ERA: Expert Retrieval and Assembly for Early Action Prediction

Lin Geng Foo[1]*, Tianjiao Li[1]*, Hossein Rahmani[2], Qiuhong Ke[3], and Jun Liu[1]

[1] ISTD Pillar, Singapore University of Technology and Design

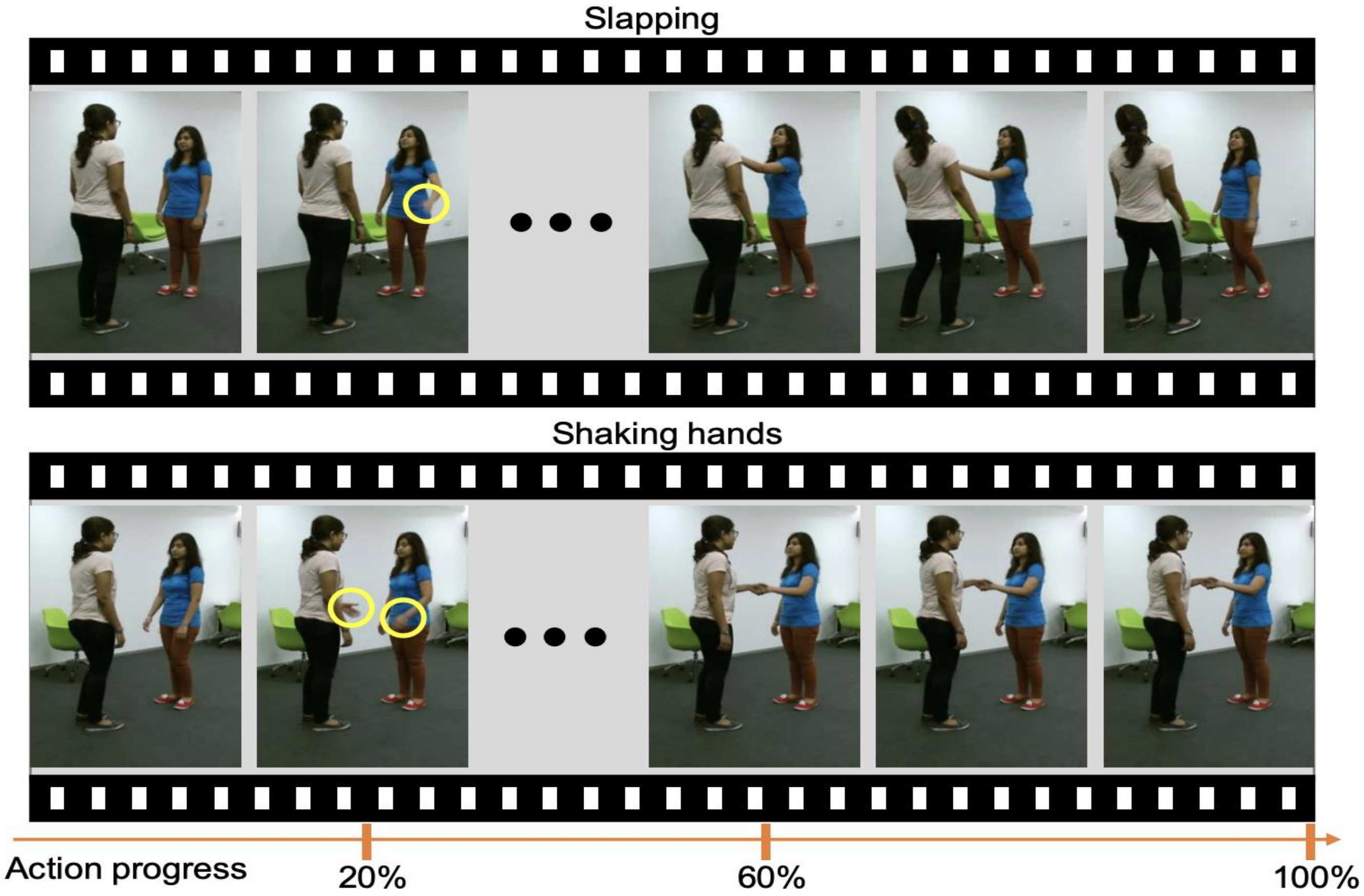
[2] School of Computing and Communications, Lancaster University

