Recognition of Actions and Paths in Figure Skating

Xiaoyan Wu, Omid Poursaeed and Serge Belongie

Introduction

Figure Skating:

- A sport in which ice skaters perform free-style movements of jumps, spins, lifts, and footwork in a graceful manner.

Figure Skating

- The first winter sport included in the Olympics
- Has developed from a practical way to get around on ice into the elegant mix of art and sport it is today
- Has become one of the most popular sports of the winter Olympics
- Considering its popularity and its huge media coverage, there is a potential for enhancing spectators' experience by using techniques from the field of Computer Vision

Figure Skating Movements

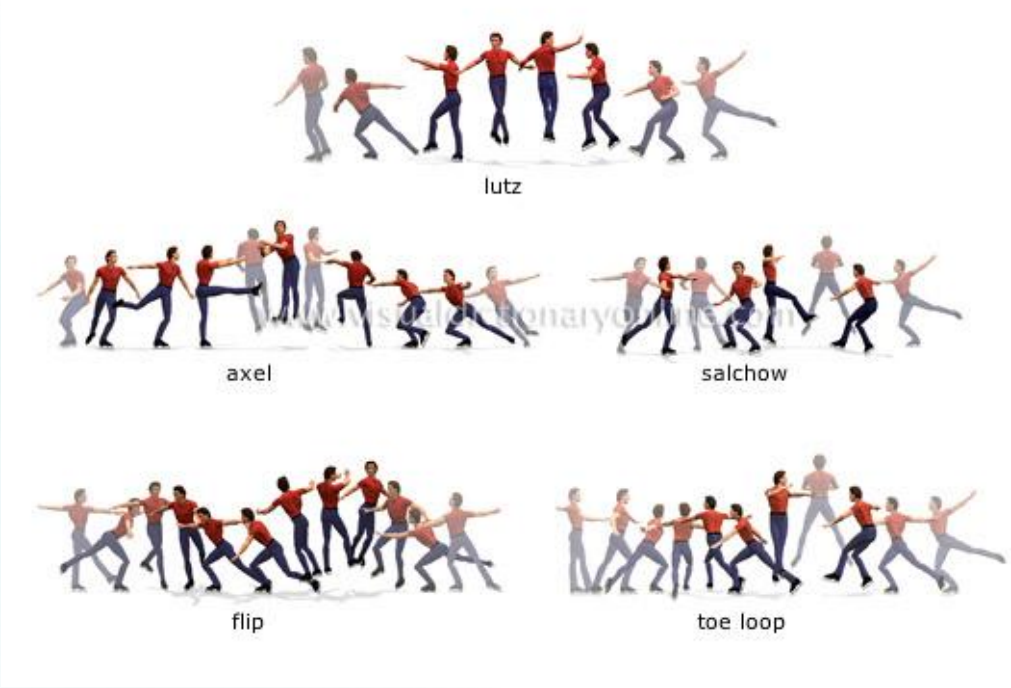
Figure 1. Professional figure skaters’ moves

Actions and Paths

- Skaters generally perform two programs: short and long
- Depending on the discipline, these programs may include spins, jumps, moves in the field and other elements or moves.
- In the short program the skater must complete a list of required elements consisting of jumps, spins and steps, and in the long program the skaters have slightly more choice of elements.

Actions and Paths

- Each move can be performed in several different ways.



Actions and Paths

- Each type of action corresponds to a particular pattern drawn by the skater.
- We can partition the path into subsections, each of which representing a specific action.
- One of the key components in determining the score of a figure skater is his ability to perform difficult actions.

