# A Deep Reinforcement Learning Framework for Identifying Funny Scenes in Movies

**Haoqi Li[1], Naveen Kumar[2], Ruxin Chen[2], Panayiotis Georgiou[1]**

[1]Signal Processing for Communication Understanding and Behavior Analysis Laboratory (SCUBA), University of Southern California

[2]SONY Interactive Entertainment LLC

haoqili@usc.edu, naveen1.kumar@sony.com, ruxin.chen@sony.com, georgiou@sipi.usc.edu

## Overview

This paper presents a novel **deep Reinforcement Learning** framework for classifying **movie scenes** based on affect using the **faces images** detected in the video stream as input.

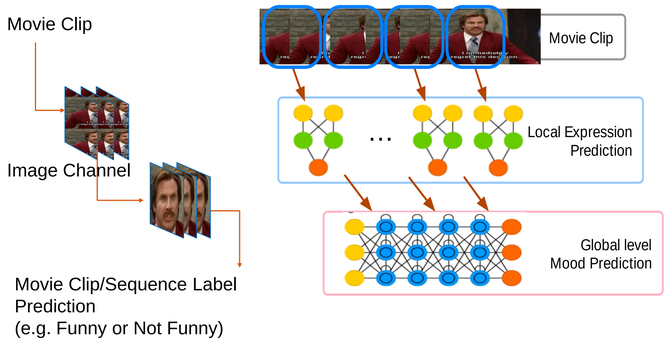
## Introduction

### ➢ Motivation

- Movie affective content analysis
- Limited labeled training data in supervised learning
- Understand movie scene's affective content through a process of "trial and error"

### ➢ Goal

- **Reinforcement Learning** for movie scene affective prediction
- Explore decision space irrespective of fine resolution labels
- Make use of human input when available




### ➢ Challenges

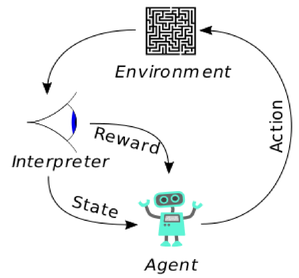
- Complexity of human interaction
- Complexity of movie data(e.g. different angles of faces)

## Dataset

- 18 movies form SONY Pictures
- Total 1471 scenes with scene level tone labels: funny, calm, exciting etc.
- Simplify the task to funny scene binary classification (sparse labels)

## Methodology

### ➢ Traditional RL framework

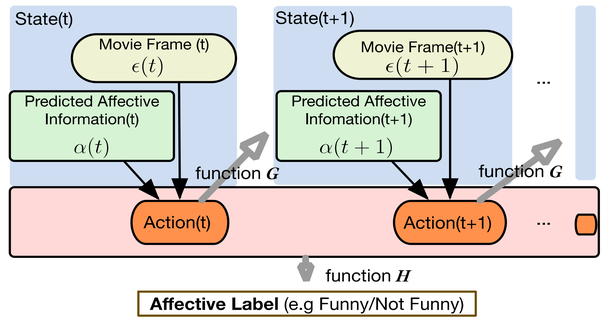



### ➢ Challenges

- Movie frame sequence is passive
- Agent's action does not change the movie sequence and "state"

### ➢ Framework of our proposed method RL framework

- Idea: Design the agent action that can interact with the environment
- Deep Q-network
- The environment state **s** = original movie face embedding **ε** + predicted affective information **α**



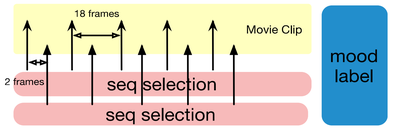

$$A(t) = \mathbf{Q}(s(t)) = \mathbf{Q}([\epsilon(t), \alpha(t)])$$
$$\mathbb{R}(t) = \mathbf{H}(A(1), \ldots, A(t), F)$$
$$\alpha(t+1) = \mathbf{G}(A(1), \ldots, A(t))$$

| | |
|---|---|
| **Q** | Deep Q-network |
| **A(t)** | Q-value over actions at time T |
| **R(t)** | Reward generated by function H based on actions taken so far |
| **G** | Mapping action to affective state |

## Experiments

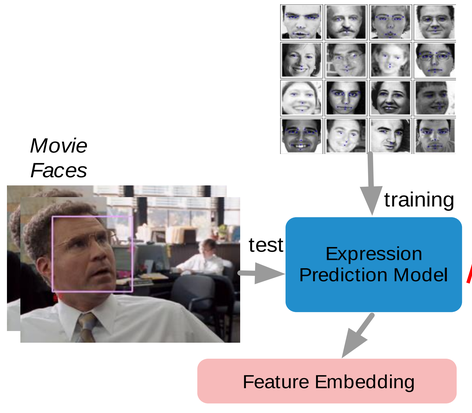
### ➢ Movie data processing

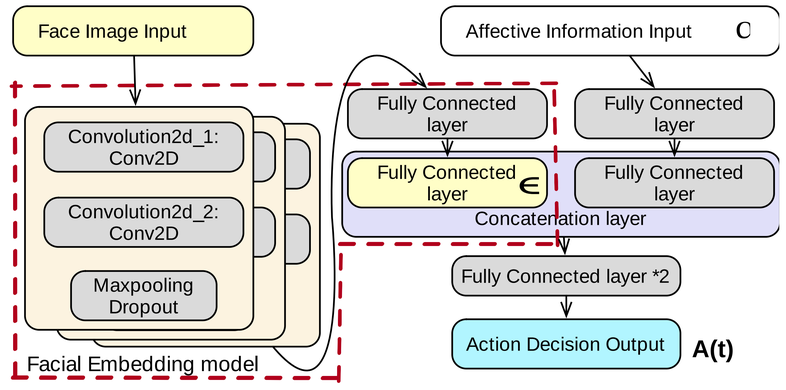
- Face extraction
- ~7000 training sequences




### ➢ Facial Expression Embedding model

- Pre-trained embedding on Kaggle FER
- 6 layer CNN with maxpooling




### ➢ RL model for affective label prediction

- Binary classification
- Feedback as G: one-hot encoding of maximum affective state activation

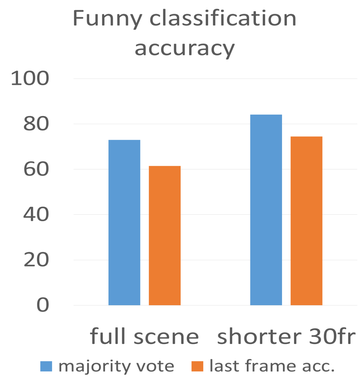
$$\alpha_i(t+1) = \begin{cases} 1 & \text{if } A_i(t) \geq A_j(t)\ \forall j \\ 0 & \text{otherwise} \end{cases}$$

e.g.

**0 0 1** ← **-1 10 21**

- Reward function (H):
  - Majority vote of frame-level results up to now.
  - Value of reward is critical ($R_{inter} = \pm 0.05$, $R_{end} = \pm 1$)

### ➢ Results




### ➢ Discussion & Conclusion

- Smaller frame length (30fr) reduces complexity of modeling sequence
- Potential applications for human in the loop training process