NeuroAI

[Approaches to Analyze Behavior]
Sparse Semi-Supervised Action Recognition with Active Learning

Presenter: Jingyuan Li {jingyli6@uw.edu},
Faculty Advisor: Eli Shlizerman {shlizee@uw.edu}
NeuroAI Shlizerman Lab
Electrical and Computer Engineering Department

Background and Motivation

Action Recognition

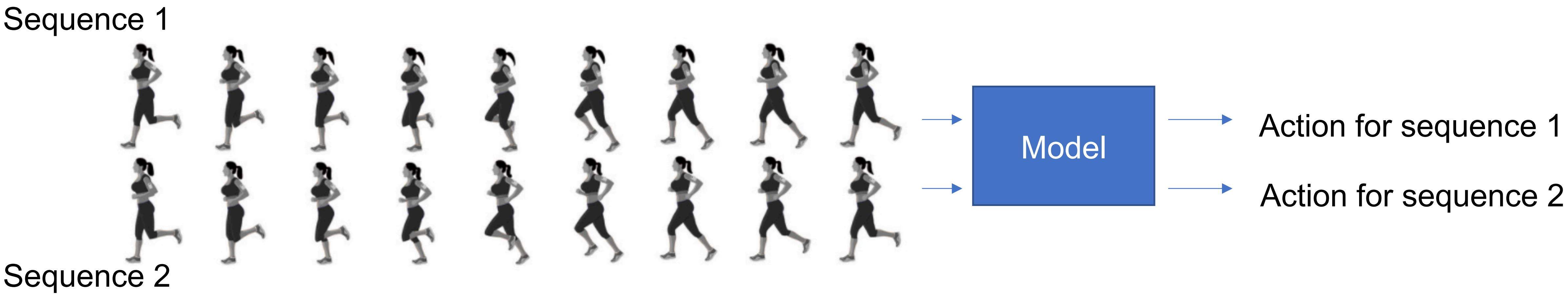


Image Source: https://www.vectorstock.com/royalty-free-vector/girl-run-cycle-animation-sequence-loop-animation-vector-26206917

Application of Action Recognition

Robotic Surveillances

Wildlife Monitoring

Action Evaluation

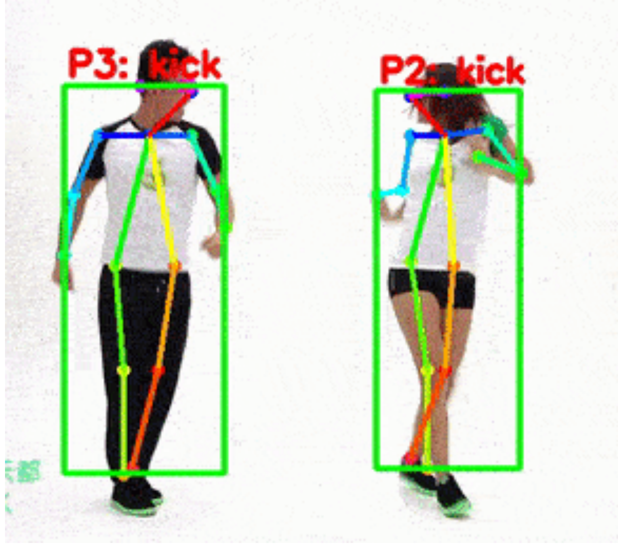
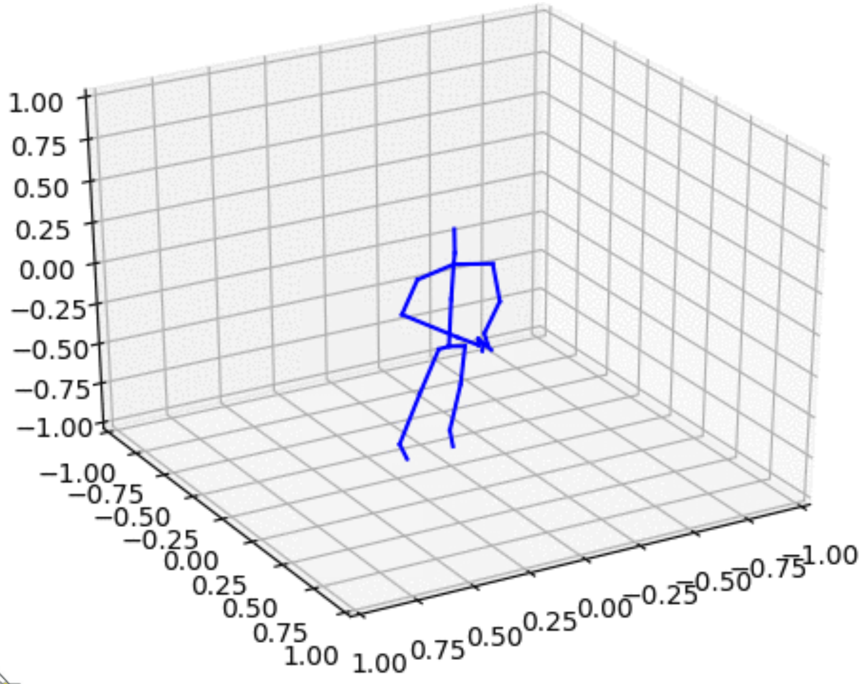
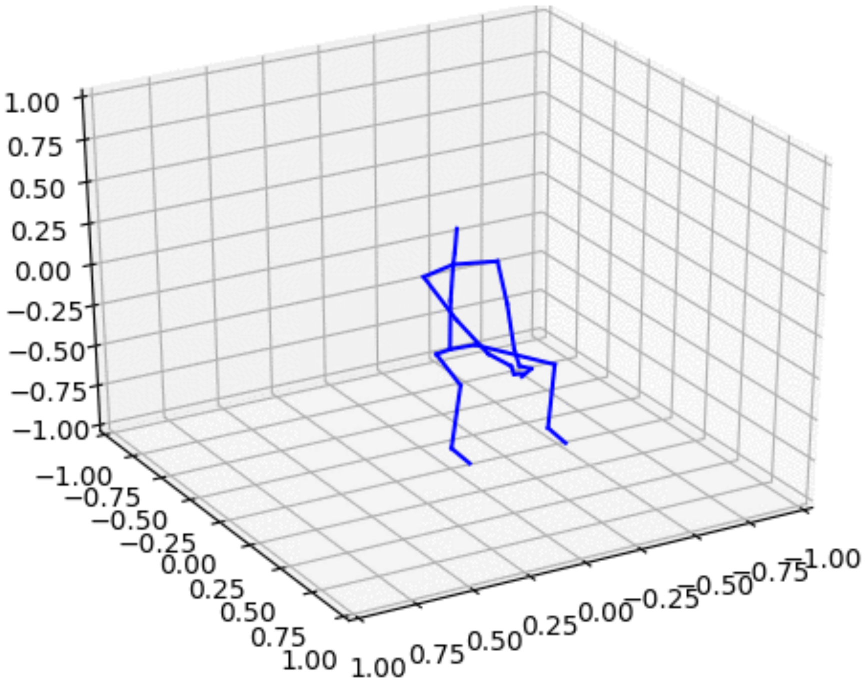


Image Source:
Left: Boston Dynamics
Middle: https://www.conservationnw.org/our-work/wildlife/wildlife-monitoring/
Right: https://pythonawesome.com/multi-person-real-time-action-recognition-based-on-human-skeleton/

