HIGHLY PRECISE MOTION TRANSITIONS DETECTION IN UNTRIMMED SPORTS VIDEOS USING SPATIO-TEMPORAL GRAPH EMBEDDINGS

1 Pattern Analysis for Transition Detection

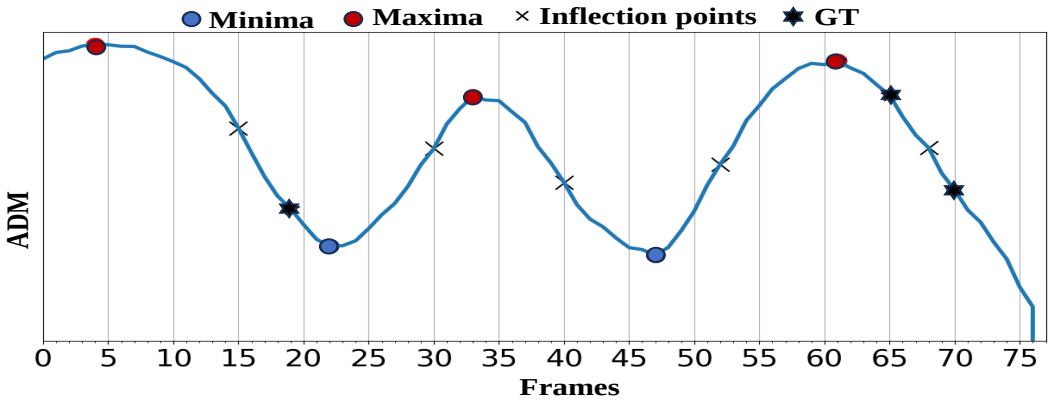

Figure 1: **Ground truth maxima, minima, and inflection points in ADM, representing key action transitions during the dive.**

The demarcations m_1, m_2, and m_3 are marked in the dataset as the ground truth (GT). Here, m_1 represents the pike position of the diver, m_2 marks the transition from the pike position (or the end of the somersault-pike rotation), and m_3 indicates the moment when the diver straightens for a controlled water entry.These points, m1, m2, and m3, are marked with a star symbol (*) for clear identification as shown in Figure 1.

To analyze these transitions, we use the Action Dynamics Metric (ADM), where we identify **minima (marked in red), maxima (marked in red), and inflection points (marked by a cross)** as shown in Figure 1. Among these, we determine which three points—whether maxima, minima, or inflection points—are closest to the GT demarcations m_1, m_2, and m_3. We denote **maxima as P, minima as V, and inflection points as F**, forming different patterns based on their proximity to the demarcation points.

We tested the following **27 patterns**:

”PPP”, ”PPV”, ”PPF”, ”PVP”, ”PVV”, ”PVF”, ”PFP”, ”PFV”, ”PFF”,
”VPP”, ”VPV”, ”VPF”, ”VVP”, ”VVV”, ”VVF”, ”VFP”, ”VFV”, ”VFF”,
”FPP”, ”FPV”, ”FPF”, ”FVP”, ”FVV”, ”FVF”, ”FFP”, ”FFV”, ”FFF”

Among these, **PPP, VVV, and FFF** represent the closest points to the GT demarcations m_1, m_2, m_3. The pattern **PPP** corresponds to the nearest maxima, **VVV** corresponds to the nearest minima, and **FFF** corresponds to the nearest inflection points.

We tested all patterns in the ADM of all videos and found that **FFF is the pattern closest to the GT**, as seen in Figure 2. The **average gap for all videos** is also shown in Figure 2. This suggests that **inflection points in the ADM are the most reliable indicators of key transitions in the diver's motion**.