[3] Department of Data Science & AI, Monash University

* Equal Contribution

Early Action Prediction

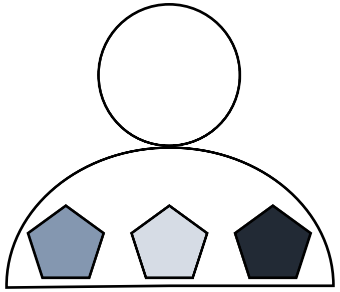
- Early Action Prediction is where we try to recognize the human action at the very early stage.



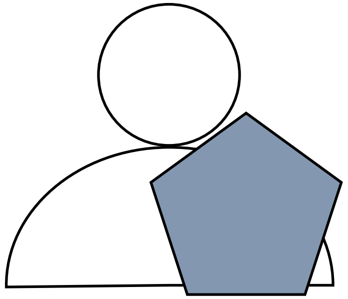
At the beginning stages of actions, there are often only subtle cues for action recognition.

Dynamic Networks for Expert Specialization

Non-Expert



Expert 1

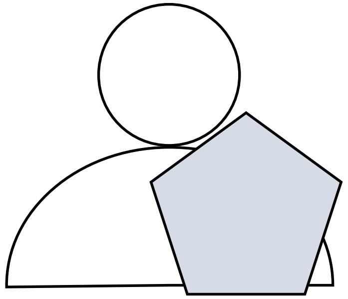


Slapping

Shaking Hands

Pointing

Expert 2

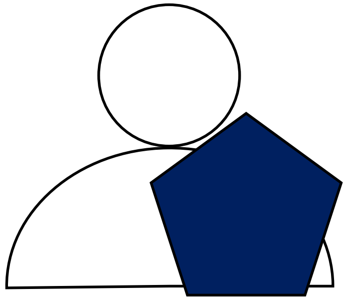


Jumping

Hopping

Squatting

Expert 3



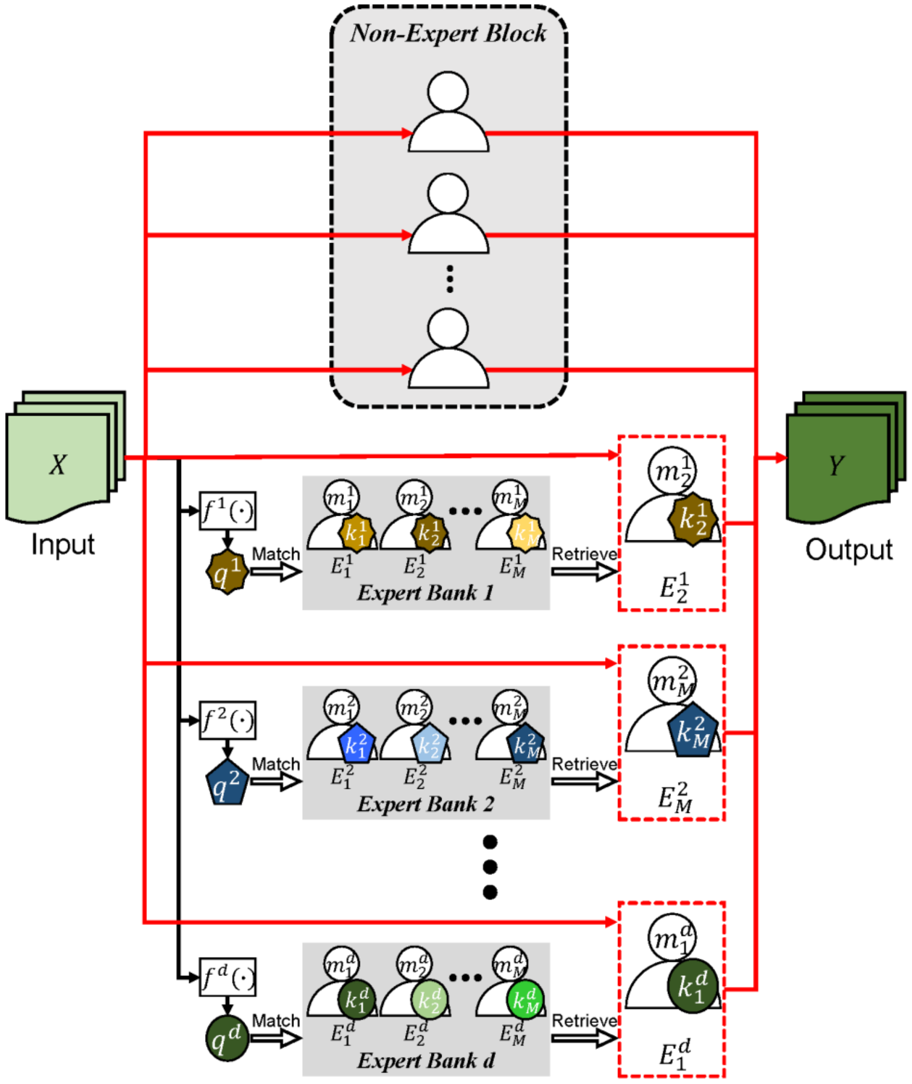
Eat meal

Drink water

Expert Retrieval and Assembly (ERA)

We design a new module that can replace convolutional layers in Convolutional Neural Networks (CNNs).

Our module **retrieves and assembles a set of *experts* most specialized at using discriminative subtle differences**, to distinguish an input sample from other highly similar samples.