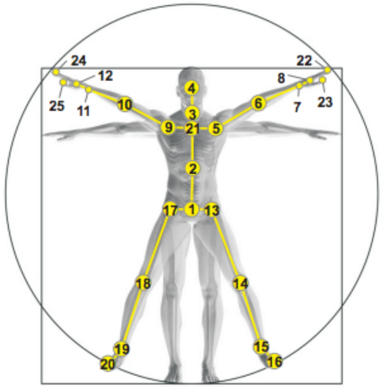
Skeleton Based Action Recognition



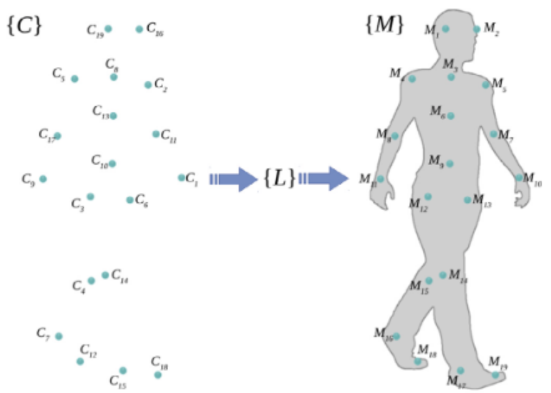
Drinking Water



Standing Up



Skeleton Data Sources

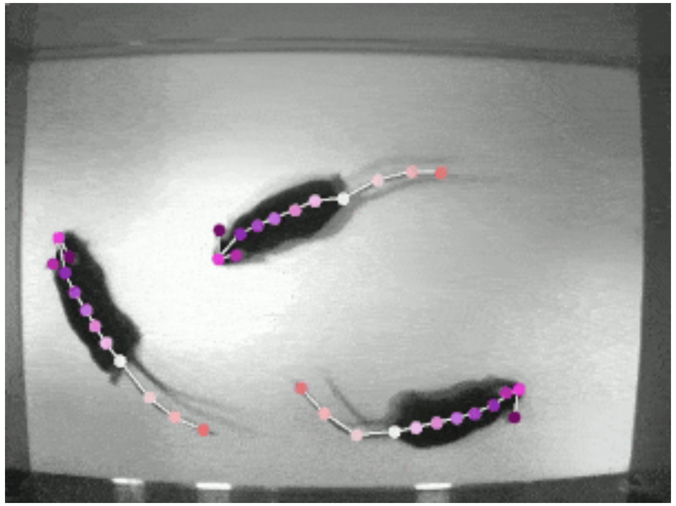


Motion Tracking with Marker

J.L. Jimenez Bascones et al., 2019

Human motion video with skeleton

Zhe Cao et al., 2017



Animal motion video with skeleton

Alexander Mathis et al., 2018

Advantageous for skeleton based action recognition:

- Reducing interference from the background
- Learning dynamics movement

Earlier Methods on Prepared Dataset

- Supervised Learning:
  -> Achieving high performance with a large labeled dataset.

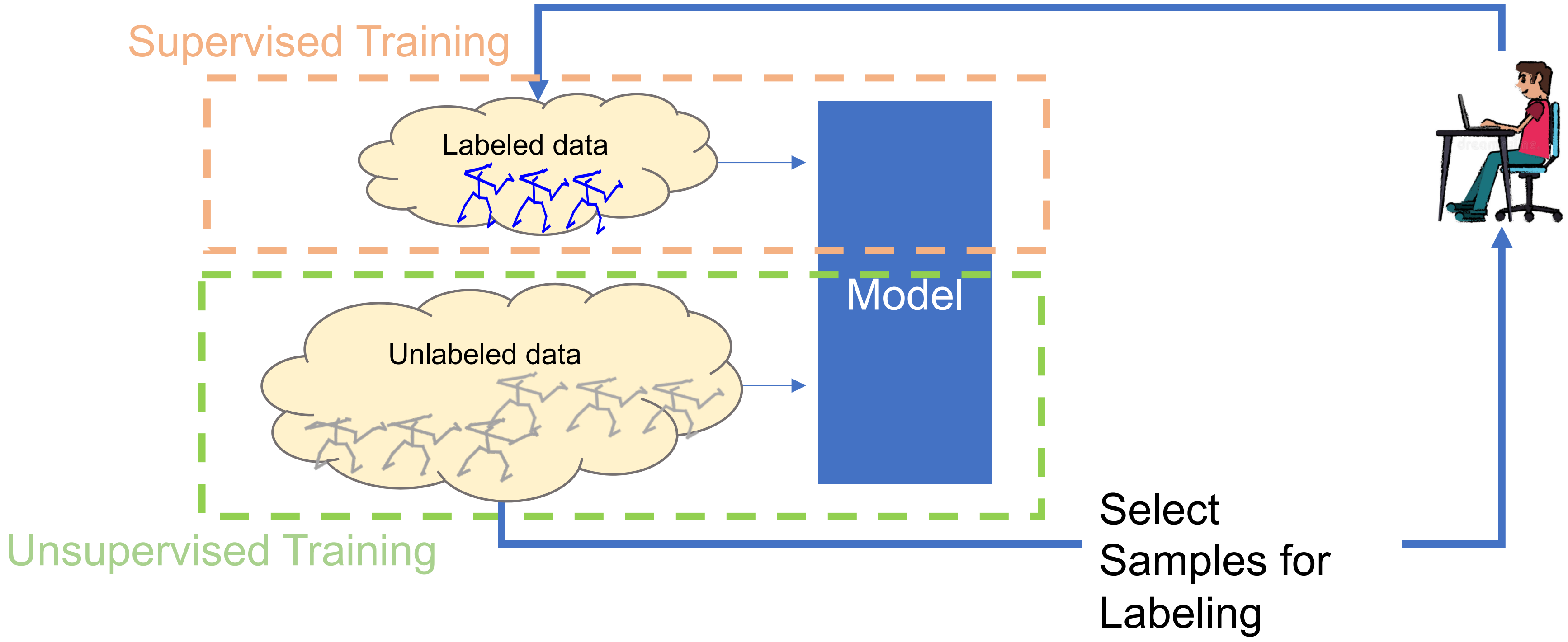
- Semi-Supervised Learning:
  -> Achieving reasonable performance with partially randomly labeled data.

No consideration of which samples should be labeled.

Tackle the problem with Active Learning

What is Active Learning (AL)

“Active learning is a special case of machine learning in which a learning algorithm can interactively query a user (or some other information source) to label new data points with the desired outputs”



Methods

1. ‘Cold Start’ in AL
2. Semi-Supervised Training for Action Recognition
3. Selection with Partially Labeled Dataset

Methods

1. ‘Cold Start’ in AL
2. Semi-Supervised Training for Action Recognition
3. Selection with Partially Labeled Dataset