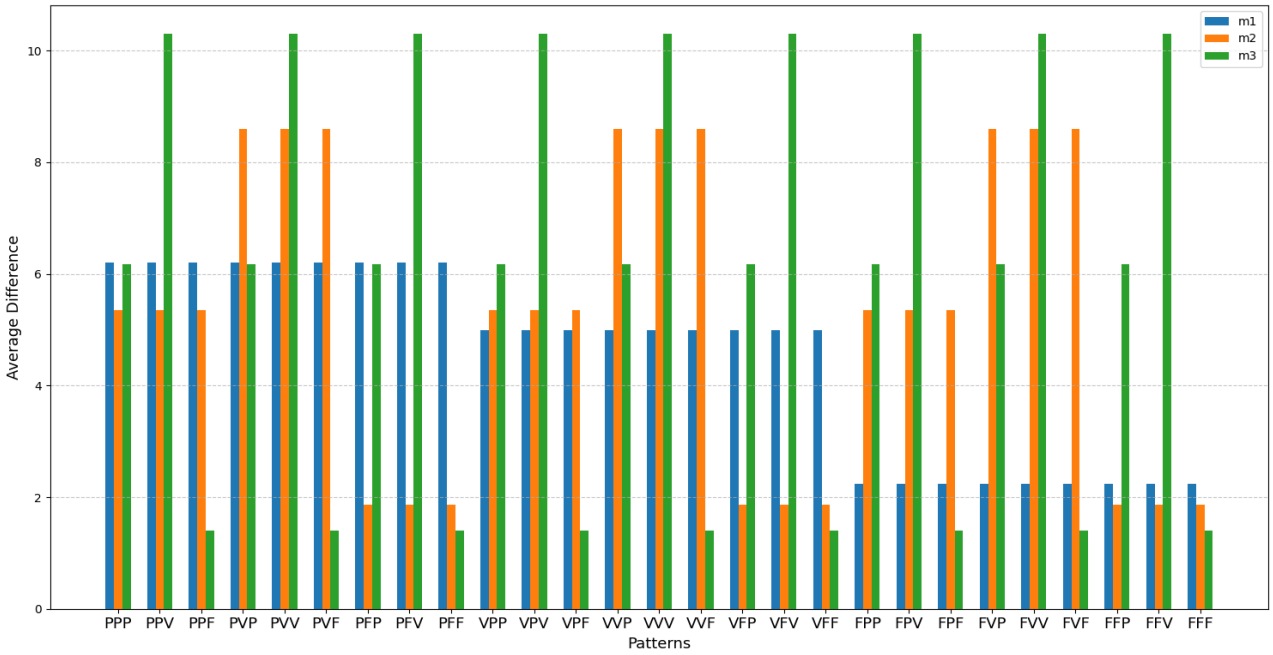

Figure 2: **Histogram of average difference for the ground truth demarcations m_1,m_2 and m_3 across all patterns for all videos.**

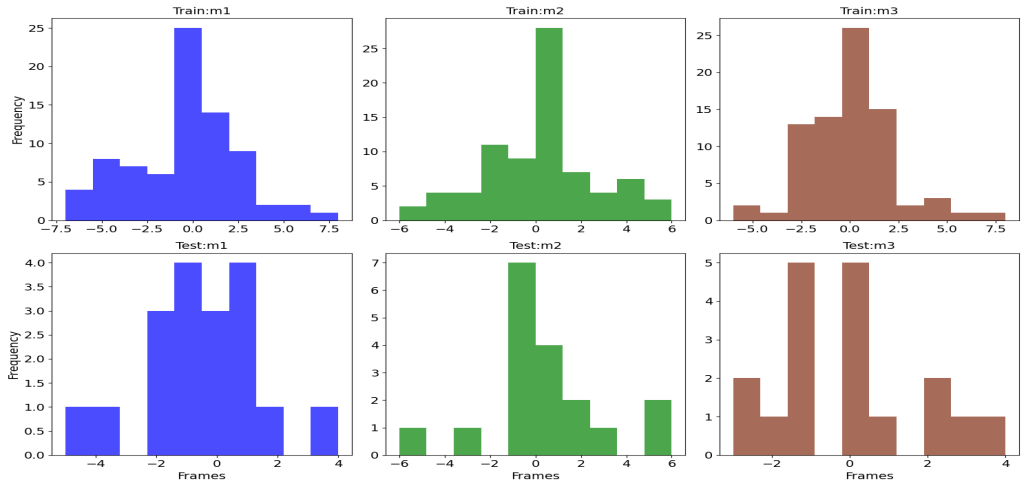

Figure 3: **The histogram shows the distribution of the number of videos following the "FFF" pattern, illustrating how often the predicted inflection points occur before, after, or exactly at the ground truth (GT) demarcations for both training and testing videos.**

2 Temporal Analysis of Predicted Demarcations

After identifying the pattern that is closest to the ground truth (GT) demarcations (m_1, m_2, and m_3), we specifically computed the number of videos in which the predicted inflection point demarcations occur before or after the actual GT demarcations. Additionally, we determined the number of videos where the predicted demarcations exactly match the GT. This analysis was performed separately for training and testing videos.

The negative values on the x-axis represent predictions occurring before the reference frame, while positive values indicate predictions occurring after the reference frame, as shown in Figure 3.

