

Olympic Highlights Dataset

A benchmark for identity-aware sports video summarization

Marcos Rodrigo Carlos Cuevas Narciso García

Grupo de Tratamiento de Imágenes (GTI), Information Processing and Telecommunications Center (IPTC),

Universidad Politécnica de Madrid (UPM), Madrid, Spain

marcos.rodrigo@upm.es

Overview

Olympic Highlights is a publicly available benchmark of full-length broadcast track-and-field competition videos with frame-level temporal annotations and athlete-identity labels. It is designed to support reproducible evaluation of identity-aware sports video summarization systems.

The dataset consists of **20 videos** spanning four events: **High Jump**, **Javelin**, **Long Jump**, and **Pole Vault**, with five videos per event. Each video is a full broadcast recording of approximately 30 minutes (roughly 10 hours of footage in total), collected from publicly available Olympic and World Athletics broadcasts on YouTube.

Each video is annotated at the frame level with three segment types:

- **Highlight (HL):** Frames covering the core athletic action/attempt and its immediate outcome.
- **Non-Highlight (NHL):** Background footage between highlight events (transitions, crowd shots, etc.).
- **Uncertainty (UN):** Boundary frames immediately surrounding each highlight, accounting for natural ambiguity in annotation boundaries (approximately 15 frames before and 30 frames after each highlight).

In addition, every highlight segment is labeled with the **athlete responsible** for the action, enabling evaluation of identity-aware summarization systems.

Contents

The dataset is distributed as a ZIP archive with the following structure:

```
OlympicHighlights/  
|-- videos.csv  
|-- High Jump/  
|   |-- high_jump_1.csv ... high_jump_5.csv  
|-- Javelin/  
|   |-- javelin_1.csv ... javelin_5.csv  
|-- Long Jump/  
|   |-- long_jump_1.csv ... long_jump_5.csv  
'-- Pole Vault/  
    |-- pole_vault_1.csv ... pole_vault_5.csv
```

Dataset Statistics

Table 1 summarises the ground-truth annotation statistics for every video in the dataset. For each segment type the table reports the average event duration (Avg, in seconds), the total cumulative duration (Tot, in minutes), and the percentage of video time occupied by that category.

Table 1: Ground-truth statistics for the Olympic Highlights dataset (20 videos; 5 per event).

Video	Highlight (HL)			Non-Highlight (NHL)			Uncertainty (UN)		
	Avg (s)	Tot (m)	%	Avg (s)	Tot (m)	%	Avg (s)	Tot (m)	%
High Jump 1	8.08	7.81	24.2	23.53	22.75	70.5	0.90	1.73	5.4
High Jump 2	5.41	5.32	16.3	25.46	25.46	78.2	0.90	1.77	5.4
High Jump 3	7.43	6.93	19.4	30.01	28.00	78.3	0.45	0.84	2.3
High Jump 4	5.79	3.09	15.3	29.26	16.09	79.9	0.90	0.96	4.8
High Jump 5	6.13	5.83	18.7	24.40	23.59	75.8	0.90	1.71	5.5
<hr/>									
Javelin 1	12.11	7.06	26.0	31.82	19.09	70.2	0.90	1.05	3.9
Javelin 2	11.77	1.11	31.1	24.88	15.76	65.3	0.90	1.11	4.6
Javelin 3	11.41	6.46	28.2	26.47	15.44	67.4	0.90	1.02	4.5
Javelin 4	8.90	4.60	24.9	24.24	12.93	70.1	0.90	0.93	5.0
Javelin 5	13.31	9.54	31.3	27.69	20.31	66.6	0.45	0.65	2.1
<hr/>									
Long Jump 1	10.81	8.65	29.1	24.01	19.61	66.0	0.90	1.44	4.9
Long Jump 2	10.99	14.84	40.8	14.13	19.07	52.5	0.90	2.42	6.7
Long Jump 3	13.45	11.21	31.4	27.03	22.98	64.4	0.90	1.50	4.2
Long Jump 4	12.95	7.26	26.1	34.28	33.14	70.3	0.90	1.71	3.6
Long Jump 5	15.18	8.86	31.1	30.85	18.51	65.0	0.94	1.09	3.9
<hr/>									
Pole Vault 1	10.97	7.86	30.4	22.84	16.75	64.7	0.90	1.29	5.0
Pole Vault 2	11.29	7.71	24.6	31.96	22.37	71.4	0.90	1.23	3.9
Pole Vault 3	12.03	9.22	26.0	32.66	25.58	72.1	0.45	0.69	1.9
Pole Vault 4	12.54	9.82	26.2	26.33	26.33	70.1	0.90	1.40	3.7
Pole Vault 5	10.40	8.32	27.9	25.48	20.81	69.7	0.45	0.72	2.4

Ground Truth File Format

Each ground truth file (e.g., `high_jump_1.csv`) contains frame-level annotations in CSV format with the following columns:

Column	Description
Event type	Segment label: Highlight, Not a highlight, or Uncertainty
First frame	Index of the first frame of the segment (inclusive)
Last frame	Index of the last frame of the segment (inclusive)
Num. frames	Number of frames in the segment

The segments are listed in chronological order and together span the full video from frame 0 to the last frame. Frame indices assume a fixed 30 fps frame rate.

Example:

```
Event type,First frame,Last frame,Num. frames
Not a highlight,0,6087,6088
Uncertainty,6088,6102,15
Highlight,6103,6298,196
Uncertainty,6299,6328,30
Not a highlight,6329,6386,58
...
```

Video Links

The videos are not distributed directly due to copyright. The file `videos.csv` (included in this archive) lists the YouTube URL for each of the 20 videos. All videos are from the World Athletics and European Athletics YouTube channels and were publicly available at the time of annotation.

`videos.csv` columns: `sport`, `ground_truth_file`, `url`.

Evaluation Protocol

The dataset was used in the associated paper to evaluate identity-aware highlight selection at the *event level*. A predicted highlight segment is counted as a true positive if:

1. Its temporal Intersection-over-Union (IoU) with a ground-truth highlight is ≥ 0.3 , *and*
2. It is attributed to the correct athlete, *and*
3. No other prediction has already been matched to that ground-truth event (one-to-one matching).

Recall, precision, and F-score are computed per video, then averaged per event and across the full dataset. Uncertainty frames are excluded from frame-level evaluation.

Citation

If you use this dataset in your research, please cite:

M. Rodrigo, C. Cuevas, and N. García, “Automatic Sports Video Summarization with Identity-Aware Highlight Selection,” *Image and Vision Computing*, under review.

```
@article{rodrigo2025pvs,  
  title   = {Automatic Sports Video Summarization with  
            Identity-Aware Highlight Selection},  
  author  = {Rodrigo, Marcos and Cuevas, Carlos and  
            Garcia, Narciso},  
  journal = {Image and Vision Computing},  
  note    = {Under review}  
}
```

Code and additional resources are available at <https://github.com/MarcosRodrigoT/PVS>.