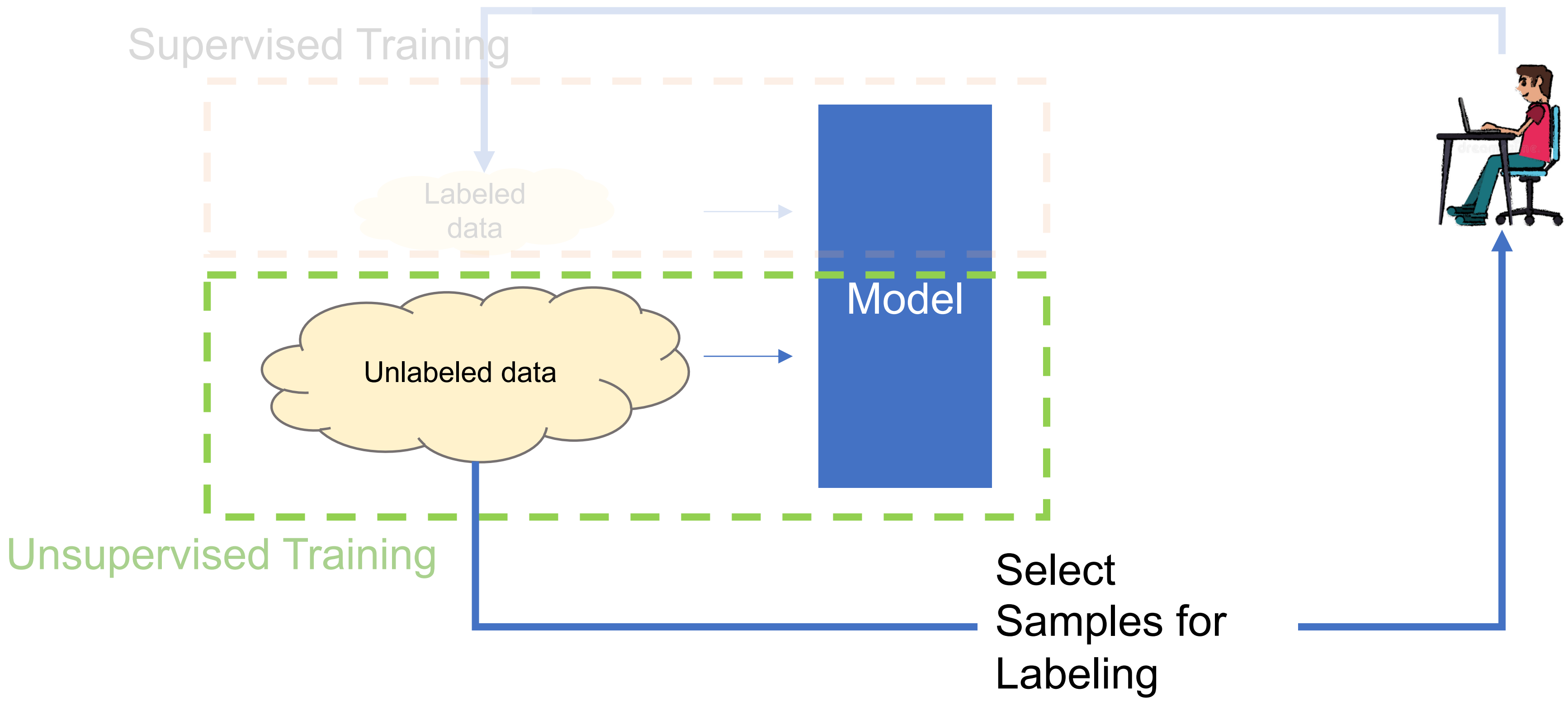
‘Cold Start’ in Initial Selection

- None of the samples is labeled at initial selection.

‘Cold Start’ in Initial Selection

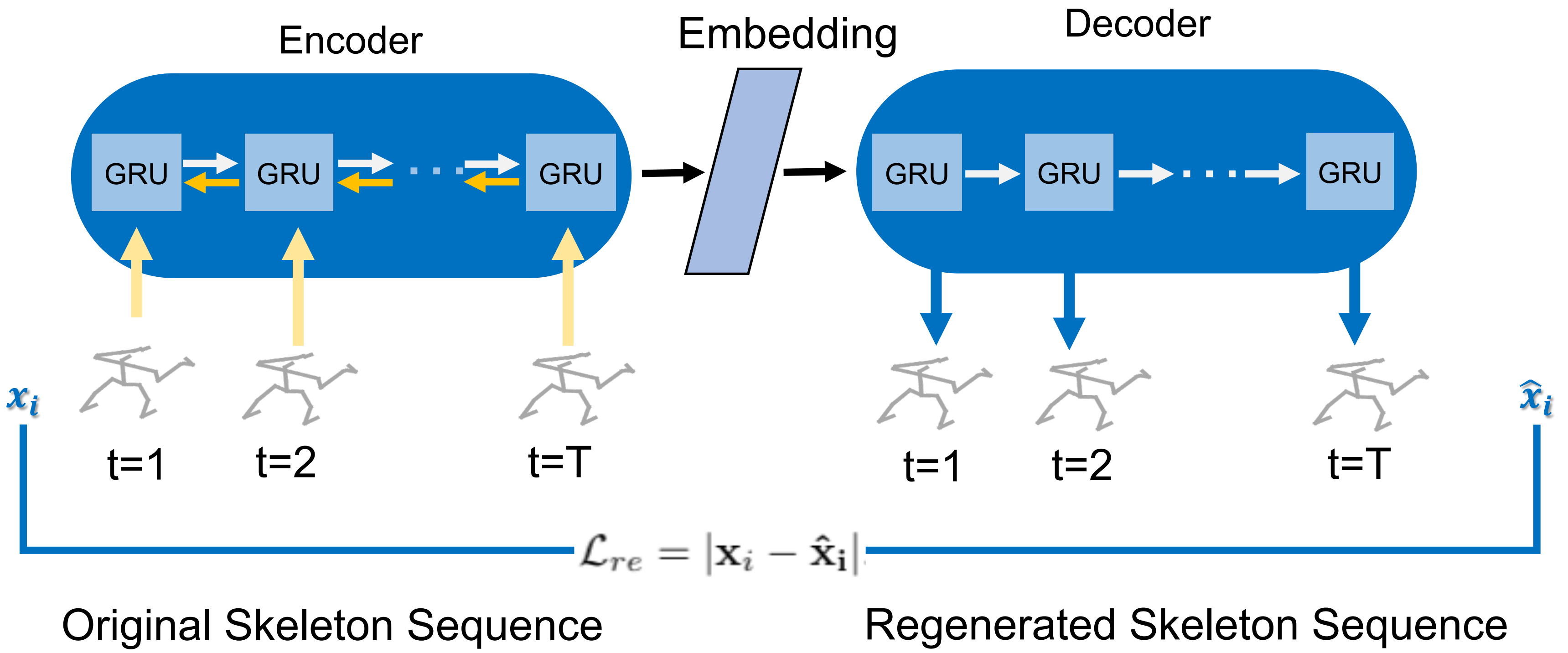
- None of the samples is labeled at initial selection.
- Quality of initial selection affects future selection.
    - Bad initialization can lead to wrong class boundary.

‘Cold Start’ in Initial Selection



Solve the ‘Cold Start’ Problem

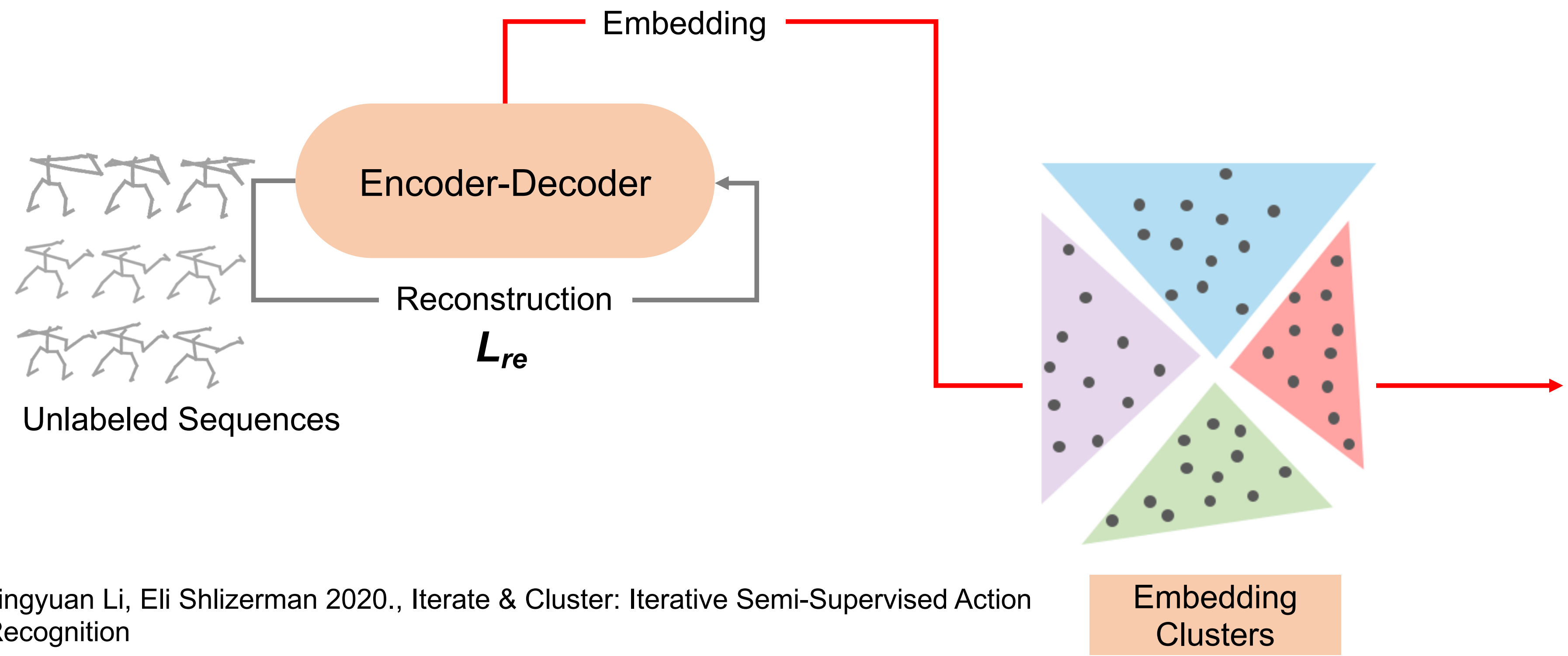
- Embeddings from unsupervised regeneration shown to form meaningful embedding space



