Widely Applicable Strong Baseline for Sports Ball Detection & Tracking

Shuhei Tarashima, Muhammad Abdul Haq, Yushan Wang, Norio Tagawa

TL; DR
- We propose a new SBDT baseline, WASB.
- We introduce a new evaluation protocol using 5 SBDT datasets from different sports. 6 SOTA methods are (re-)implemented for fair comparison.
- Experiments show that WASB substantially outperforms SBDT SOTAs on all the datasets.

Dataset & Codebase
- SBDT datasets from 5 different sport categories: Volleyball, Basketball, Soccer, Badminton, and Volleyball are newly introduced by us.
- for Soccer and Basketball, new annotations are provided.

1. High-Resolution Feature Extraction Model
   - High-Resolution Modules (HRMs) of small HRNet are used.
   - Stem without strides to feed higher-resolution features to HRMs.
   - Multi-In Multi-Out (MIMO) design (N = 3).

2. Position-Aware Model Training
   - Train a model that predicts heatmaps representing ball positions.
   - Focal-loss [6] with binary ground truth (GT) during the first T epochs.

3. A Bunch of Tricks during Inference
   - Prediction of each ball position (i.e., (x, y)-coordinate) as a center of heatmap values in a detected blob.
   - online tracking with local motion model to take long-term temporal consistency into account.
   - oversampling the same image in different MIMO combinations to produce diverse detection candidates.

Input: a (sports) video clip. Output: a (x, y)-coordinate of a sports ball (if visible) for each frame.

Results
- Comparison on 5 datasets from different sports (distance threshold τ = 4[pixel]).
- Comparison on 5 sports datasets with different τ.

WASB
- DeepBall [1], DeepBall-Large, BallSeg [2], TrackNetV2 [3], ResTrackNetV2, MonoTrack [4].

Qualitative results (blue: ground truth, red: prediction).