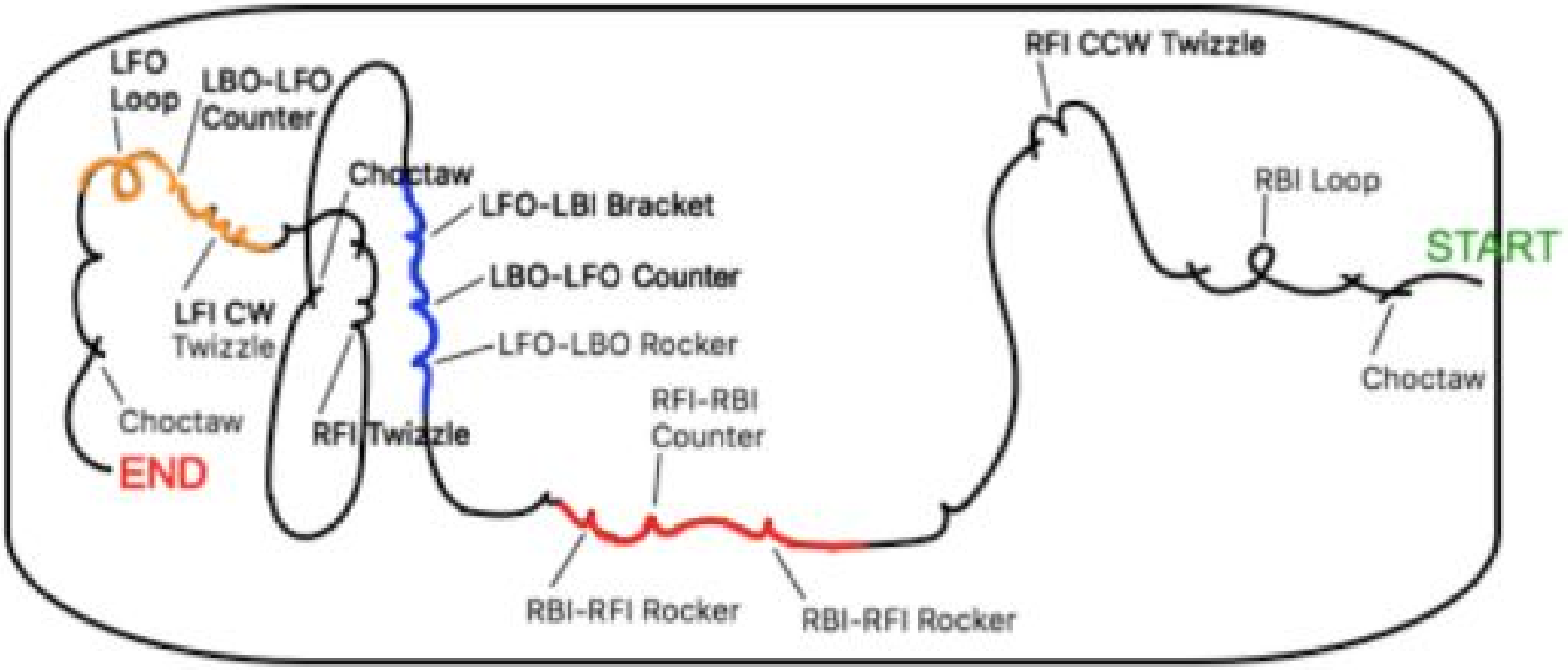


Figure 3. the path drawn by the skater and the specific type of action performed in each part of the path

Objectives

- Delivering enhanced sport-viewing experience to figure skating fans
- Building an automated method for determining the path drawn by the skaters, measuring their speed and specifying the action they perform at each part of the path.

