Medium Scale Benchmark for Cricket Excited Actions Understanding

Sejong University
Seoul, South Korea

Altaf Hussain, Noman Khan, Muhammad Munsif, Min Je Kim, and Sung Wook Baik
Sejong University, Republic of Korea

CVPR SEATTLE, WA JUNE 17-21, 2024

Introduction

- The Sports Action Recognition (SAR) domain is crucial in research, with applications in assisting coaching, enhancing athletic performance, entertainment, and generating highlights.
- Prominent datasets like ActivityNet and Kinetics-400 focus mainly on daily life activities such as walking, running, and sitting. These datasets lack the granularity needed for specific sports analyses like cricket.
- Existing sports action datasets cover a wide range of sports but fail to provide detailed and applicable data for Cricket Action Analysis (CAA).

Catch

Clean Bowled

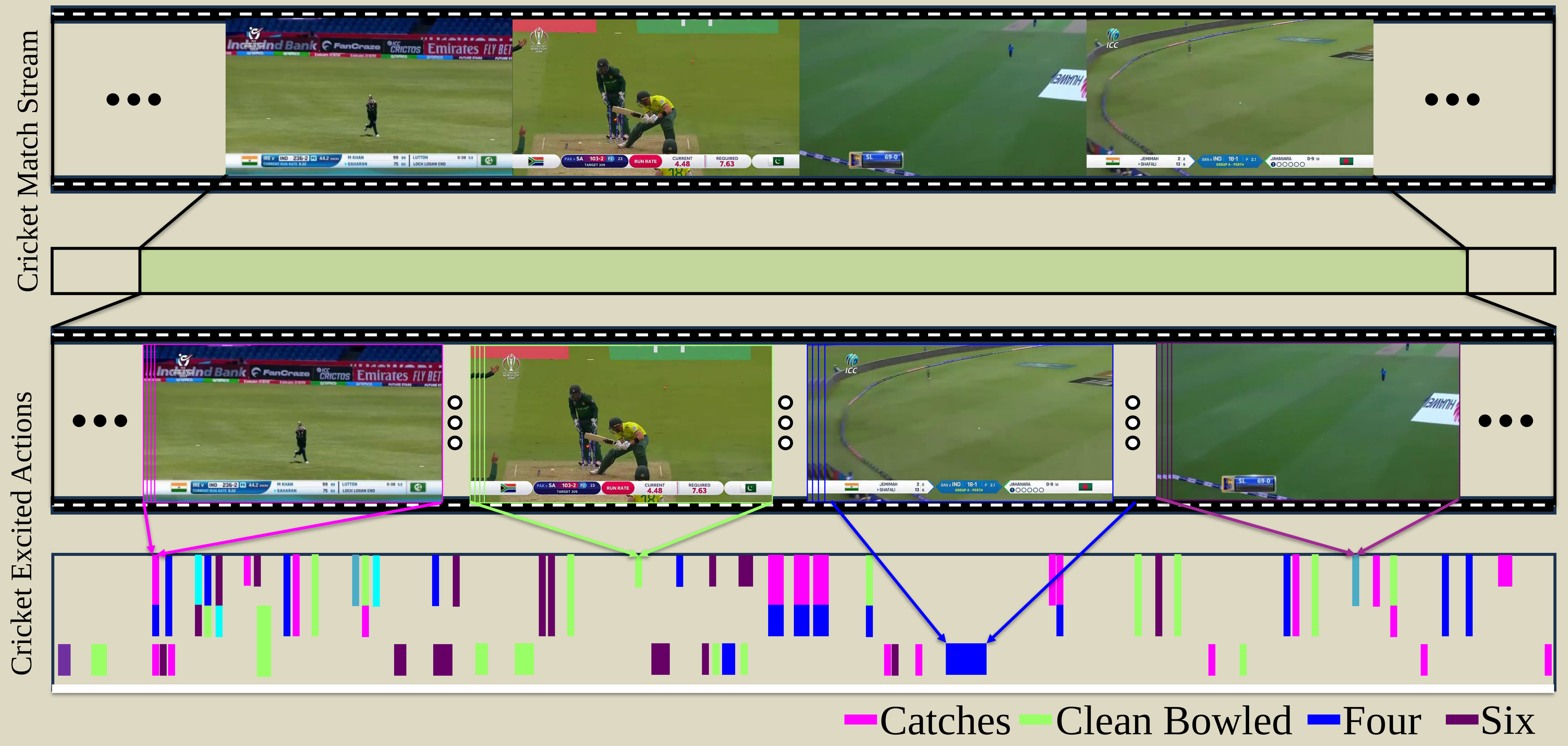
Four

Six

Only three datasets have been proposed for CAA, each with significant limitations:

- **DPC Images:** Contains 8,646 images but covers only two classes (delivery and play).
- **EXINP:** An audio dataset with 868 clips categorized into Excited, Interval, and Normal Play.
- **CKT:** A small dataset with 722 video clips with only 150 videos per category.
- These datasets are constrained by their generality, limited classes, and small scale, restricting their practical use in real world CAA applications.

Proposed Cricket Excited Actions Dataset



- The study introduces the Cricket Excited Actions (CEA) dataset, developed with professional cricket players, to overcome the limitations of existing datasets.
- The CEA dataset includes comprehensive and realistic multi-person cricket actions, covering key moments with accurate human pose information, potential player interactions, and occlusion challenges.

Contributions:

- Introduced the CEA dataset to advance spatiotemporal CAA through deep learning, bridging academia with real-world sports applications by optimizing player performance.
- Conducted detailed experiments on the CEA dataset using baseline methods to analyze their performance. These experiments elucidate the primary challenges associated with recognizing complex sports activities, offering valuable insights for the research community to leverage in future endeavors within this domain.

Experimental Results

- **CNN Models:** Evaluated C3D, TSN, C2D, I3D, R2+1D, CSN, and SlowOnly, with top-1 accuracy scores from 41.41 to 85.90. The CEA dataset posed significant challenges compared to the CKT dataset.
- **Transformer Models:** Evaluated state-of-the-art TimeSformer, VideoSwin, and UniFormerV2, with top-1 (%) accuracy scores of 76.65, 34.80, and 62.11. CNN models outperformed transformers, highlighting the complexity of the CEA dataset.

Methods	Year	Res	Top-1 (%)	
			CEA	CKT
C3D [43]	2015	112 × 112	83.26	81.13
TSN [48]	2016	224 × 224	80.18	--
C2D [50]	2018	224 × 224	81.50	84.91
I3D [8]	2018	224 × 224	80.62	97.17
R2+1D [44]	2018	112 × 112	85.90	91.51
CSN [45]	2019	224 × 224	75.77	99.06
SlowOnly [12]	2019	256 × 256	41.41	99.06
TimeSformer [5]	2021	224 × 224	76.65	96.23
VideoSwin [31]	2022	224 × 224	34.80	--
UniFormerV2 [25]	2022	224 × 224	62.11	--

- **Model Performance Summary:**

CNN models outperformed transformer models in CAA due to their inductive biases. The evaluation with the CEA dataset highlights the strengths of both CNN and transformer models, providing valuable insights into sports analytics and deep learning for complex sports activities.

Catch Clean Bowled Four Six

The data and code are available at
https://github.com/Altaf-hucn/Cricket-Excited-Actions-Benchmark