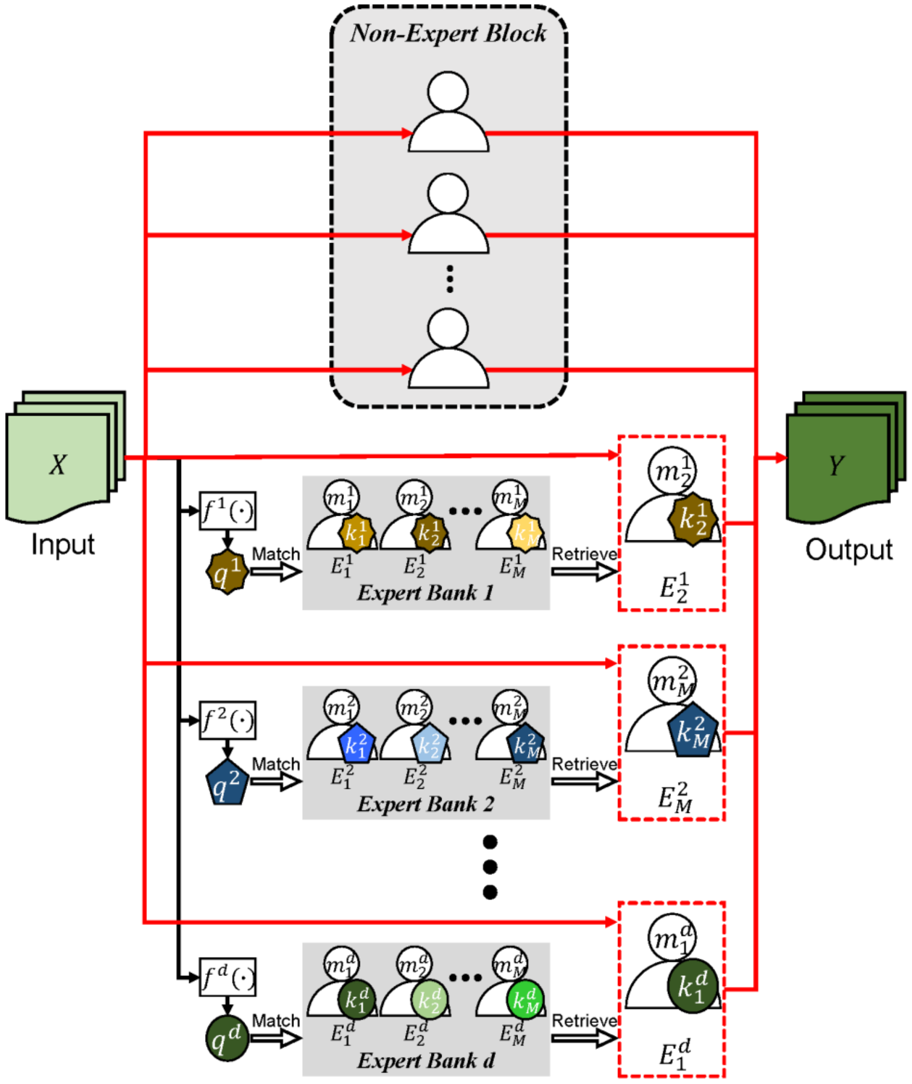
ERA Module Design

Highlights of the Design

- Both *non-experts* and *experts* are used. Each *expert/non-expert* outputs a channel

- Conditioned on the input, an *expert* is retrieved from each *Expert Bank,* while other experts are not used

