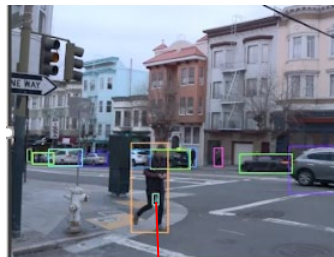
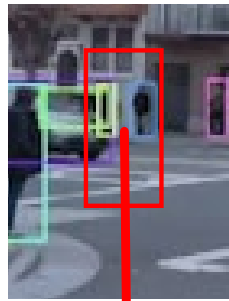


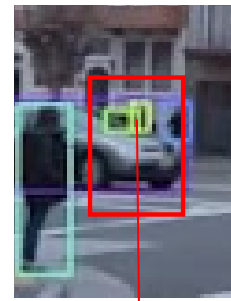
seen and not moving



Redundant box



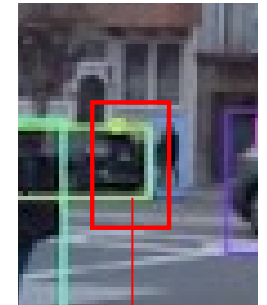
Before Occlusion



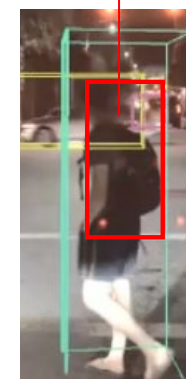
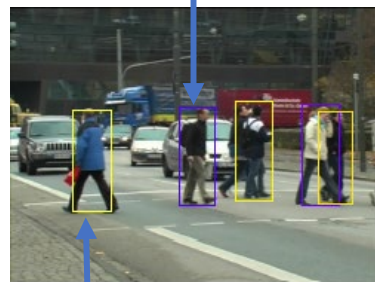
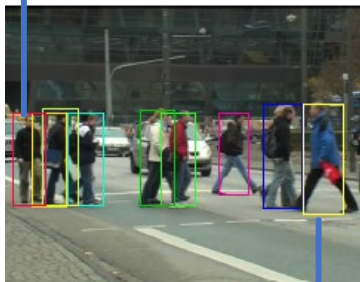
Partial Occlusion



Full Occlusion

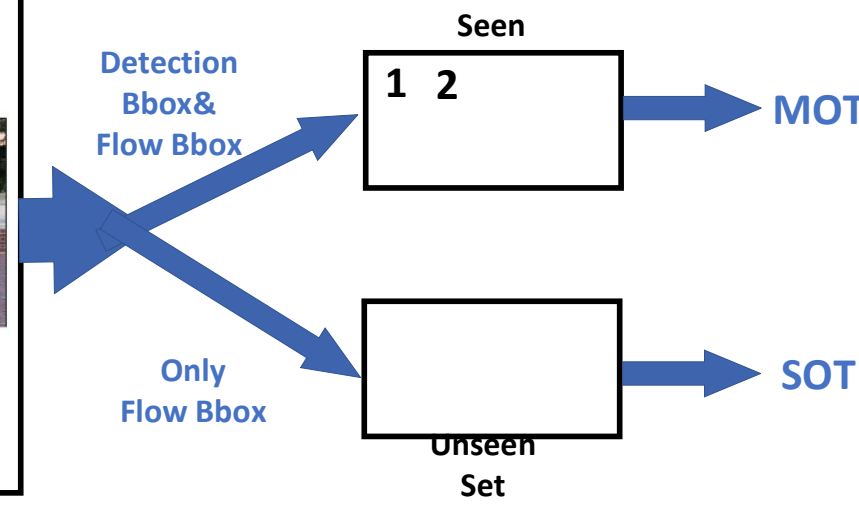
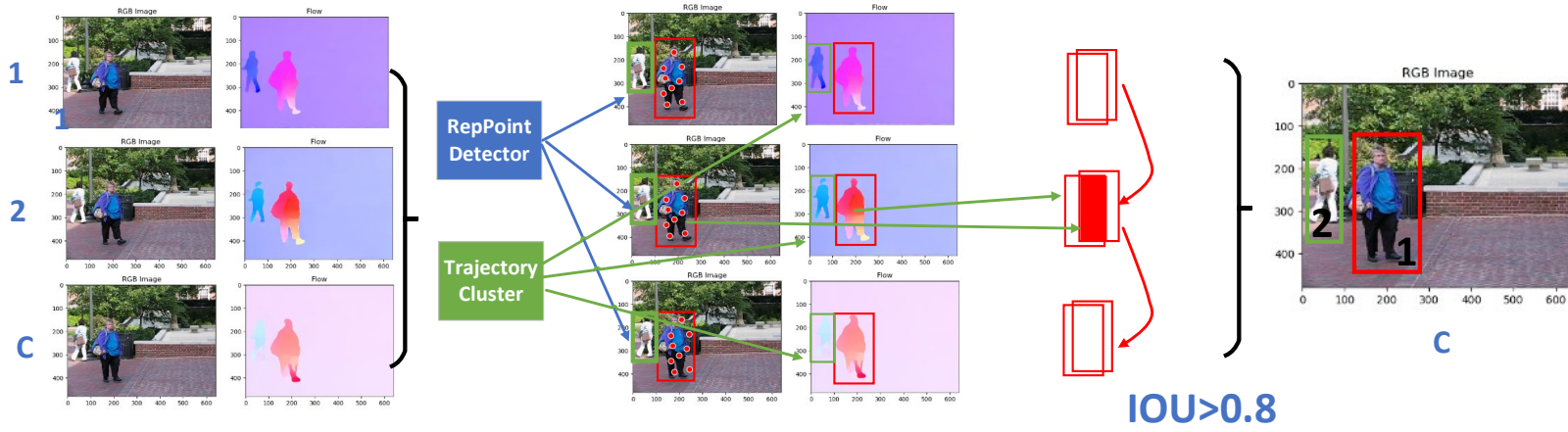


