

# Scene Separation & Data Selection: Temporal Segmentation Algorithm for Real-time Video Stream Analysis

Yuelin Xin<sup>12</sup>   Zihan Zhou<sup>1</sup>   Yuxuan Xia<sup>1</sup>

<sup>1</sup>SWJTU-Leeds Joint School, CS  
Southwest Jiaotong University

<sup>2</sup>School of Computing  
University of Leeds

Spatio-Temporal Reasoning and Learning, 2022

# Outline

## Introduction

- The problem

- Our motivation

# Introduction

- ▶ The problem (Background & What we want to achieve)
- ▶ Our motivation (Why not neural networks?)

## Remark

**Scene separation** is a problem in which we want to separate a video stream into different scenes. **A scene** is defined as a group of similar-looking frames that are temporally adjacent to each other.

## The problem

- ▶ **Background:** real-time video stream interpretation, including video semantics / video accessibility / surveillance footage auto-interpretation, etc.
- ▶ **What's missing:** algorithms do not see video as a continuous stream of images, but as discrete frames.



Figure 1: Video semantics.

# The problem

## Our motivation

Let  $X_1, X_2, \dots, X_n$  be a sequence of independent and identically distributed random variables with  $E[X_i] = \mu$  and  $\text{Var}[X_i] = \sigma^2 < \infty$ , and let

$$S_n = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{1}{n} \sum_i^n X_i$$

denote their mean. Then as  $n$  approaches infinity, the random variables  $\sqrt{n}(S_n - \mu)$  converge in distribution to a normal  $\mathcal{N}(0, \sigma^2)$ .