Kun Su, Xiulong Liu, Eli Shlizerman 2020., PREDICT & CLUSTER: Unsupervised Skeleton Based Action Recognition

Solve the 'Cold Start' Problem

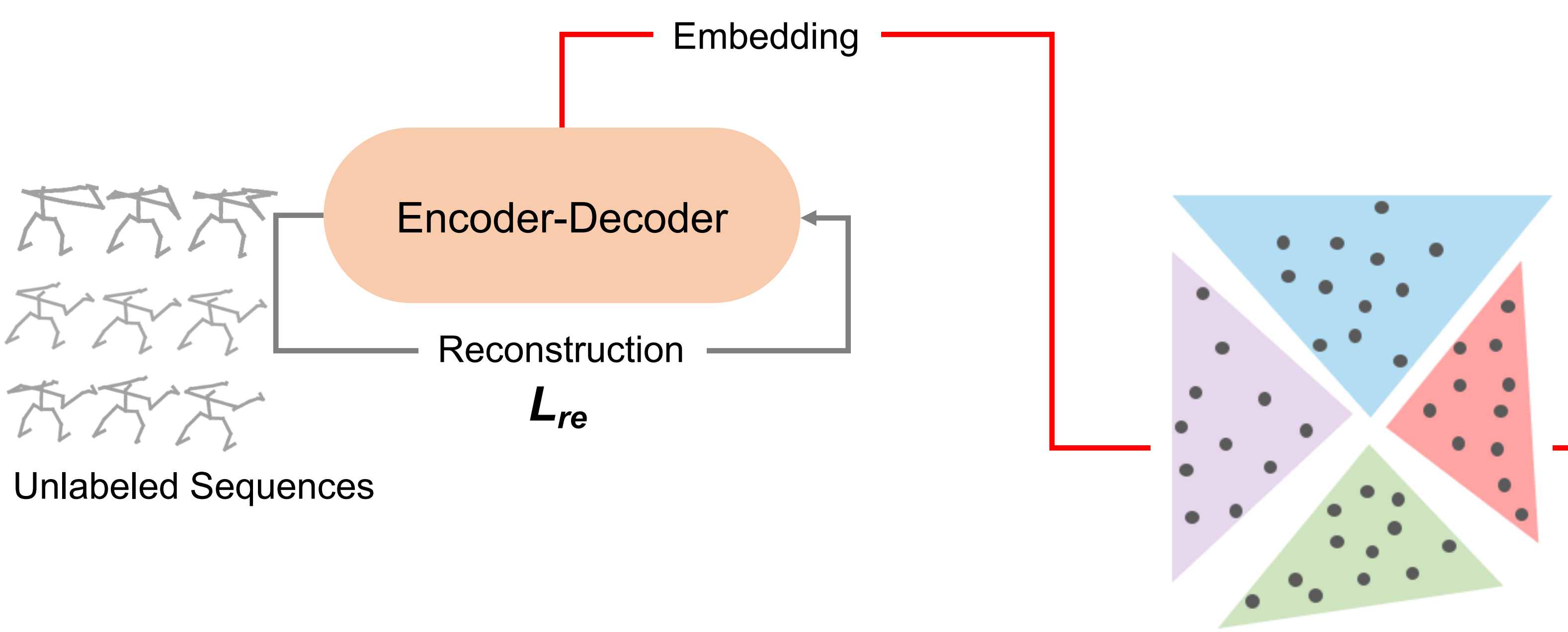
- Clustering samples in the embedding space

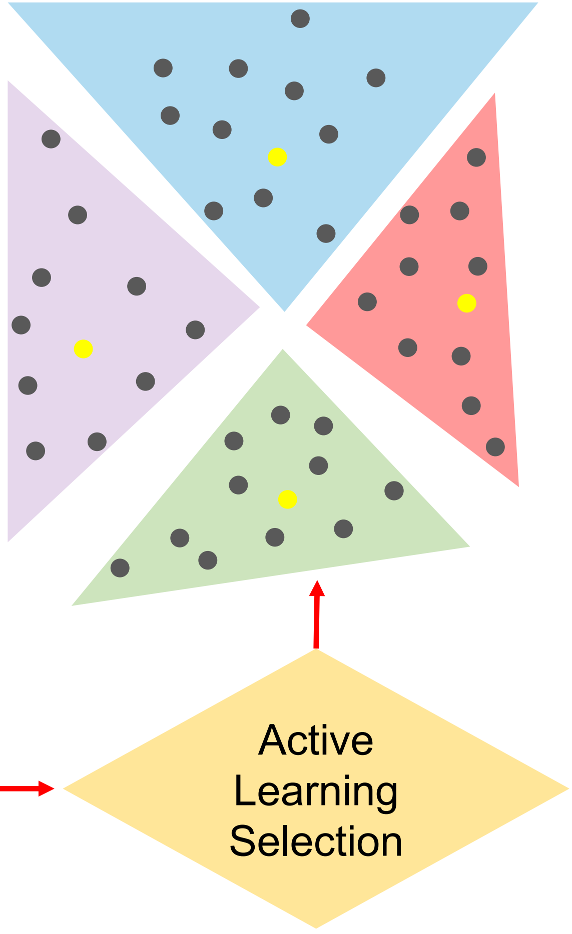


Jingyuan Li, Eli Shlizerman 2020., Iterate & Cluster: Iterative Semi-Supervised Action Recognition

Solve the ‘Cold Start’ Problem

- Clustering samples in the embedding space
- Selecting samples closest to cluster centers



