Supplementary Material: Generating Personalized Summaries of Day Long Egocentric Videos

Pravin Nagar, Anuj Rathore, C.V. Jawahar, Chetan Arora

Algorithm 1 Policy Gradient Framework

```
Initialize θ and learning rate α.
for For each sliding window do
    Calculate S_p and S_f according to the position of W_s
    Get M probability scores from the neural network
    for For each episode do
        Sample M actions from probability scores
        Compute cost and reward
```

$$cost{+}= \sum_{m=1}^{M} R(S) \nabla_\theta log \pi_\theta(a_m|h_m)]$$

```
    end for
    Compute episodic cost and episodic reward
    if episodic cost improves then
        update summary by picking top |S| sub-shots
    end if
    if For each mini batch then
        Back-propagate pseudo batch cost
    end if
end for
```

Algorithm 2 Q Learning Framework

```
Initialize θ, γ and learning rate α.
for For each sliding window do
    Calculate S_p and S_f according to the position of W_s
    Get M Q values from the Q value network
    Get M Q values from the target Q value network
    for For each episode do
        Sample M actions from probability scores
        Compute correction (TD error) for actions
```

$$\delta_{1:M} = R(S) + \gamma \sum_{m=1}^{M-1} Q_{\theta^-}(s_{m+1}, a_{m+1}) - \sum_{m=1}^{M-1} Q_\theta(s_m, a_m)$$

```
        Compute cost and reward R(S)
```

$$cost{+}= \delta_{1:M} \sum_{m=1, a\in A}^{M} \nabla_\theta Q_\theta(s_m, a_m)$$

```
    end for
    Compute episodic cost and episodic reward
    if episodic reward improves then
        update summary by picking top |S| subshots
    end if
    if For each mini batch then
        Back-propagate pseudo batch cost
    end if
end for
```

Though we have discussed the proposed approaches in section 3.2 of the main text but the Algorithm 1, Algorithm 2, and Algorithm 3 precisely describes the training process of Policy Gradient, Q Learning, and AC framework.

References

[1]

Subjects	Video Name	Dataset	Events Included	Events Excluded	Score (1 to 5)	User Experience
S01-S1	Alin	Disney	-			
S01-S1	P01	UTE				
S01-S1	Yair	HUJI				
S02-S1	Alin	Disney	-			
S02-S1	P01	UTE				
S02-S1	Yair	HUJI				
S03-S1	Alin	Disney	-			
S03-S1	P01	UTE				
S03-S1	Yair	HUJI				
S04-S1	Alin	Disney	-			
S04-S1	P01	UTE				
S04-S1	Yair	HUJI				
S05-S1	Alin	Disney	-			
S05-S1	P01	UTE				
S05-S1	Yair	HUJI				
S06-S1	Alin	Disney	-			
S06-S1	P01	UTE				
S06-S1	Yair	HUJI				
S07-S1	Alin	Disney	-			
S07-S1	P01	UTE				
S07-S1	Yair	HUJI				
S08-S1	Alin	Disney	-			
S08-S1	P01	UTE				
S08-S1	Yair	HUJI				
S09-S1	Alin	Disney	-			
S09-S1	P01	UTE				
S09-S1	Yair	HUJI				
S10-S1	Alin	Disney	-			
S10-S1	P01	UTE				
S10-S1	Yair	HUJI				

TABLE 1

The table shows the Likert score when specific events are included or excluded in the summary. S01-S1 represents subject 1 in scenario 1. The detail results of all 10 participants are shown in the supplementary material.

Algorithm 3 Actor Critic Framework

Initialize θ, w, γ and learning rates α_a, α_c.
for For each sliding window **do**
  Calculate S_{p} and S_{f} according to the position of W_{s}
  Get Q values from the Critic Network
  Get Policy distribution from Actor network
  Get Q values from the target Critic network
  for For each episode **do**
    Sample M actions from Policy distribution
    Actor cost calculation

$$cost_{ac} + = \sum_{m=1}^{M} Q_c(s_m, a_m) \nabla_\theta log(\pi_a(s_m, a_m))$$

    Compute correction (TD error) for actions

$$\delta_{1:M} = R(S) + \gamma \sum_{m=1}^{M-1} Q_{w^-}(s_{m+1}, a_{m+1}) - \sum_{m=1}^{M-1} Q_w(s_m, a_m)$$

    Compute cost and reward R(S)

$$cost_{cri} + = \delta_{1:M} \sum_{\substack{m=1, \\ a \in A}}^{M} \nabla_w Q_w(s_m, a_m)$$

  end for
  Compute episodic cost$_{ac}$, cost$_{cri}$ and episodic reward of actor and critic
  if episodic reward improves **then**
    update summary by picking top $|S|$ subshots
  end if
  if For each mini batch **then**
    Back-propagate pseudo batch cost$_{ac}$, and cost$_{cri}$
  end if
end for