After Occlusion

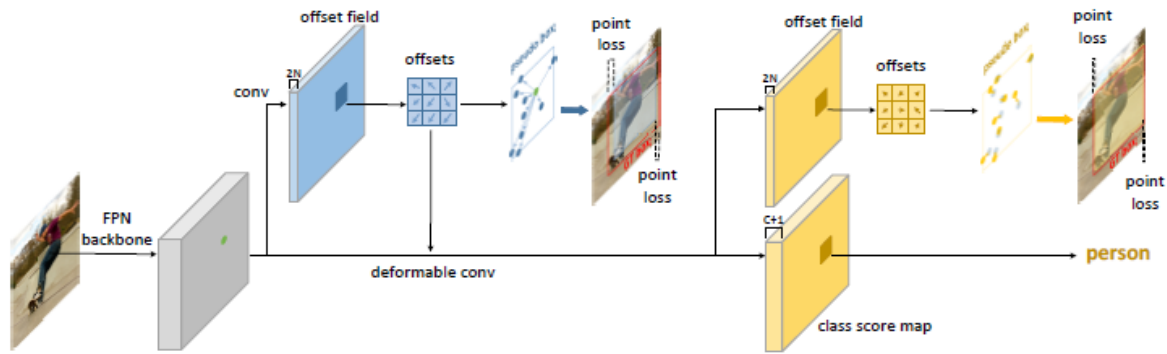


Association with similar Appearance

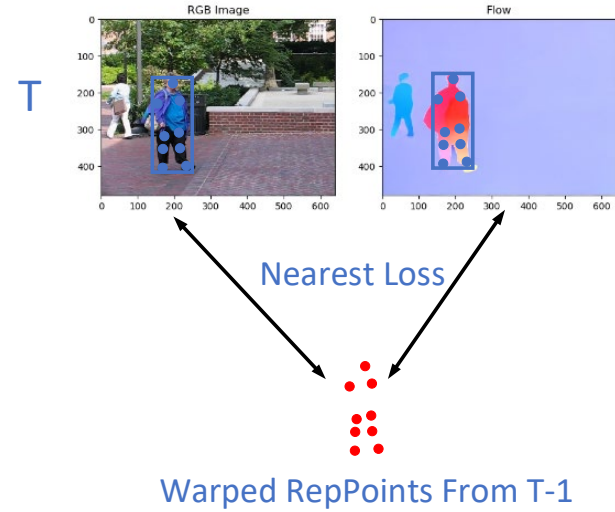
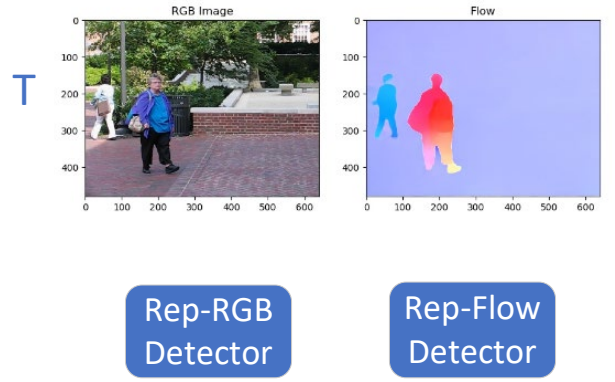
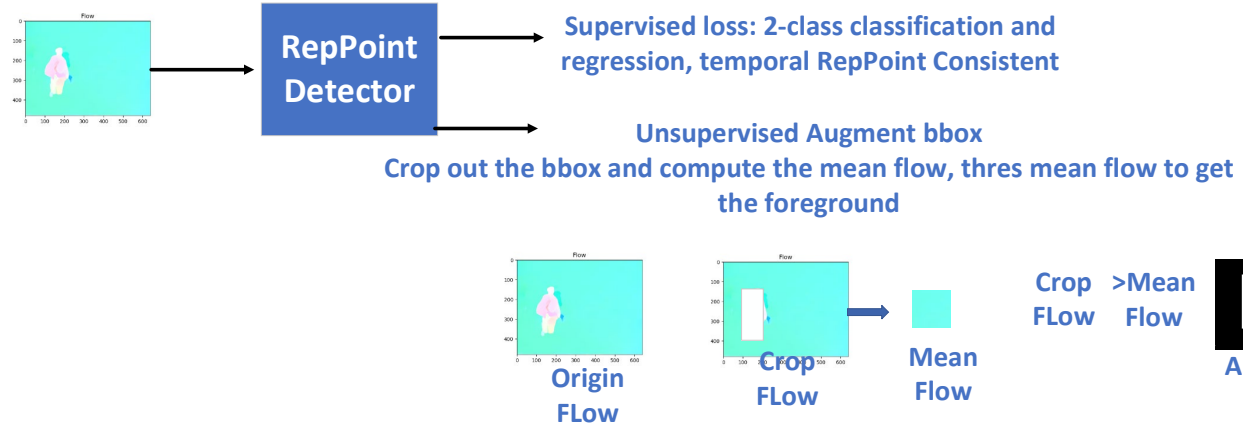
Init



