

Input Image or
Video



Feature
Extraction

Input

LSTM
Generator

**Final
Caption**