CVPR Nashville JUNE 11-15, 2025

DANIEL KIENZLE, ROBIN SCHÖN, RAINER LIENHART, SHIN'ICHI SATOH

TOWARDS BALL SPIN AND TRAJECTORY ANALYSIS IN TABLE TENNIS BROADCAST VIDEOS VIA PHYSICALLY GROUNDED SYNTHETIC-TO-REAL TRANSFER

NII UNIVERSITY OF STUTTGART

MOTIVATION

- Ball Spin and 3D Trajectory are key to gameplay analytics
  → Improve training, extract statistics, enable virtual replay
- **Ball motion alone provides sufficient information**
  → No need for analyzing human behavior
- **Ball trajectories are deterministic**
  → Fully governed by physics and can be simulated

➔ **Train on synthetic data only**

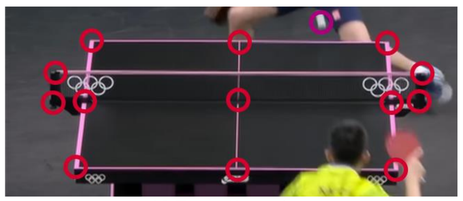
KEY IDEAS

- Train a neural network to **predict the ball's initial spin and 3D trajectory**
- **Uplifting approach**: Use 2D trajectory and table points as input → Built on top of established 2D keypoint methods
- **Fully synthetic training:** Generate physically accurate 2D trajectories using MuJoCo [1, 2] → No real-world data required
- **Zero-Shot Generalization**

ZERO-SHOT GENERALIZATION

- **Physically accurate training data** from simulation
- **2D pixel coordinates as input**
- **Targeted augmentations:** Motion blur, annotation noise, early trajectory cutoffs

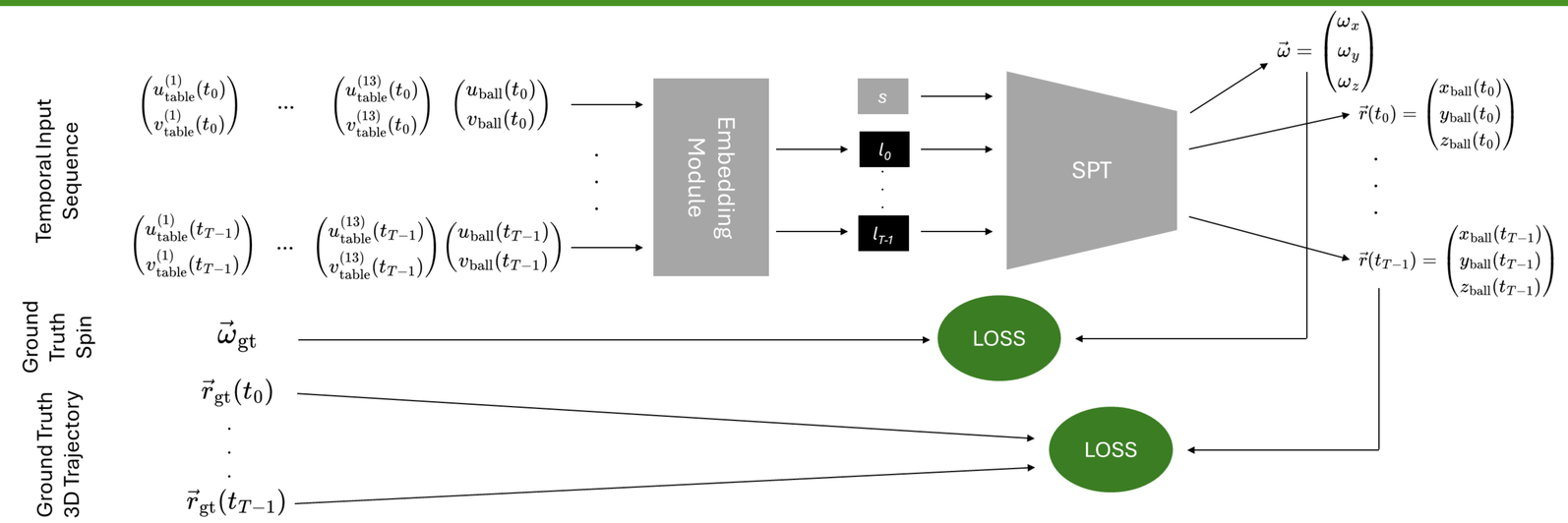
➔ **Real and synthetic data become indistinguishable**



13 Table Keypoints

References:
- [1]: E. Todorov, T. Erez and Y. Tassa, "MuJoCo: A physics engine for model-based control,"IEEE/RSJ International Conference on Intelligent Robots and Systems, 2012
- [2]: D. B. D'Ambrosio, S. W. Abeyruwan, L. Graesser, et.al., "Achieving Human Level Competitive Robot Table Tennis", 7th Robot Learning Workshop: Towards Robots with Human-Level Abilities, 2025

METHOD



- **For each frame at time t_i :**
  → Extract 2D ball coordinates and table keypoints
  → Embed into a **location token** l_i
- Prepend a **learnable spin token** s to the sequence $\{s, l_0, ..., l_{T-1}\}$
- Process the sequence with the **Spin Prediction Transformer (SPT). Output:**
  → 3D position $\vec{r}(t_i)$ for each time t_i
  → Initial ball spin $\vec{\omega}$
- Training: **Fully supervised** using only physically accurate **synthetic data**

EVALUATION DATA

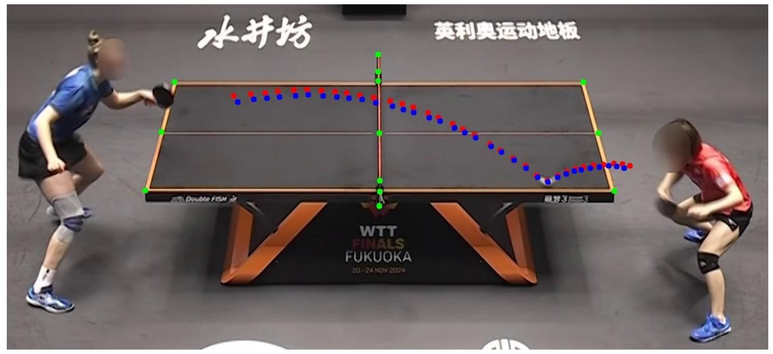
- Annotations: 2D trajectory, table keypoints, binary spin class
- Evaluation: Trajectory reprojection error and spin classification

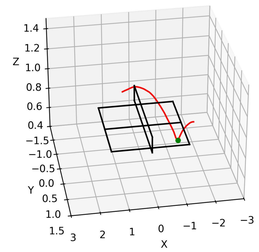
Backspin

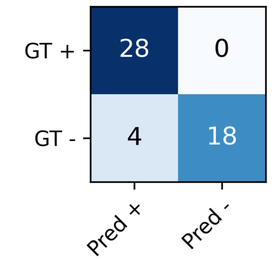
Topspin

RESULTS

Evaluation on real broadcast videos







Method: motion blur	Method: sudden end	Method: gaus. blur	spin: *acc* ↑	spin: F_1 ↑	trajectory: $\Delta\vec{r}_{\text{img}}$ ↓	trajectory: rel. $\Delta\vec{r}_{\text{img}}$ ↓
×	×	×	74.0 %	0.740	15.6 px	0.53 %
✓	×	×	88.0 %	0.875	16.1 px	0.55 %
×	✓	×	96.0 %	0.959	9.9 px	0.34 %
×	×	✓	80.0 %	0.800	18.9 px	0.64 %
✓	✓	✓	92.0 %	0.917	5.6 px	0.19 %