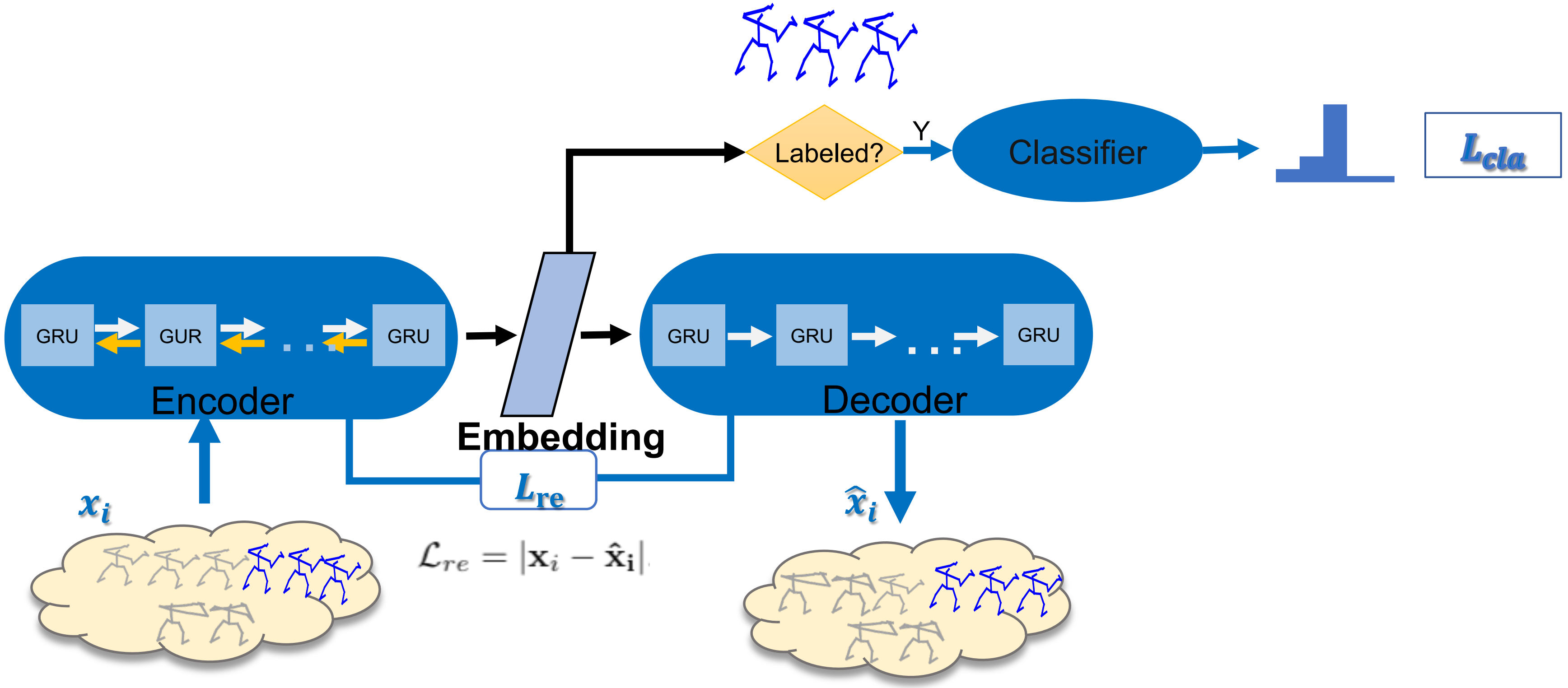


Jingyuan Li, Eli Shlizerman 2020., Iterate & Cluster: Iterative Semi-Supervised Action Recognition

Methods

1. 'Cold Start' in AL
2. Semi-Supervised Training for Action Recognition
3. Selection with Partially Labeled Dataset

Semi-Supervised Learning with Partially Labeled Samples

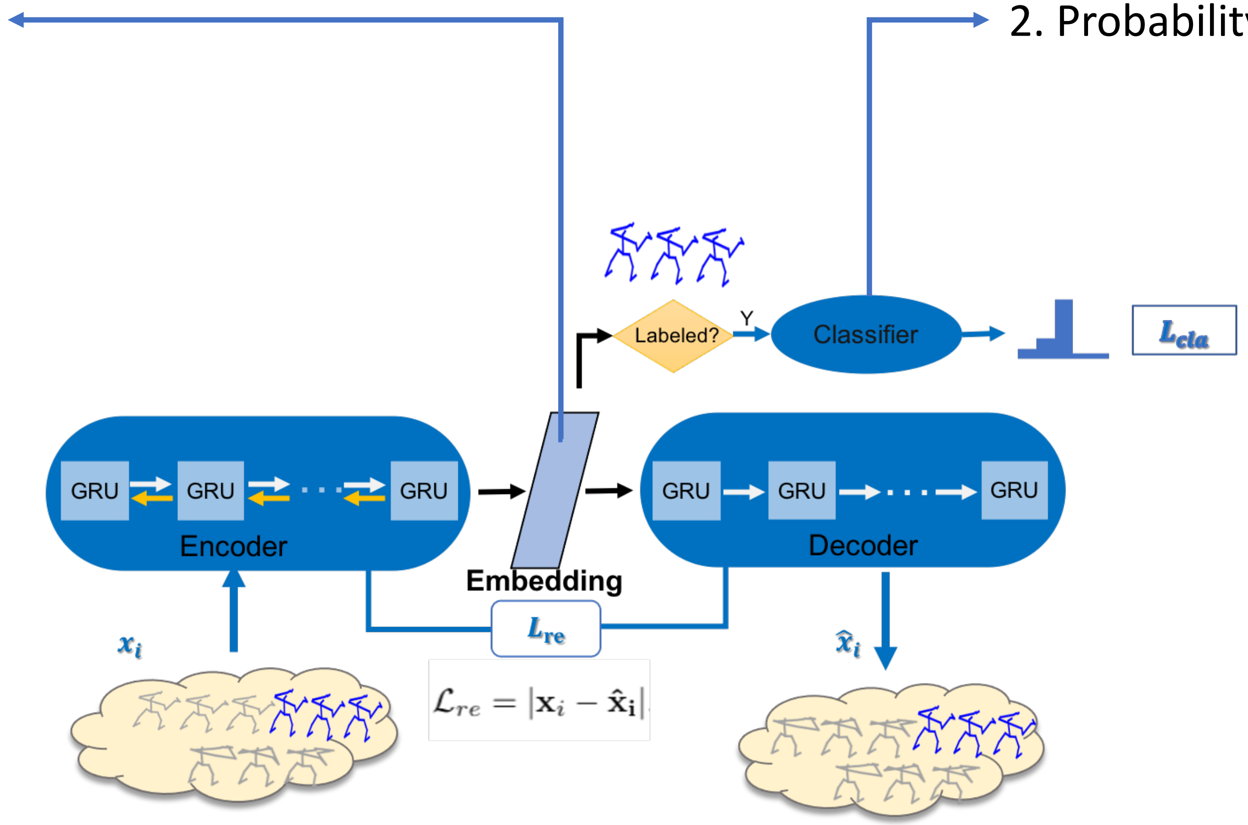


Training with loss: L_{re}+L_{cla}

Two Representations

1. Embedding from encoder

2. Probability distribution



Methods

1. ‘Cold Start’ in AL
2. Semi-Supervised Training for Action Recognition
3. Selection with Partially Labeled Dataset (Distance & Uncertainty)

Selecting According to Distance

- Step 1: Form clusters in embedding space.

Selecting According to Distance

- Step 1: Form clusters in embedding space.
- Step 2: Measure the **distance** of unlabeled samples to the labeled samples **within the same cluster** with predicted class probability.

Selecting According to Distance

- Step 1: Form clusters in embedding space.
- Step 2: Measure the **distance** of unlabeled samples to the labeled samples **within the same cluster** with predicted class probability.
- Step 3: For each cluster, select a subset of samples with the largest distance.

Jingyuan Li, Eli Shlizerman 2020, Sparse Semi-Supervised Action Recognition with Active Learning

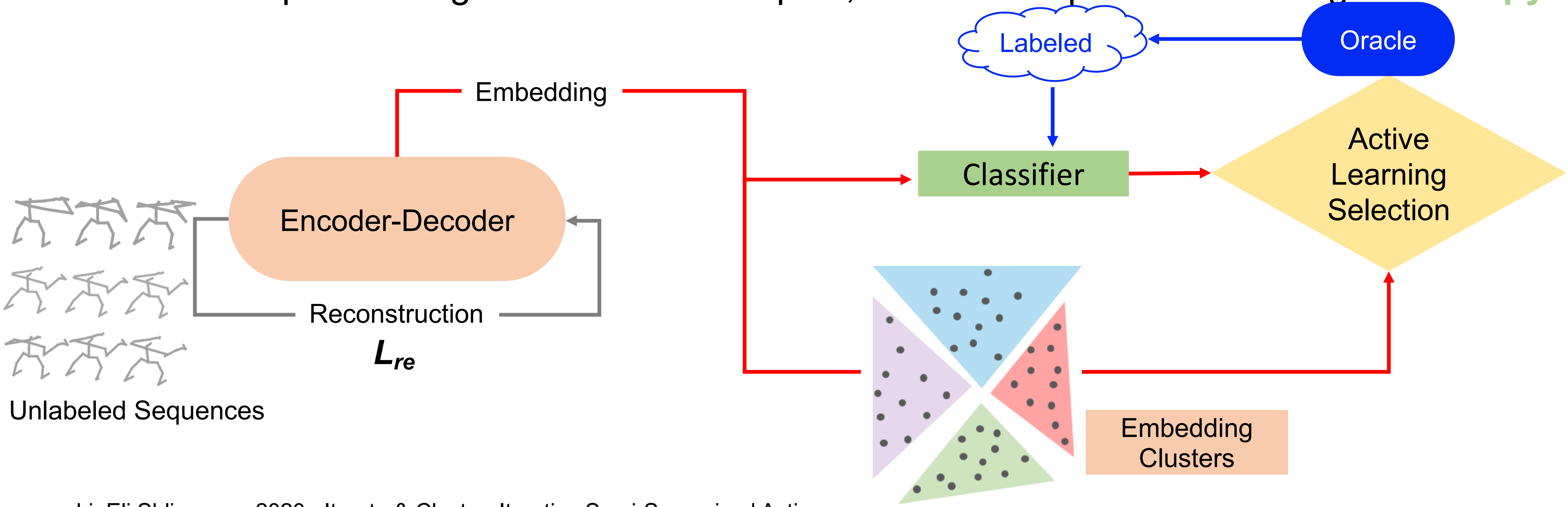
Selecting According to Uncertainty

- Step 1: Form clusters in embedding space.
- Step 2: Measure the **distance** of unlabeled samples to the labeled samples **within the same cluster** with predicted class probability.
- Step 3: For each cluster, select a subset of samples with largest distance.
- Step 4: Among the subset of samples, choose samples with the largest **entropy**.

