Multi-Modal Trajectory Prediction of NBA Players

Authors:
Sandro Hauri, Temple University
Nemanja Djuric, Uber ATG
Vladan Radosaljevic, Spotify
Slobodan Vucetic, Temple University

NBA Dataset Overview

- Recent improvements in CV allow for precise localization of NBA players
- Human trajectory prediction
- Basketball player
  - face difficult decisions at every time step
  - take into account other players on the court
  - show a wide range of human locomotion

- Data of more than 300 games publicly available
- Dataset contains 25 updates per second
- 113,760 possession
- Approximately 305 hours of game play

Input and Output Data

- Predict trajectory of single offensive NBA player
- Take into account available data at each time step
- Predict *M* possible paths
- Predict a probability for each generated path

Table 1: Input data for one time step, containing a total of 23 values

Input	Variable	Unit	Range
ball	x	[ft]	[0, 94]
	y	[ft]	[0, 50]
player P	x	[ft]	[0, 94]
	y	[ft]	[0, 50]
4 team mates	x	[ft]	[0, 94]
	y	[ft]	[0, 50]
5 opponents	x	[ft]	[0, 94]
	y	[ft]	[0, 50]
shot clock	t	[s]	[0, 24]

Table 2: Output contains a time series of H output velocities in x- and y-directions, as well as per-mode probability for M modes

Output	Variable	Unit	Range
M trajectories for player P	$\hat{\nu}_m$	[ft/s]	[-25, 25]
M probabilities	$\hat{p}_m$	[none]	[0, 1]

Loss Function

- We adapt previously published multi-modal loss function
- Addition of Relaxed Kronecker Delta allows to ramp up reward for best mode

$$\mathcal{L}^{\text{MSE}}(\boldsymbol{\nu}, \hat{\boldsymbol{\nu}}) = \frac{1}{2H} \|\boldsymbol{\nu} - \hat{\boldsymbol{\nu}}\|_2^2$$

$$\mathcal{L}^{\text{MTP}} = \sum_{m=1}^{M} \delta_\epsilon(m = m^*) \Big(\log \hat{p}_m + \alpha \mathcal{L}^{\text{MSE}}(\boldsymbol{\nu}_t^P, \hat{\boldsymbol{\nu}}_{t,m}^P) \Big)$$

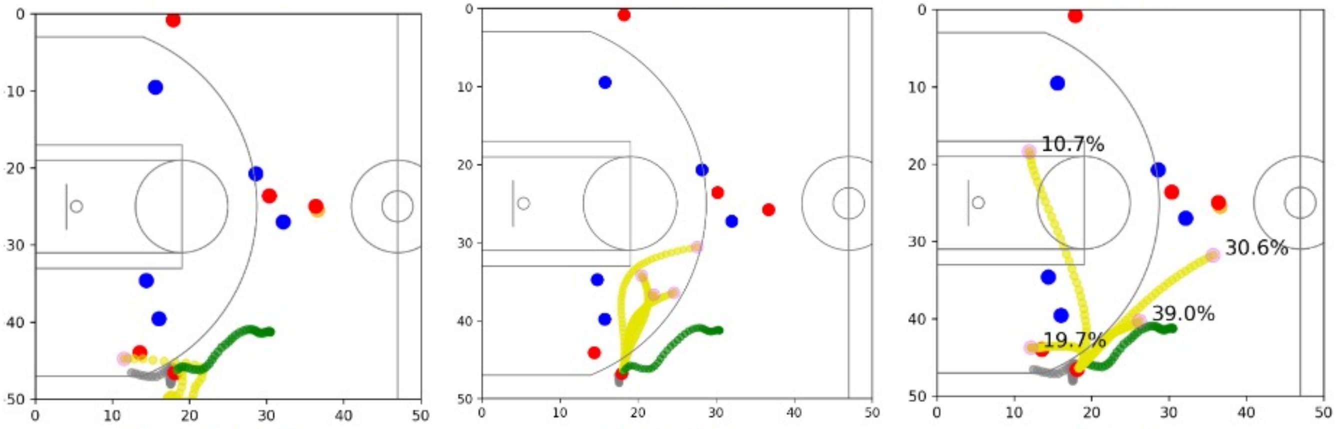
$$m^* = \underset{m \in \{1,\dots,M\}}{\arg\min} \; dist(\boldsymbol{\nu}_t^P, \hat{\boldsymbol{\nu}}_{t,m}^P)$$

$$\delta_\epsilon(cond) = \begin{cases} 1 - \epsilon, & \text{if condition } cond \text{ is true,} \\ \frac{\epsilon}{M-1}, & \text{otherwise.} \end{cases}$$

TEMPLE UNIVERSITY

Qualitative Results

- Generated trajectories are
  - smooth
  - plausible
  - physically achievable
  - diverse



(d) MACRO VRNN$_1$ (e) SocialGAN$_4$ (f) MBT$_{4l}$

Figure 1: Visualization of predicted trajectories with $H = 40$ using several state-of-the-art methods: a) location-LSTM, b) CNN, c) MBT$_1$, d) MACRO VRNN$_1$, e) SocialGAN$_4$, f) MBT$_{4l}$ (ours), red: attackers, blue: defenders, orange: ball, grey: input history of predicted player, yellow: prediction, green: ground truth; a video animation is included in the Supplementary Material

Quantitative Results

- Detailed results table in full paper
- Results show reduction in error measures on
  - Average Displacement Error of waypoints
  - Final Displacement Error of final trajectory location
  - Mean Squared Error of velocities
- Maximum acceleration of our approach: $12.2\ m/s^2$
- Maximum acceleration of ground truth (after denoising): $14.5\ m/s^2$
- Maximum acceleration of baseline (MACRO VRNN): $54.9\ m/s^2$

Thank you for your attention!

TEMPLE UNIVERSITY