

Dataset

visual



textual

$[w_1, w_2, \dots, w_L]$

ActivityNet-Captions

Charades-STA

TACoS

Model

Video Encoder

Text Encoder

Multi-Modal
Fusion

TGN

CMIN

LGI

Utils

Checkpoint,
Loss,
Optimizer,
Metric, Log,
Configure,
Timer

Dataloader.py

BuildModel.py

Runner

Train

Evaluate

Recall@1, IoU

Recall@5, IoU

Recall@10, IoU

mIoU