# An Implementation of Video Sementaition and Classifcation based on Video Features Extraction and Machine Learning

**Jiaming Song** JIAMING.TSONG@GMAIL.COM
Computer Science and Technology Department, Tsinghua University

**Gerald Yankai Zhang** ZYK12@MAILS.TSINGHUA.EDU.CN
Computer Science and Technology Department, Tsinghua University

## Abstract

## 1. Introduction

In these days on the Internet, we have witnessed the continue increase of available network bandwidth. The network is capable to deal with video streams with higher and higher bitrates. Thus we are not surprised to see that video data are taking larger part of total network data. In compare with the fast development of network capability and the convenience brought by the feature of uploading self-made videos, the lack of detailed video description information is still an open problem that urges to be solved. For instance, most of the users are only interested in some specific parts of a video stream. Usually, users do not like the repeating opening and ending of a series. It is impossible to tag them by human hands. If there is a tool that can tag them automatically once the videos are uploaded, the experience of video watching will be greatly increased.

On the other hand, thanks to the flourish of machine learning, many problems which are once believed unsolvable are neatly settled. For instance, the computers are able to classify video and audio information into multiple genres with an incredible precision. But the true power of deep learning is still waiting to be developed.

In this application condition and theoretical background, we believe it is a right time to present our research topic, a video segmentation and classification system based on video features extraction and deep learning. It is a system that can segment a long video into individual programs and classify these programs into various genres. For each program, users can choose to watch some selected part of programs and ignore some of them, e.g. opening and ending of a program.

Our work is with the following contributions: 1) A set of video and audio features that can be used to solve programs classification and video segmentation problems in the future. 2) We present a novel application scenario for machine learning. 3) A fully functioned video processing and playback system that can be used in further study purpose.

The challenges that we are facing are listed as follows: 1) To combine multimedia information, we need to handle audio and video information properly at the same time. 2) To increase the precision of segmentation and classification, we need to try various ways in order to enhance the performance of our system. 3) To deal with the large amount of information, we need to find the balance point between running time and integrity of information.

The rest of this article will be arranged as follows: In Section 2, we introduce our methodology with explanations of features and algorithms we use. In Section 3, we briefly demonstrate our system. With dataset information presented, we describe our experiment environment and estimate our experiment in detail. In Section 4, we propose our plan for future work and time table.

## 2. Methodology

## 3. Estimation

### 3.1. Dataset

### 3.2. Demo System

### 3.3. Experiment

## 4. Future work

## References