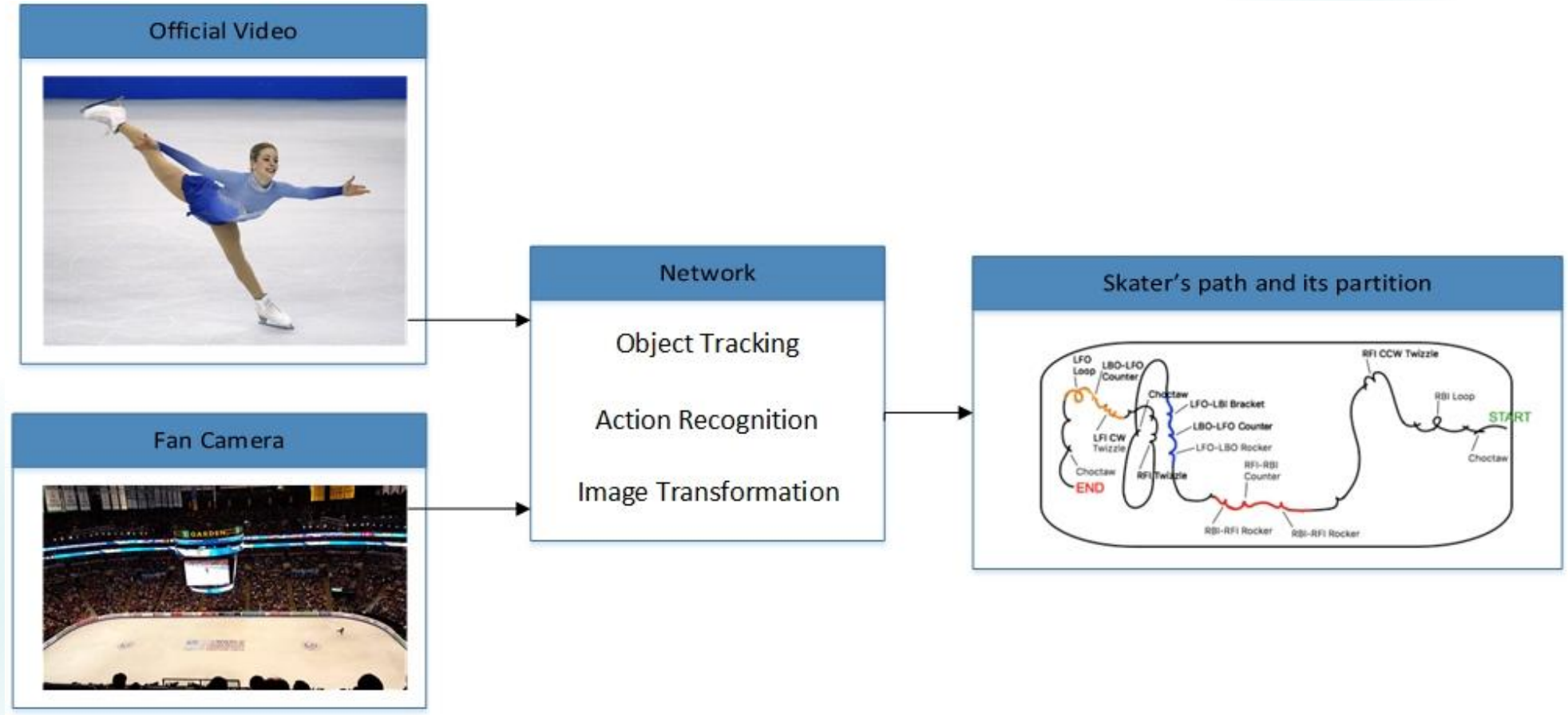


Figure 4. Schematic diagram of our network

Objectives

In order to train our model, we use both videos obtained from official sources and those recorded by fans.

Official videos:

- Have higher quality
- The camera moves very often which makes it challenging to use them in our model.

Videos recorded by fans:

- Generally of lower quality
- The camera is either fixed or moves slowly

We use both types of videos in order to obtain more accurate results.

Objectives

- Using object tracking, action recognition and image transformation, we process the inputs to estimate the path drawn by the skater.
- We divide the path into several parts each representing a specific action performed by the skater.

Preliminary Results

14

- **Object Tracking:**

As the first step in recognizing the actions and paths of the skater, we track the skater, i.e. we locate him at each frame and draw a bounding box around him.