Jingyuan Li, Eli Shlizerman 2020, Sparse Semi-Supervised Action Recognition with Active Learning

Selection According to Distance & Uncertainty

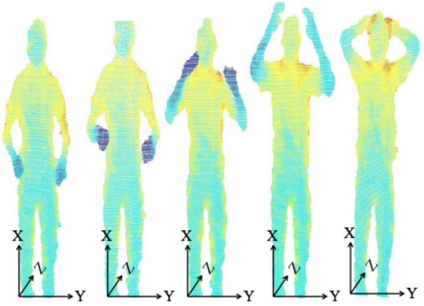
- Step 1: Form clusters in embedding space.
- Step 2: Measure distance of unlabeled samples to the labeled samples **within the same cluster** with predicted class probability.
- Step 3: For each cluster, select a subset of samples with largest distance.
- Step 4: Among the subset of samples, choose samples with the largest entropy.



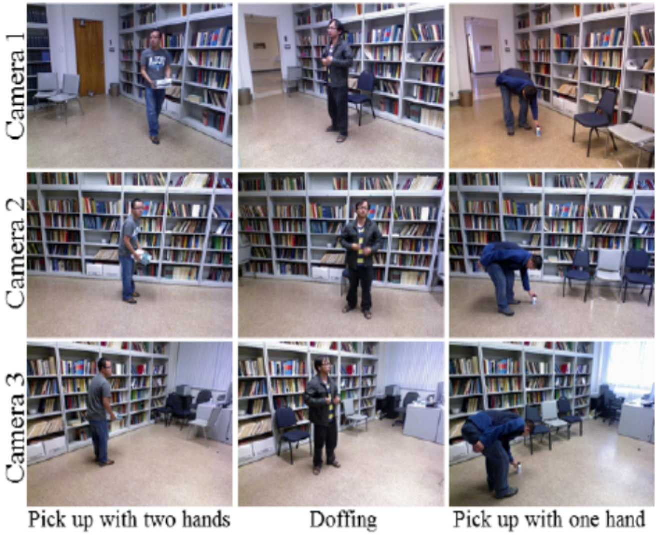
Jingyuan Li, Eli Shlizerman 2020., Iterate & Cluster: Iterative Semi-Supervised Action Recognition

Experiments

Datasets



UWA3D
30 Classes



NW-UCLA
10 Classes

NTU RGBD
60 Classes

Compare with Semi-supervised and other AL Methods

UWA3D VIEW3					
	% Labels #Labels	5% 25	10% 50	20% 100	50% 250
SSL	C	18.3	21.9	32.1	44.3
	RC	19.5	30.0	26.9	46.3
	IRC	20.0	36.4	37.6	51.1
AL(our)	SESAR-DIS	18.5	29.2	40.9	55.6
	SESAR-U	21.8	31.3	41.0	55.8
	SESAR-CS	26.9	**37.1**	41.2	55.8
AL+K(our)	SESAR-KT	22.8	34.6	**51.8**	58.8
	SESAR-KJS	**28.3**	36.0	49.5	**59.5**

Table 1. Performance of different semi-supervised approaches (top), SESAR with STOA AL methods (middle) SESAR with AL+K methods (bottom) on UWA3D dataset.

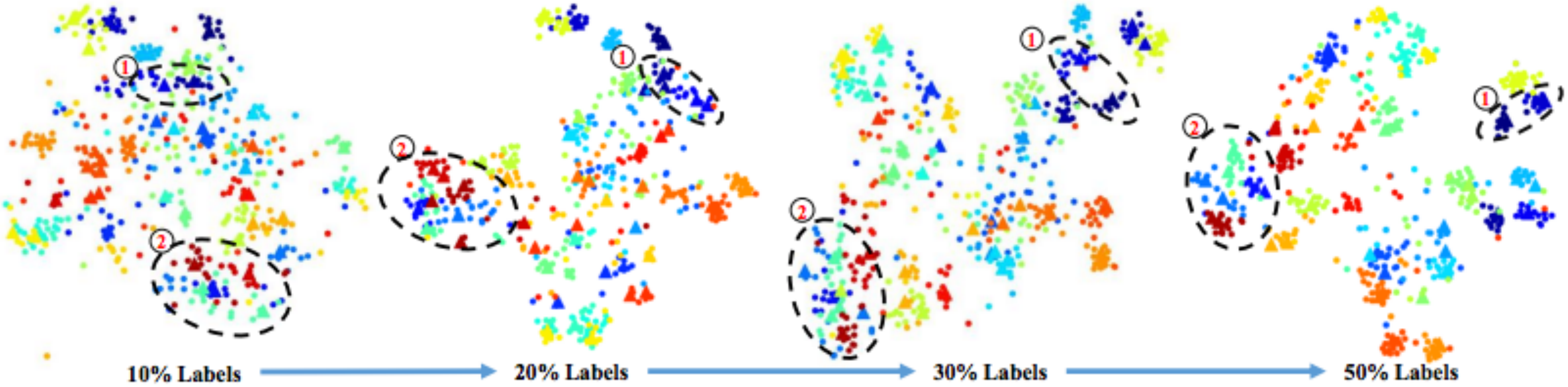
NW-UCLA					
	% Labels # Labels	5% 50	15% 150	30% 300	40% 400
SSL	C	44.6	56.0	70.9	72.9
	RC	51.5	63.1	77.0	76.7
	ASSL[26]	52.6	74.8	78.0	78.4
	RIC	56.2	70.5	73.5	81.0
	MS^2L[14]	–	60.5 –	–	–
AL(our)	SESAR-DIS	55.3	73.3	**80.3**	83.3
	SESAR-U	62.7	74.0	77.9	80.3
	SESAR-CS	**63.9**	71.5	77.5	82.3
AL+K(our)	SESAR-KJS	58.1	76.6	**80.0**	**85.0**
	SESAR-KT	**63.6**	**76.8**	77.2	78.9

Table 2. Performance of different semi-supervised approaches (top), SESAR with STOA AL methods (middle) SESAR with AL+K methods (bottom) on NW-UCLA dataset.

NTU RGB+D 60 Cross Subject					
	% Labels # Labels	1% 400	2% 800	5% $2K$	10% $4K$
SSL	C	21.8	37.2	49.6	56.7
	MS^2L[14]	33.1	–	–	**65.2**
	RC	33.8	41.6	47.8	60.0
	IRC	36.7	42.7	53.9	61.2
	ASSL[26]	–	–	57.3	64.3
AL(our)	SESAR-CS	17.6	23.1	37.0	49.6
	SESAR-DIS	34.9	39.5	53.8	60.4
	SESAR-U	36.1	42.5	53.9	60.8
AL+K(our)	SESAR-KJS	38.2	45.0	**57.8**	62.9
	SESAR-KT	**41.8**	**46.1**	55.0	58.2

