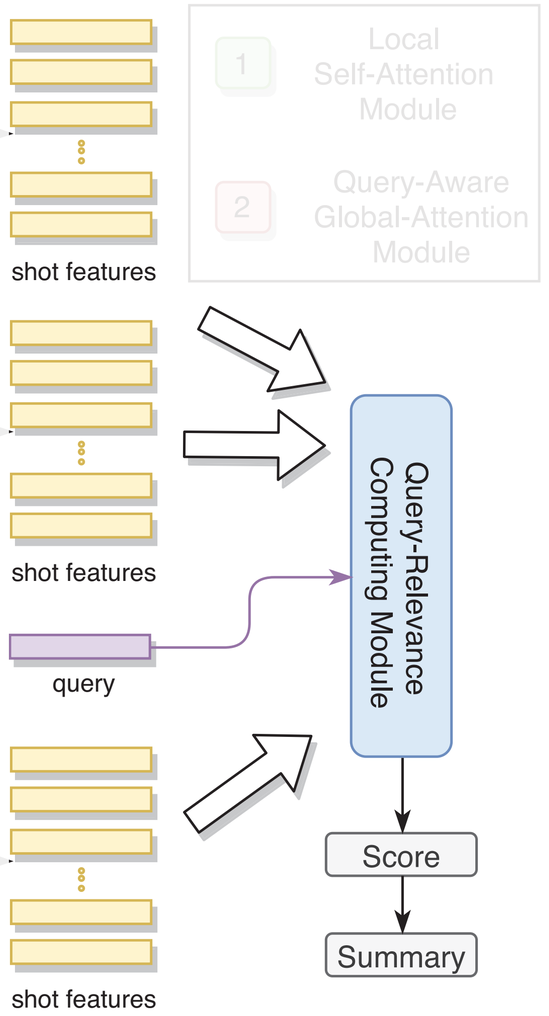
Proposed Method

Query-Relevance Computing Module

- Then let the output pass a MLP and get the concept-relevant score between i-th segment and concept c.
- The average of two concept-relevant score is taken as the query-relevant score $s = \{s_1, s_2, \ldots, s_n\}$
- Given the ground truth annotations $\hat{s} = \{\hat{s}_1, \hat{s}_2, \ldots, \hat{s}_n\} \in [0,1]$
- Loss: $L_{summ} = \frac{1}{T}\sum_{t=1}^{T} \hat{s}_t \log s_t + (1-\hat{s}_t)\log(1-s_t)$
- By minimizing the loss, module can focus on the most concept-related video shots.



Experiments

Datasets

- Sharghi, A.; Laurel, J. S.; and Gong, B. (CVPR 2017): *Query-Focused Video Summarization: Dataset, Evaluation, and A Memory Network Based Approach*
- Contains videos taken from the first person perspective.
- The dataset has 4 videos containing different daily life scenarios, each of which lasts 3~5 hours.
- Provide a set of concepts for user's queries (total 48), the concepts are concise and diverse, related to some common objects in our daily life.
- Each query is composed of two concepts and there are 46 queries in the dataset.