

```
graph TD; A[Event Detection from Video] --> B[Video Representation]; B --> C[Sub-sequence Representation]; C --> D[SB: An effective method to represent video for event detection [1].]; C --> E[SM: An efficient method to aggregate local features into video feature representation [2].]; C --> F[EDMIL: A method to leverage the event description to learn key evidences for event detection [3].];
```

Event Detection
from Video

Video
Representation

Sub-sequence
Representation

SB: An *effective*
method to represent
video for event
detection [1].

SM: An *efficient*
method to aggregate
local features into
video feature
representation [2].

EDMIL: A method to
leverage the *event
description* to learn
key evidences for
event detection [3].