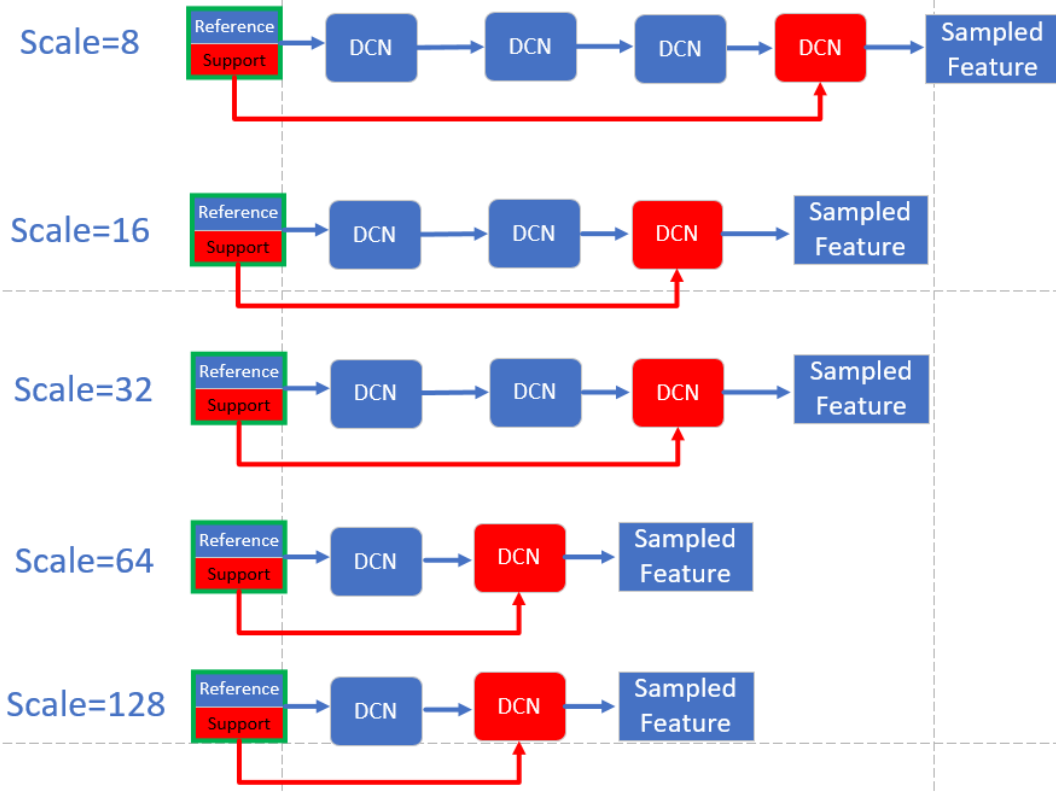
RepPoint Detector



Flow RepPoint



Video object detection by temporal sampling