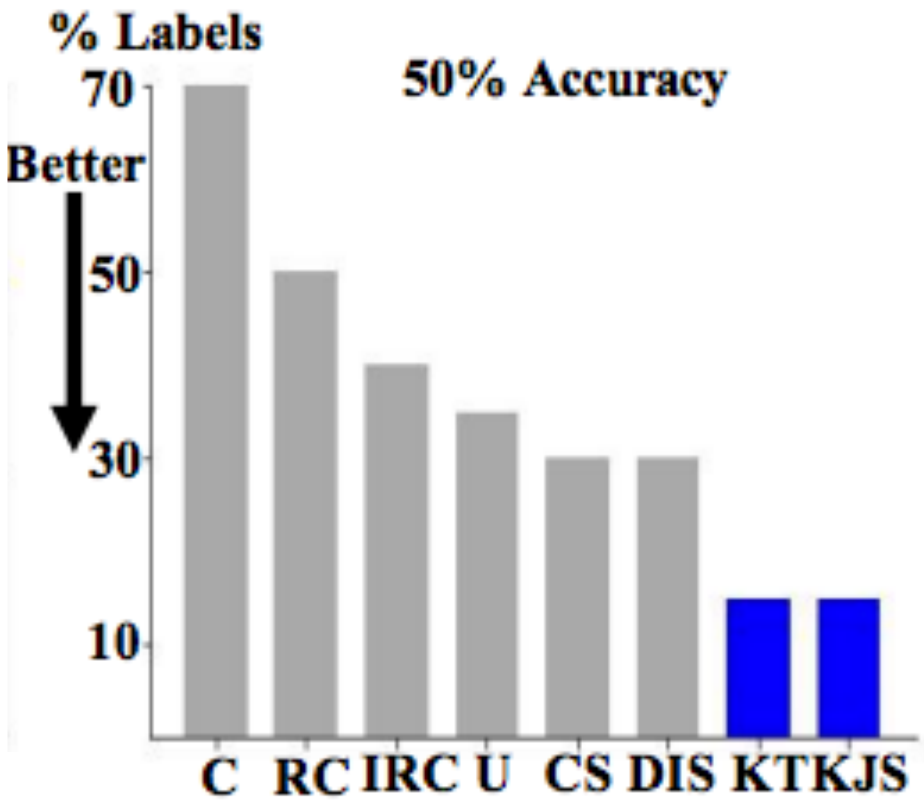
Table 3. Performance of different semi-supervised approaches (top), SESAR with STOA AL methods (middle) SESAR with AL+K methods (bottom) on NTU RGB+D 60 dataset.

Learning to Form Clearer Clusters



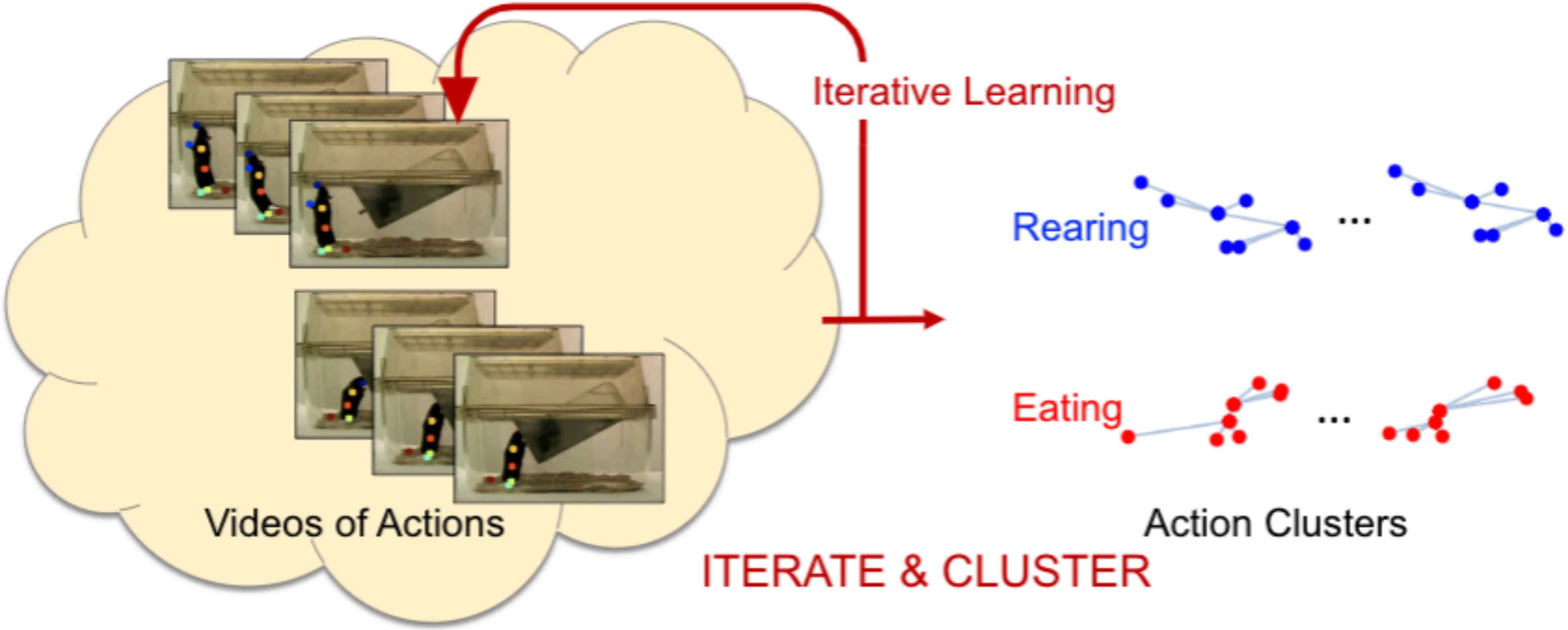
2D latent embedding (with t-SNE) for UWA-3D dataset across three training iterations.

Label Efficiency

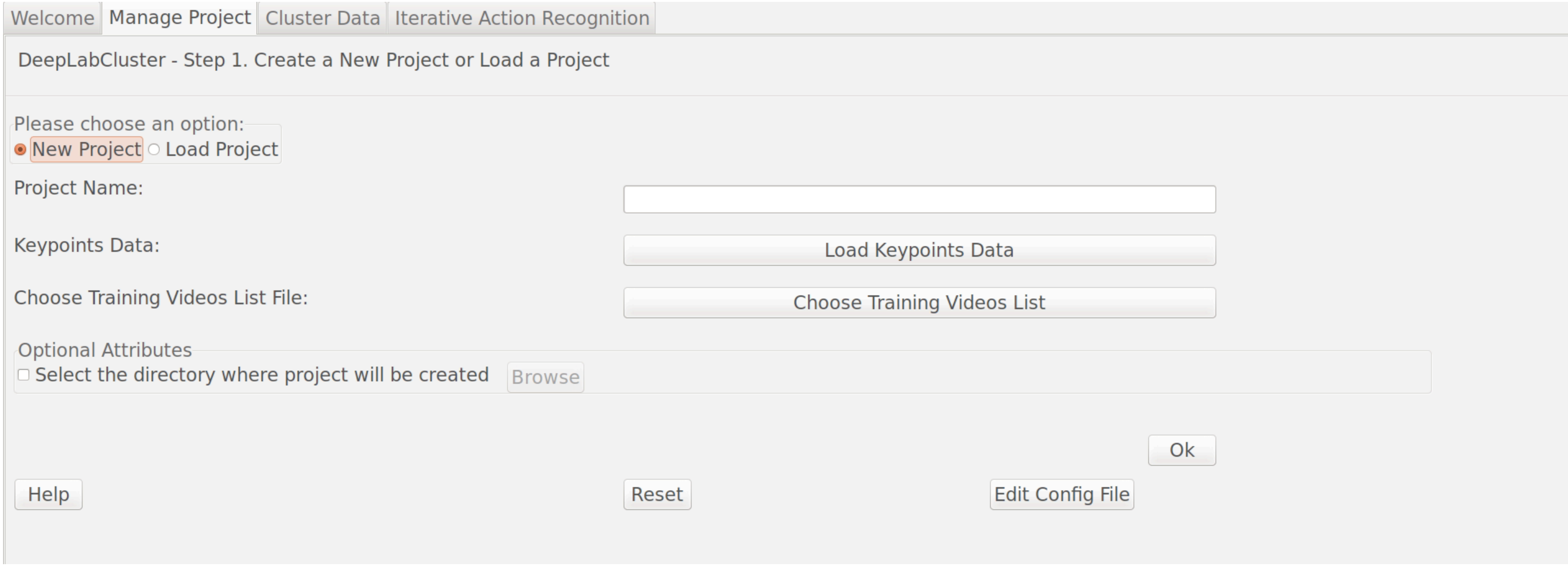


Label requirement to achieve 80% accuracy on UWA-3D dataset.