Figure 5. Tracking results for two frames of the fan-camera

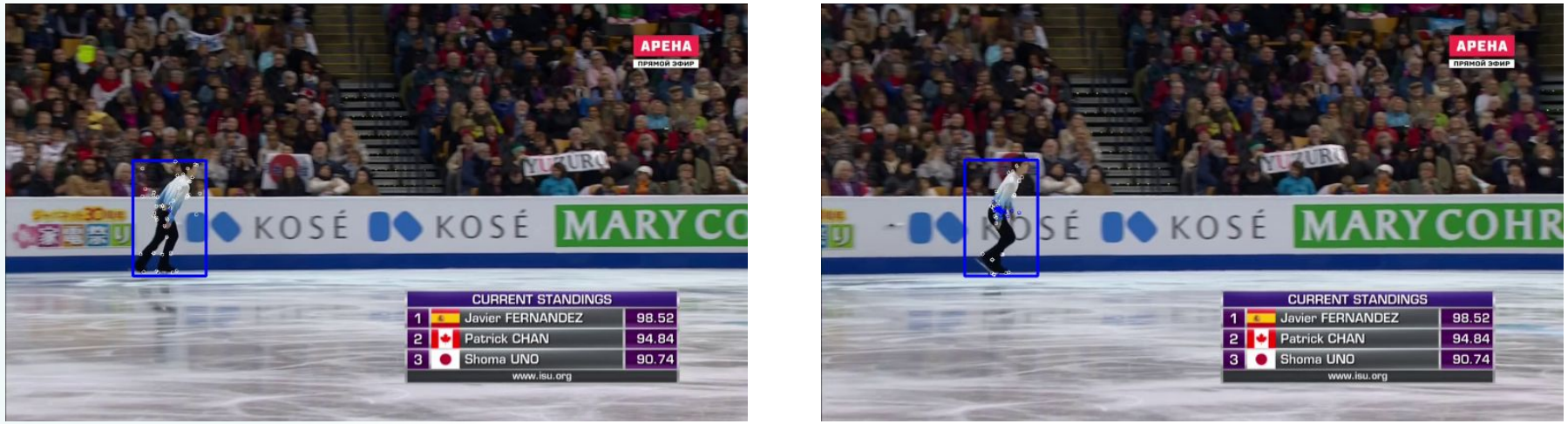


Figure 6. Tracking results for two frames of official videos

Preliminary Results

- **Skater’s path:**
- Using the middle point of the bottom of the bounding box as an approximation of the feet’s location.
- Detecting key-points on feet or the skates and track them as the skater moves.

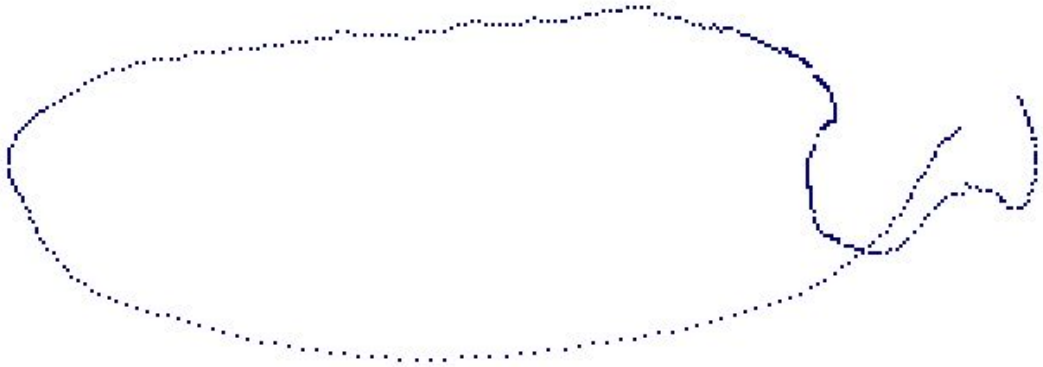


Figure 7. Skater’s approximate path

Preliminary Results

- **Rectification:**

To counter the effect of different camera viewpoints, we processed the videos into rectified, overhead view by applying a homography to each frame.