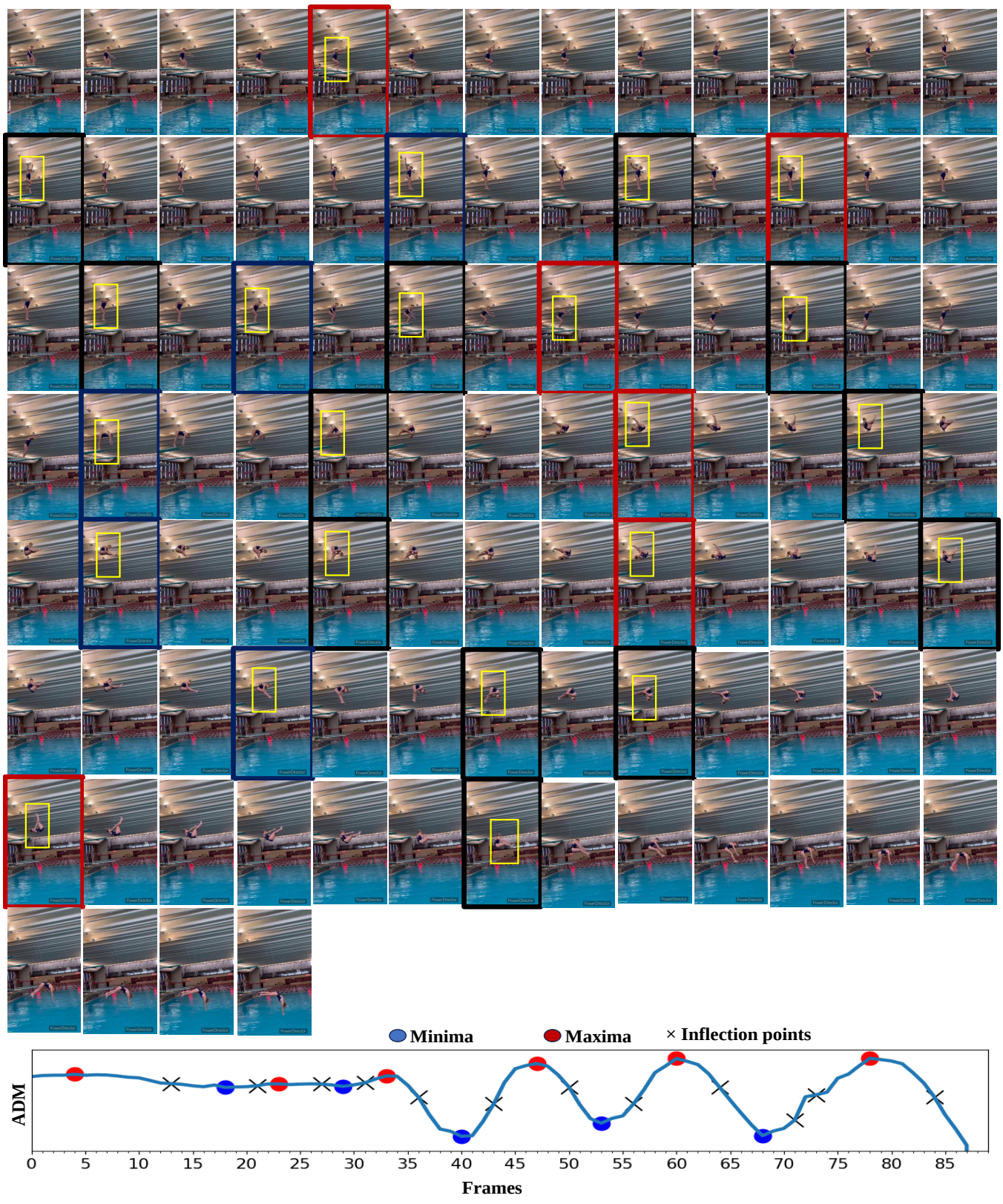

Figure 4: **Frames from the YouTube diving video with ADM, highlighting key action transitions based on maxima, minima, and inflection points.**

3 Detecting Dynamic Transitions in Diving via ADM

This figure 4 presents all the frames from the YouTube diving videos alongside the ADM (Action Dynamics Metric) of the respective video. The red-marked frames indicate the maxima points in the ADM, while the blue-marked frames correspond to the minima. Additionally, the black-marked frames represent the inflection points in the ADM. These marked frames serve as key action transition detectors, capturing the moments where the diver changes actions during the dive. The 3rd to the 4th minima in the ADM represent one complete rotation of the somersault pike, while the 4th to the 5th minima indicate the second rotation of the somersault pike, as shown in the figure!4. The ADM effectively highlights these transitions, providing insights into the dynamic shifts in movement throughout the diving sequence.