Vehicle Interaction Learning

Xiaosong Jia

07/22/2019

About Window Based Model

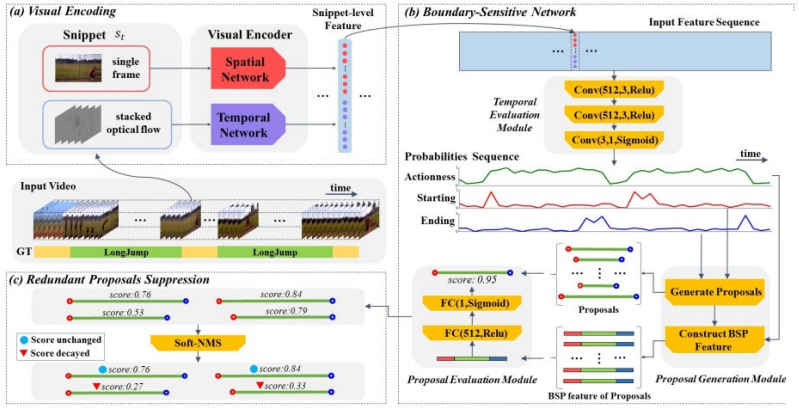
- Advantage:

1. Memory-saving

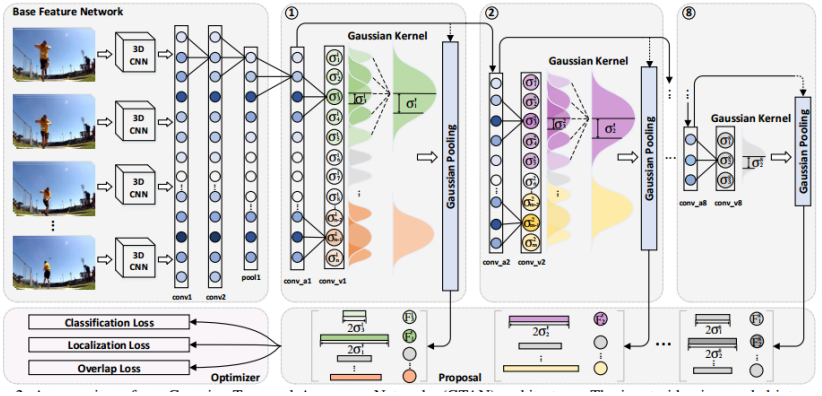
2. Multiple interactions in one pairs of interaction

- Disadvantage:

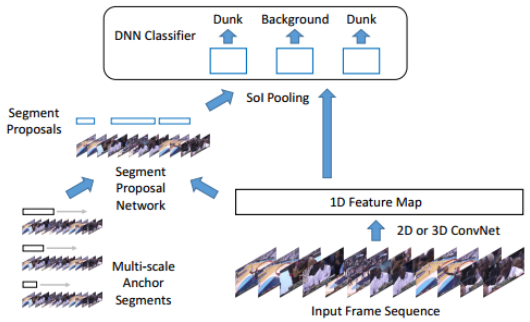
1. Lack of global information

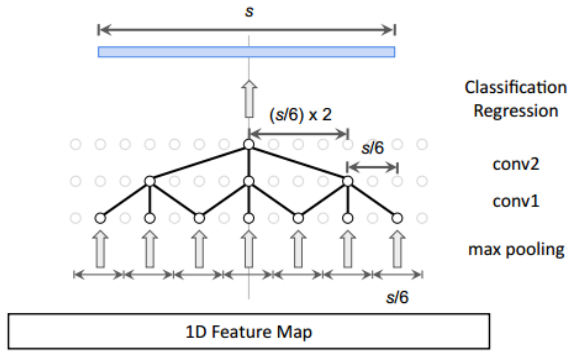


Lin et. al. ECCV 2018



Long et. al. CVPR 2018





Lin et. al. CVPR 2018 (state-of-the-art)

HighD Further Tuning (10304 samples)

	IOU > 0.9	Change Lane Acc	Interaction Cls Acc
Original model	90.0%	95.8%	98.4%
Only (x, y)	90.8%	93.5%	97.4%
Only (x, y)+ Pure transformer	95.2%	94.8%	97.5%
Only (x, y)+ Pure transformer+ 2 task	94.2%	94.4%	97.6%
…+Double Depth	94.5%	94.4%	93.6%

NGSIM (7916 samples)

	NGSIM	HighD
X1_mean	-2.40	-5.93
Y2_mean	786.23	0.095
X2_mean	-3.74	-7.04
Y2_mean	750.97	0.0090
X1_std	12.99	199.14
Y1_std	502.76	4.61
X2_std	12.78	178.44
Y2_std	494.22	4.68
SampleLen_mean	606.22	298.08
SampleLen_std	96.97	268.78
LabelLen_mean	30.46	69.22
LabelLen_std	2.60	13.16
SampleLen_max	2097	1958

- Use mean and std of NGSIM itself to normalize each frame and use model trained by HighD -> poor

- Clip? (max=2500, batch_size=2, sample_num=32972 -> 80min 1 epoch) 30~40 epoch to converge

- Frequency

- Train on NGSIM(28/40):

Change Lane Acc: 56.6%
Interaction Cls Acc: 92.0%

- Train on NGSIM and HighD (22/40):

Change Lane Acc: 70.0%
Interaction Cls Acc: 95.0%

Data Processing

- ref_x = (x1[0]+x2[0])/2 then x1 -= ref_x, x2 -= ref_x (√)

- Data Augmentation:

1. (x1,y1, x2, y2) and (x2, y2, x1, y1) have the same label (√)
2. Exchanging x and y still has the same label (√)
3. Set Normalization v2 (×): x_mean=(x1.mean + x2.mean)/, x_std=(x1.std + x2.std)/2, then standardize x1, x2respectively

   feature(x1, x2, y1, y2, x_mean, x_std, y_mean, y_std) -> some std = 0

Future Work

- Train on NGSIM
- Data Processing for different scale
- Data Augmentation: rotate coordinate system randomly
- Data from similar datasets/ Data from dissimilar datasets/Domain Adaption/Domain Generalization
- Multi-task weights?