓ **Object Detection:** We use object detection models like YOLO or Faster R-CNN to detect and classify objects in each keyframe.
- ✓ **Action Recognition:** Apply action recognition models (e.g., C3D, I3D) to identify actions or events occurring in the keyframes.
- ✓ **Event Classification:** Classify detected events using a pre-trained classifier (e.g., SVM, Random Forest) based on the features extracted from the keyframes.

3. RNN-Based Sequence Modeling

Sequence Preparation: Arrange the keyframes in chronological order to form a sequence.

Feature Encoding: Encode the features of each keyframe using a CNN and pass them to an RNN (e.g., LSTM or GRU).

Temporal Analysis: Use the RNN to capture temporal dependencies and generate a sequence representation.

4. Summary Generation: Generate a summary by selecting key events and actions identified by the RNN, ensuring the summary is coherent and concise.

An example method Workflow for simulatiom

Input: Video sequence

Output: Summarized video highlighting key events and actions.

Steps:

1. Extract frames from the video.

2. Extract features from each frame using a CNN.

3. Cluster frames and select keyframes.

4. Detect objects and recognize actions in keyframes.

5. Classify events based on detected objects and actions.

6. Encode keyframe features and pass them to an RNN.

7. Analyze temporal dependencies using the RNN.

8. Generate a summary by selecting key events and actions

Dataset:

- ✓ **Drive&Act:** For driver behavior analysis.
- ✓ **BDD100K:** For diverse driving scenarios.
- ✓ **DAD (Driver Attention Dataset):** For attention and distraction analysis.

Simulation tools for traffic management:

VISSIM: A microscopic traffic simulation software that allows for detailed modeling of traffic flow and driver behavior.

CARLA: An open-source simulator for autonomous driving research.

SUMO (Simulation of Urban MObility): An open-source, highly portable, microscopic and continuous traffic simulation package designed to handle large road networks.

Implementation tool:

TensorFlow, Scikit-learn, PyTorch in **Python**.

OpenCV: An open-source computer vision library that provides tools for image and video processing.

Evaluation Parameters

Precision, Recall, and F1-Score, BLEU and ROUGE Scores.

Github codes for more information that must be considered:

https://github.com/yashkolli/Video-Summarization-Using-Attention

https://github.com/topics/video-summarization?o=desc&s=updated

References:

1. Kumar, Raman, and Anuj Jain. "Driving behavior analysis and classification by vehicle OBD data using machine learning." *The Journal of Supercomputing* 79, no. 16 (2023): 18800-18819.
2. Peppes, Nikolaos, Theodoros Alexakis, Evgenia Adamopoulou, and Konstantinos Demestichas. "Driving behaviour analysis using machine and deep learning methods for continuous streams of vehicular data." *Sensors* 21, no. 14 (2021): 4704.
3. De Oliveira, Wictor Gomes, Pedro Pedrosa Rebouças Filho, and Elias Teodoro da Silva Junior. "Driver behavior analysis: Abnormal driving detection using mlp classifier applied to outdoor camera images." In *International Conference on Intelligent Systems Design and Applications*, pp. 1142-1152. Cham: Springer International Publishing, 2021.
4. Carmona, Juan, Fernando García, David Martín, Arturo de la Escalera, and José María Armingol. "Data fusion for driver behaviour analysis." *Sensors* 15, no. 10 (2015): 25968-25991.
5. Peppes, Nikolaos, Theodoros Alexakis, Evgenia Adamopoulou, and Konstantinos Demestichas. "Driving behaviour analysis using machine and deep learning methods for continuous streams of vehicular data." *Sensors* 21, no. 14 (2021): 4704.
6. Qu, Fangming, Nolan Dang, Borko Furht, and Mehrdad Nojoumian. "Comprehensive study of driver behavior monitoring systems using computer vision and machine learning techniques." *Journal of Big Data* 11, no. 1 (2024): 32.
7. Bouhsissin, Soukaina, Nawal Sael, and Faouzia Benabbou. "Driver behavior classification: a systematic literature review." *IEEE Access* 11 (2023): 14128-14153.
8. Sabha, Ambreen, and Arvind Selwal. "Towards machine vision-based video analysis in smart cities: a survey, framework, applications and open issues." *Multimedia Tools and Applications* 83, no. 22 (2024): 62107-62158.
9. Razi, Abolfazl, Xiwen Chen, Huayu Li, Hao Wang, Brendan Russo, Yan Chen, and Hongbin Yu. "Deep learning serves traffic safety analysis: A forward-looking review." *IET Intelligent Transport Systems* 17, no. 1 (2023): 22-71.
10. Sonth, Akash, Abhijit Sarkar, Hirva Bhagat, and Lynn Abbott. "Explainable Driver Activity Recognition Using Video Transformer in Highly Automated Vehicle." In *2023 IEEE Intelligent Vehicles Symposium (IV)*, pp. 1-8. IEEE, 2023.
11. Ingle, Palash Yuvraj, and Young-Gab Kim. "Multiview abnormal video synopsis in real-time." *Engineering Applications of Artificial Intelligence* 123 (2023): 106406.