Figure 8. Transforming the image to rectified view

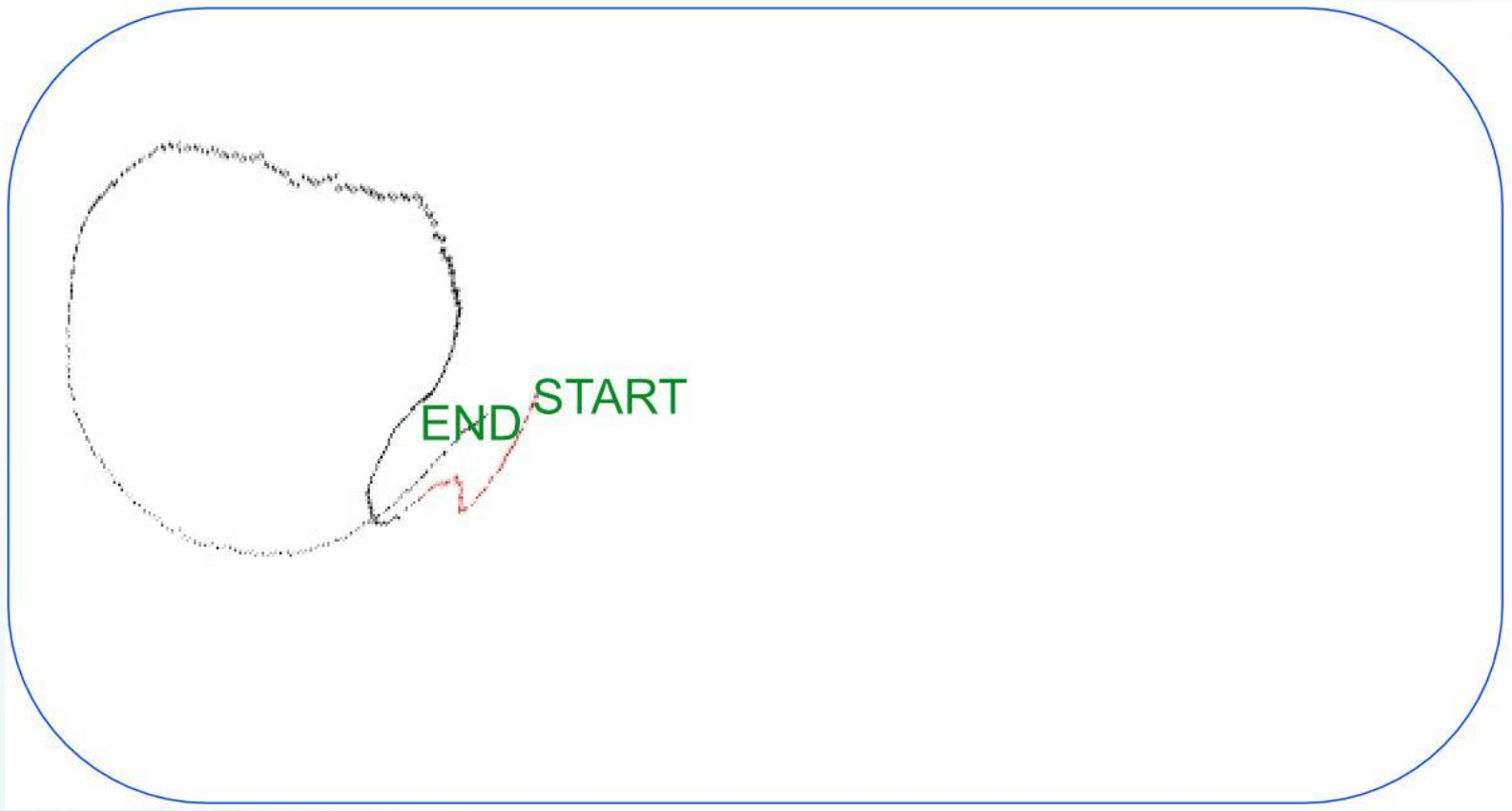


Figure 9. Skater’s path in rectified view

Future Work

- **Action Recognition:**

Recognizing the specific action the skater performs at each part of the path

- **Official videos for path estimation:**

Use official videos alongside with fan cameras to make the estimated path more accurate

Future Work: Action Recognition

Using Convolutional Neural Networks for pose estimation

- Stacked Hourglass Networks for Human Pose Estimation (Newell, et al.)

