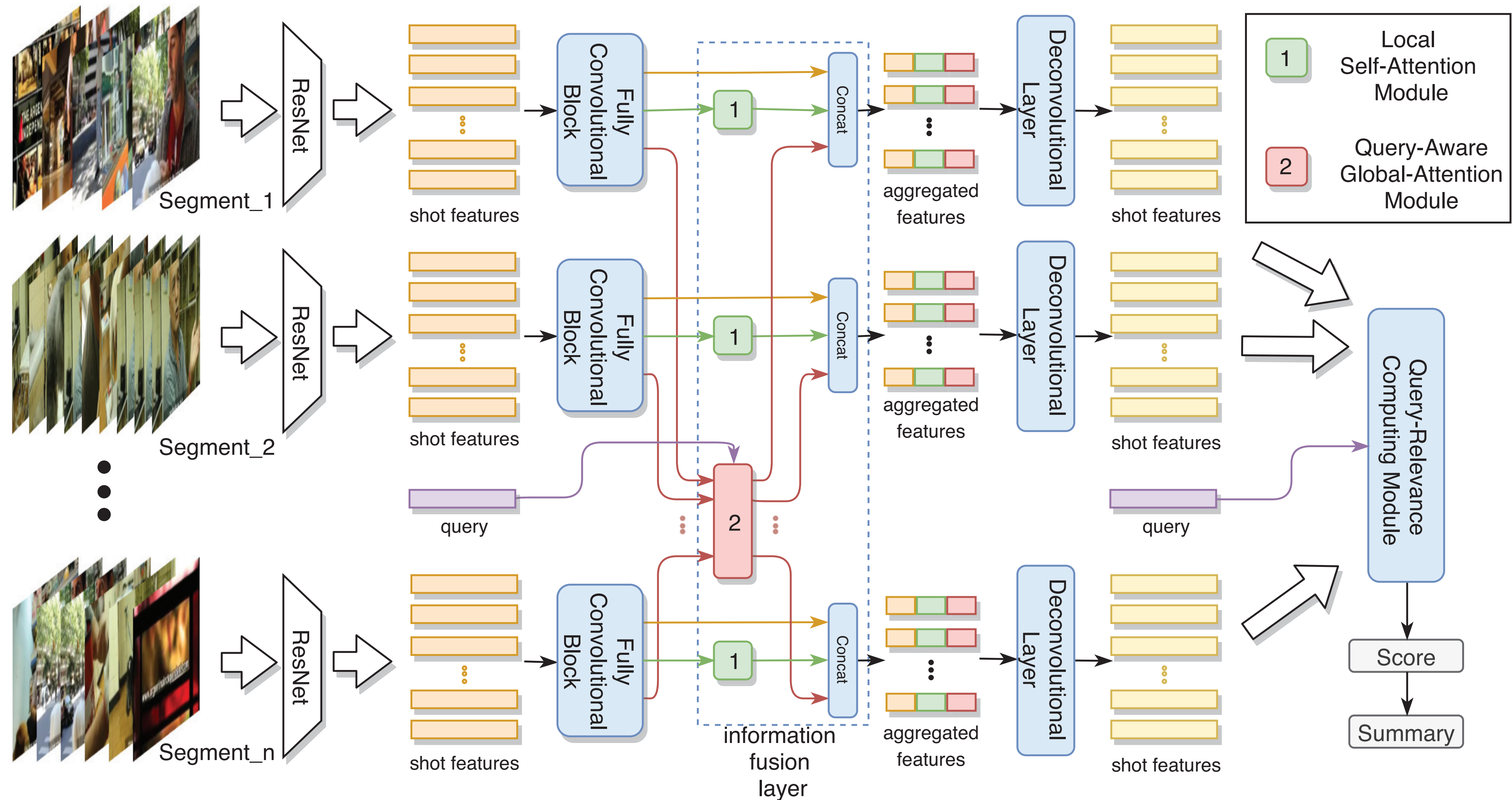


Proposed Method

Convolutional Hierarchical Attention Network (CHAN)



Proposed Method

Problem Formalization

- Given a long video v and a query q , find out the diverse and representative subset of query-related frames or video shots.
- Denote the task as a problem of calculating the shot-query similarity.
- Denote the video as a sequence of video shots (s_1, s_2, \dots, s_n)
 - n is the number of video shots
 - Each video shot is a small clip of the original video
- Denote h_q as representation of query