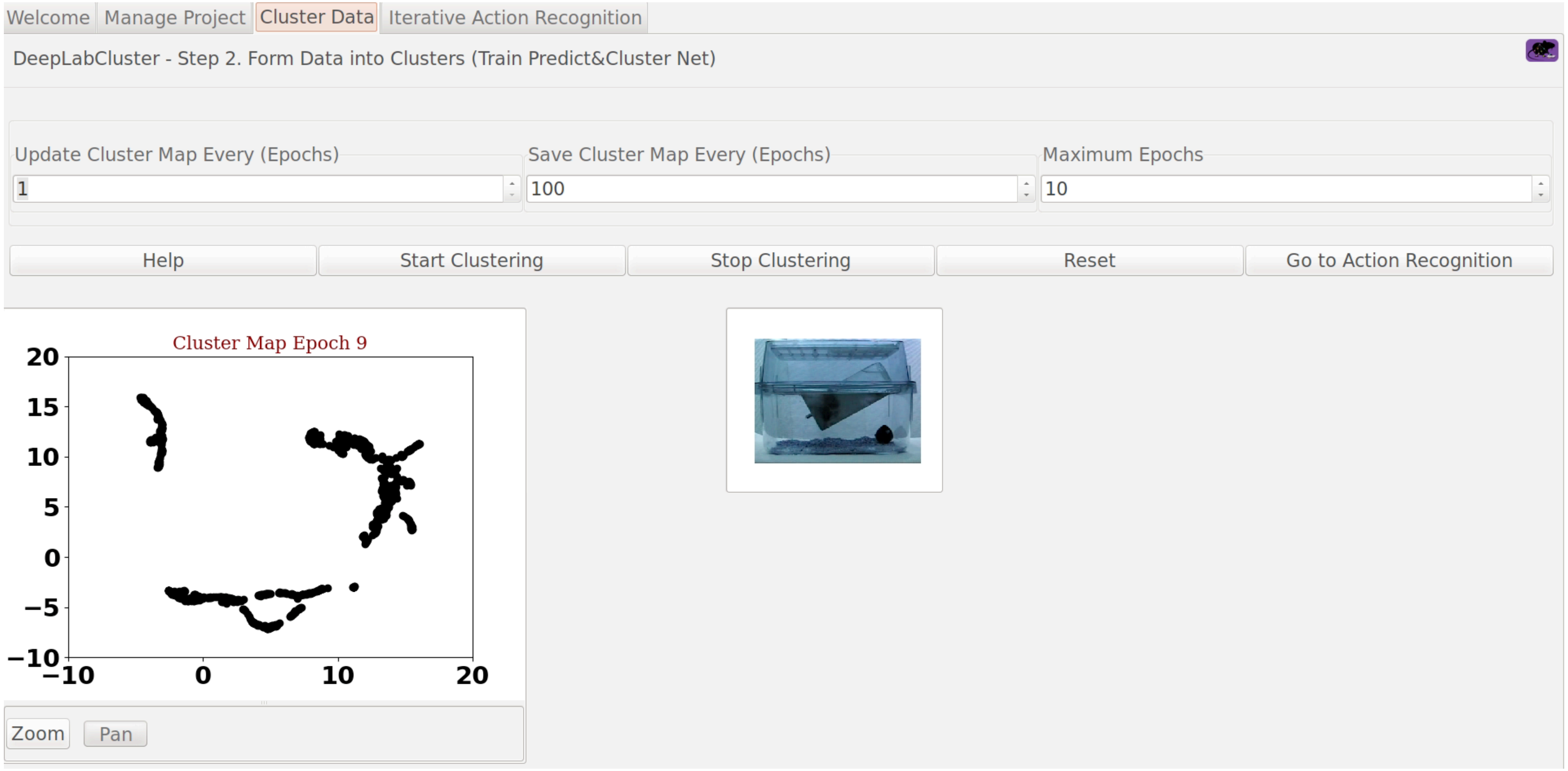
Application: GUI DeepLabCluster



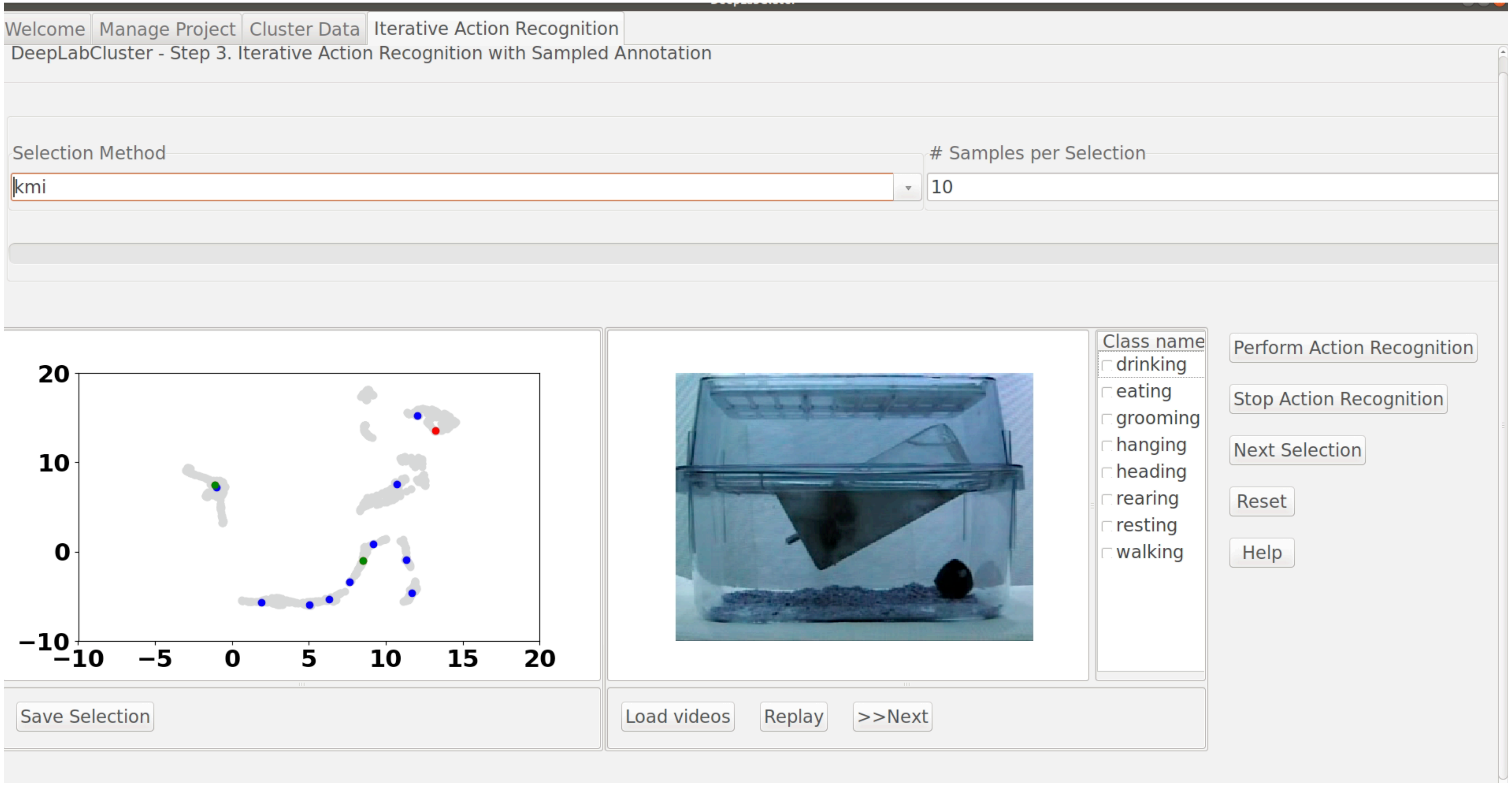
Application



Application



Application



Thanks!