- Retrieval is done using *Key-Query Mechanism* to select similar *experts* for similar inputs, such that *experts* specialize in distinguishing using subtle cues

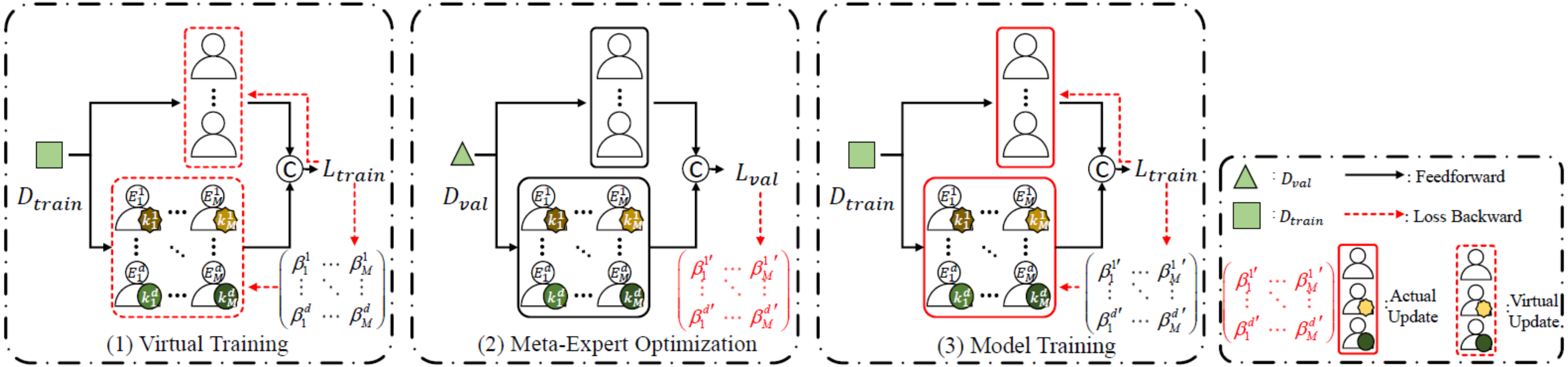


Expert Learning Rate Optimization (ELRO)

Motivation for New Training Scheme

Moreover, it is non-trivial to balance the training among the many different *experts* in the module.

