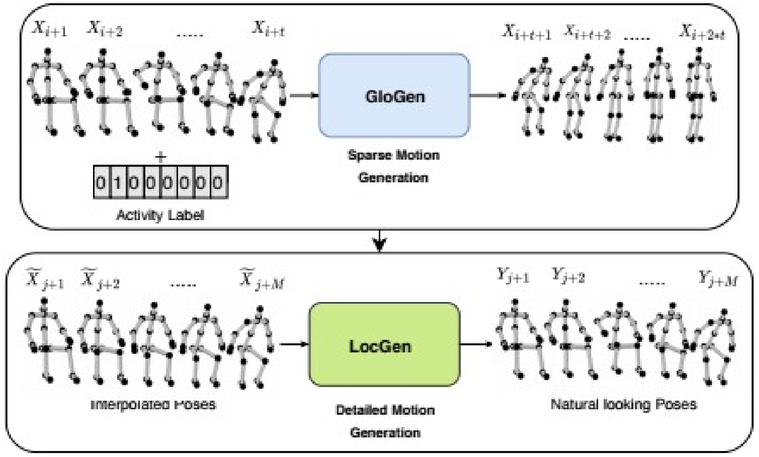
INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD
R&D SHOWCASE 2021
Technology, Social Impact

GlocalNet: Class-aware Long-term Human Motion Synthesis

Motivation

- **Aim** - Synthesis of long-term (> 6000 ms) human motion skeleton sequences across a large variety of human activity classes (> 50) to aid human-centric video generation.
- **Applications** - Augmented Reality, 3D character animations, pedestrian trajectory prediction...
- **Challanges** - long-term temporal dependencies among poses, cyclic repetition across poses, bi-directional and multi-scale dependencies among poses, variable speed of actions, and a large as well as partially overlapping space of temporal pose variations across multiple class/types of human activities.



Overview of our two-stage framework, GlocalNet.

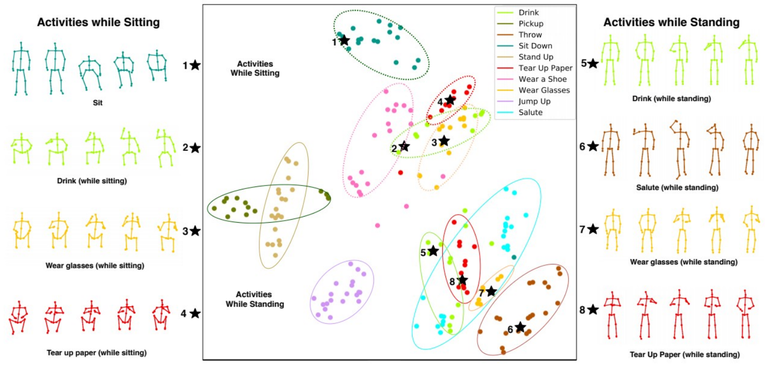
GlocalNet Architecture

- In the first stage, GloGen generates the sparse motion trajectory of an activity, followed by the second stage, LocGen, that predicts the dense poses from the generated sparse motion.
- Loss function with Joint Loss (L_J) and Motion Flow Loss (L_{MF}) as:

$$L = (\lambda_1 * L_J) + (\lambda_2 * L_{MF})$$

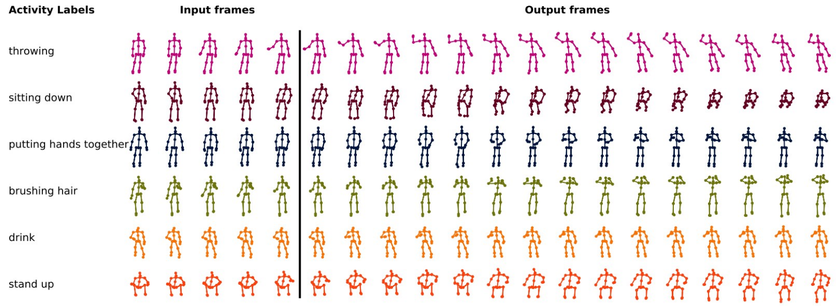
$$L_J = \sum_{i=1}^{t} \left\| X[i] - \hat{X}[i] \right\|_2 \qquad L_{MF} = \sum_{i=1}^{t-1} \left\| V[i] - \hat{V}[i] \right\|_2$$

$$\hat{V}_i = \hat{X}_{i+1} - \hat{X}_i$$



The t-SNE plot of GloGen embedding subspace along with the plot of selected motion trajectories where multiple samples for different classes are represented as color-coded 3D points.

Results



Output of GloGen using different activity labels and initial poses.

Models	cross-view		cross-subject	
	MMD_{avg}	MMD_{seq}	MMD_{avg}	MMD_{seq}
SkeletonVAE	1.079	1.205	0.992	1.136
SkeletonGAN	0.999	1.311	0.698	0.788
c-SkeletonGAN	0.371	0.398	0.338	0.402
SAGCN	0.316	0.335	0.285	0.299
Ours (L_J)	0.213	0.218	0.201	0.212
Ours (L_{MF})	0.646	0.647	0.601	0.625
Ours (L_J + L_{MF})	**0.195**	**0.197**	**0.177**	**0.187**

Comparison of Our Method (GloGen) in terms of MMD on NTU RGB+D (2D)

Authors: Neeraj Battan, Yudhik Agrawal, Sai Soorya Rao, Aman Goel, and Avinash Sharma
Research Center Name: Centre for Visual Information Technology

Publication: Battan, Neeraj, et al. "GlocalNet: Class-aware Long-term Human Motion Synthesis." IEEE/CVF Winter Conference on Applications of Computer Vision. 2021.