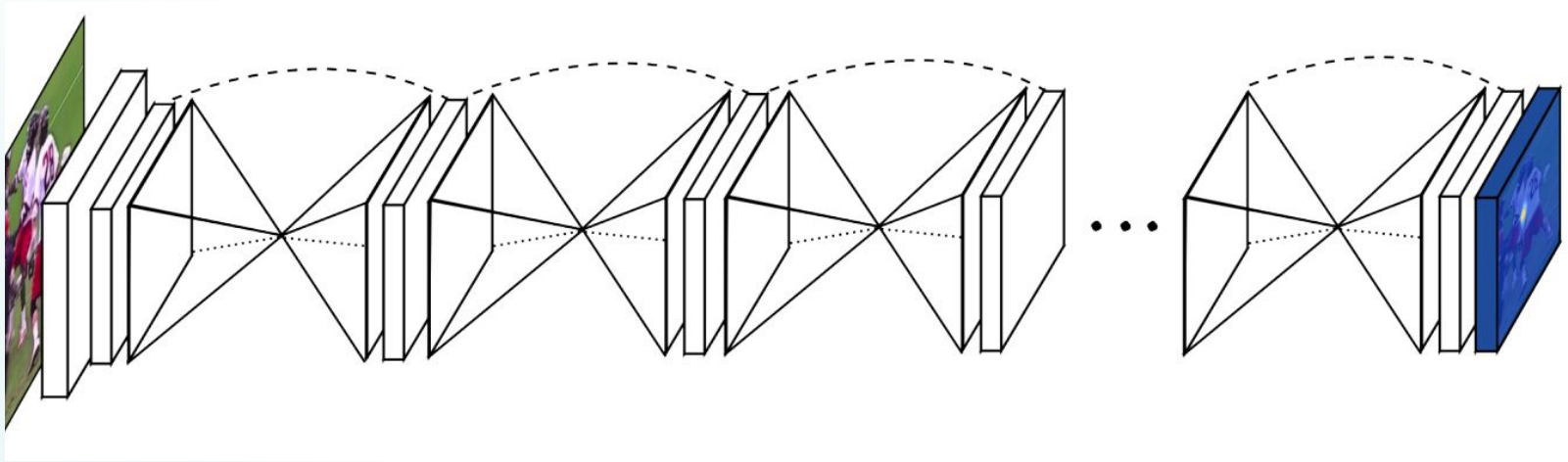


Figure 10. The network for pose estimation. It consists of multiple stacked hourglass modules which allow for repeated bottom-up, top-down inference

Pose Estimation

Figure 11. Example output produced by the network. On the left we see the final pose estimate provided by the max activations across each heatmap. On the right we show sample heatmaps. (From left to right: neck, left elbow, left wrist, right knee, right ankle)

Pose Estimation

Figure 12. Example Pose Estimation Output

Pose Estimation in Figure Skating

Assessing the Quality of Actions (Pirsiavash, et al.)

Collecting training data

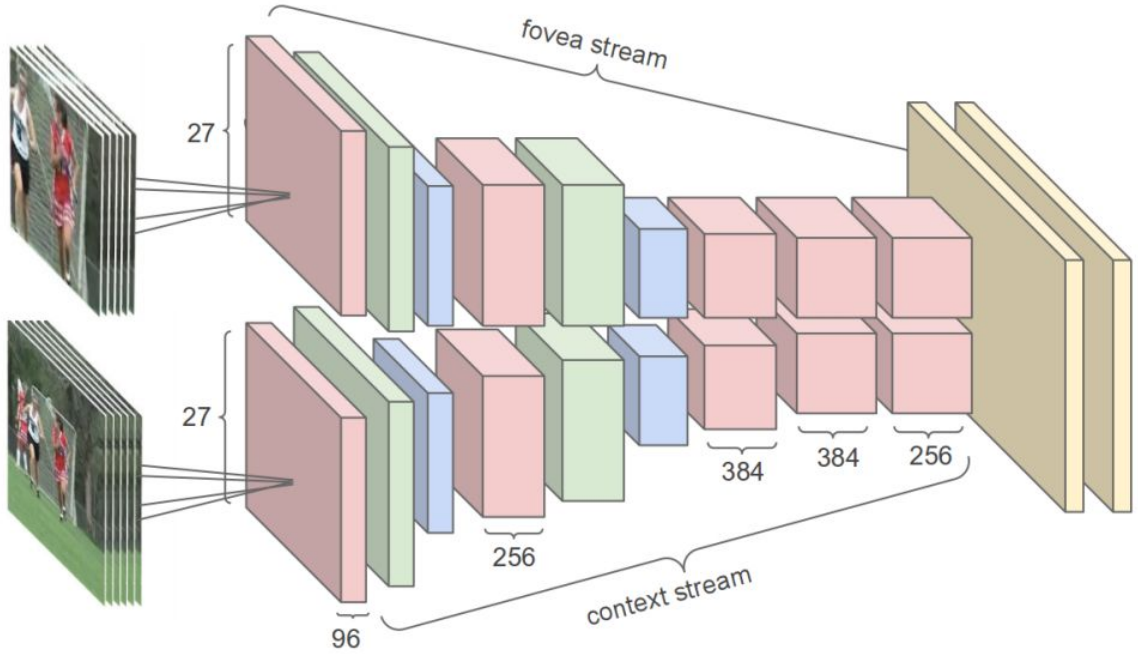
Jump

Spin

Action Classification

Using pose estimation to obtain feature vectors for action classification

Large-scale Video classification with Convolutional Neural Networks



Classification Results

Summary

- Collecting Training data using crowd-sourcing
- Pose estimation using convolutional neural networks
- Action classification using pose estimation and training labels