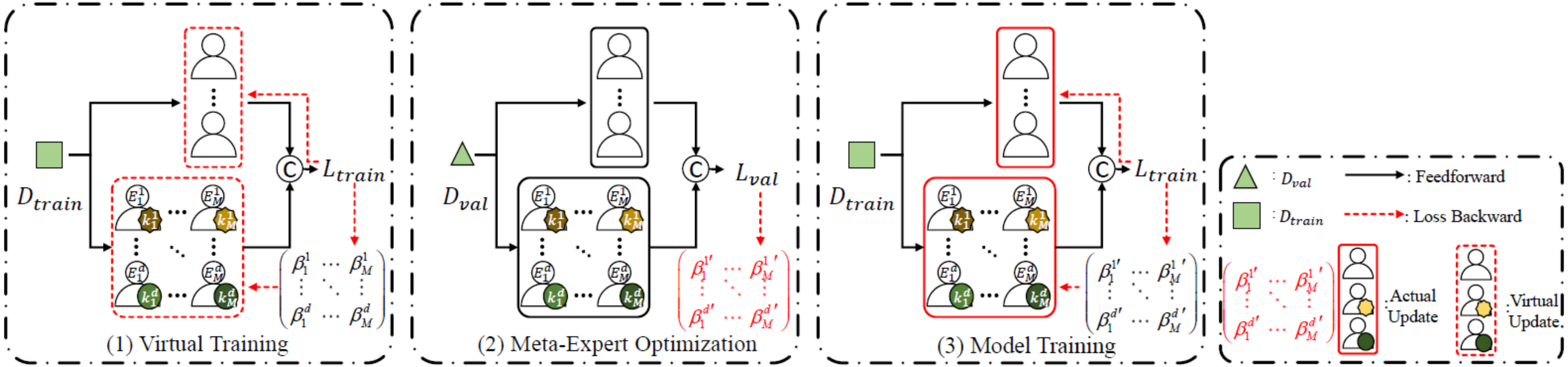
For instance, as some subtle cues may be more common, a few *experts* are selected more often and might be better trained.



Expert Learning Rate Optimization (ELRO)

Thus, we introduce **individual expert learning rates** for more balanced training of our experts. **ELRO is implemented** during the training of the *experts*, which tunes their individual learning rates together with the rest of the model parameters.

Overall, ELRO is **a 3-step procedure** shown below:



Experiments

Table 1: Results on NTU60 and SYSU

Methods	Observation Ratios on NTU60						Observation Ratios on SYSU					
	20%	40%	60%	80%	100%	AUC	20%	40%	60%	80%	100%	AUC
Jain *et al.* [18]	7.07	18.98	44.55	63.84	71.09	37.38	31.61	53.37	68.71	73.96	75.53	57.23
Ke *et al.* [21]	8.34	26.97	56.78	75.13	80.43	45.63	26.76	52.86	72.32	79.40	80.71	58.89
Kong *et al.* [26]	-	-	-	-	-	-	51.75	58.83	67.17	73.83	74.67	61.33
Ma *et al.* [34]	-	-	-	-	-	-	57.08	71.25	75.42	77.50	76.67	67.85
Weng *et al.* [53]	35.56	54.63	67.08	72.91	75.53	57.51	-	-	-	-	-	-
Aliakbarian *et al.*[40]	27.41	59.26	72.43	78.10	79.09	59.98	56.11	71.01	78.39	80.31	78.50	69.12
Hu *et al.* [16]	-	-	-	-	-	-	56.67	75.42	80.42	82.50	79.58	71.25
Wang *et al.* [52]	35.85	58.45	73.86	80.06	82.01	60.97	63.33	75.00	81.67	86.25	87.92	74.31
Pang *et al.* [38]	33.30	56.94	74.50	80.51	81.54	61.07	-	-	-	-	-	-
Tran *et al.* [49]	24.60	57.70	76.90	85.70	88.10	62.80	-	-	-	-	-	-
Ke *et al.* [22]	32.12	63.82	77.02	82.45	83.19	64.22	58.81	74.21	82.18	84.42	83.14	72.55
HARD-Net [29]	42.39	72.24	82.99	86.75	87.54	70.56	-	-	-	-	-	-
Baseline	38.09	66.36	78.67	83.29	84.10	66.43	60.71	73.04	77.81	83.88	84.32	72.20
ERA-Net w/o ELRO	43.94	73.23	84.53	87.61	87.97	71.62	63.50	80.82	82.70	86.33	87.10	75.78
ERA-Net	**53.98**	**74.34**	**85.03**	**88.35**	**88.45**	**73.87**	**65.30**	**81.27**	**85.67**	**89.17**	**89.38**	**77.73**

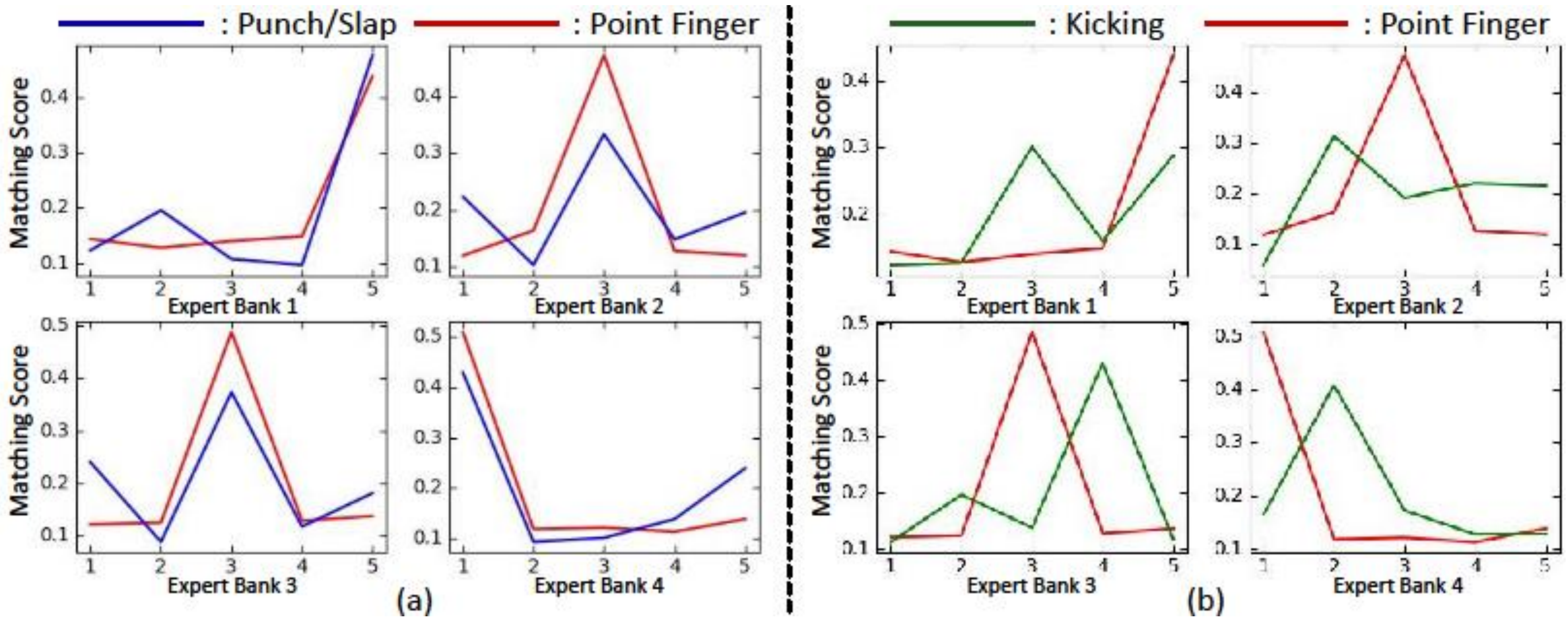
Experiments

Table 2: Results on NTU120 and UCF101

Methods	Observation Ratios on NTU120						Observation Ratios on UCF101					
	20%	40%	60%	80%	100%	AUC	10%	30%	50%	70%	90%	AUC
MSRNN [16]	-	-	-	-	-	-	68.01	88.71	89.25	89.92	90.23	80.89
Wu *et al.* [56]	-	-	-	-	-	-	80.24	84.55	86.28	87.53	88.24	80.57
Wu *et al.* [57]	-	-	-	-	-	-	82.36	88.97	91.32	92.41	93.02	84.66
Wang *et al.* [54]	-	-	-	-	-	-	83.32	88.92	90.85	91.28	91.31	89.64
Baseline	23.14	32.49	59.07	75.61	81.18	50.03	82.88	89.02	89.64	91.12	91.96	89.30
ERA-Net w/o ELRO	29.60	43.45	65.14	78.03	82.01	55.52	86.99	91.49	93.63	94.24	94.40	92.51
ERA-Net	**31.73**	**45.67**	**67.08**	**78.84**	**82.43**	**57.02**	**89.14**	**92.39**	**94.29**	**95.45**	**95.72**	**93.64**

Qualitative Validation

Supplementary Figure 1: Visualization of experts selection.



Thank You!