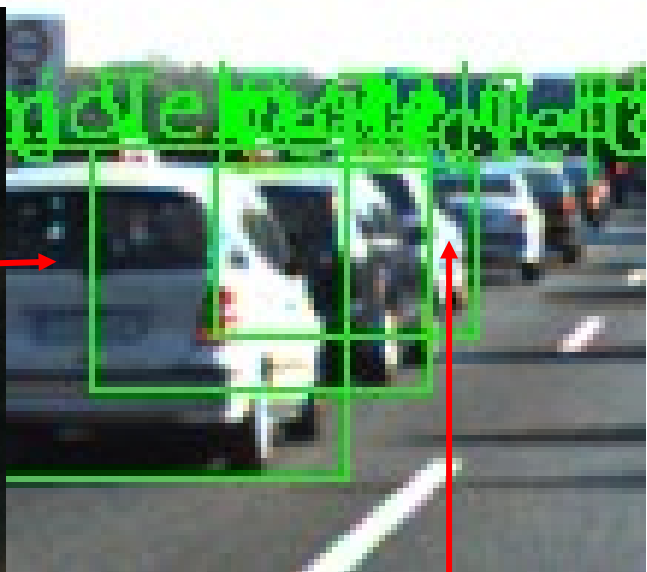
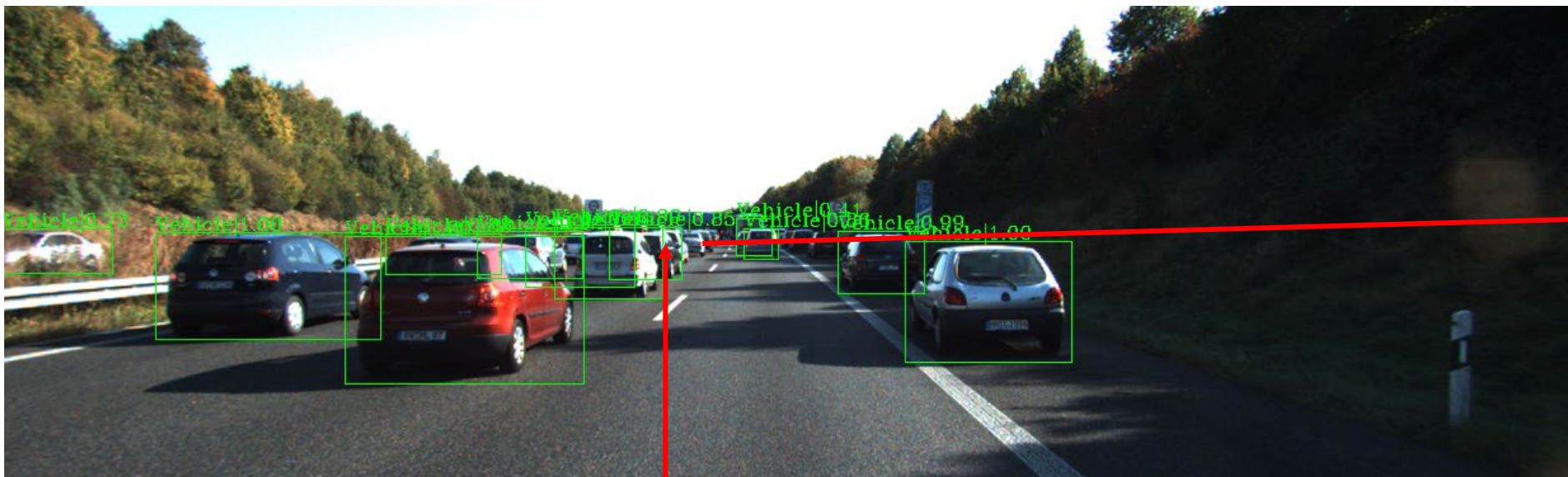


Without video

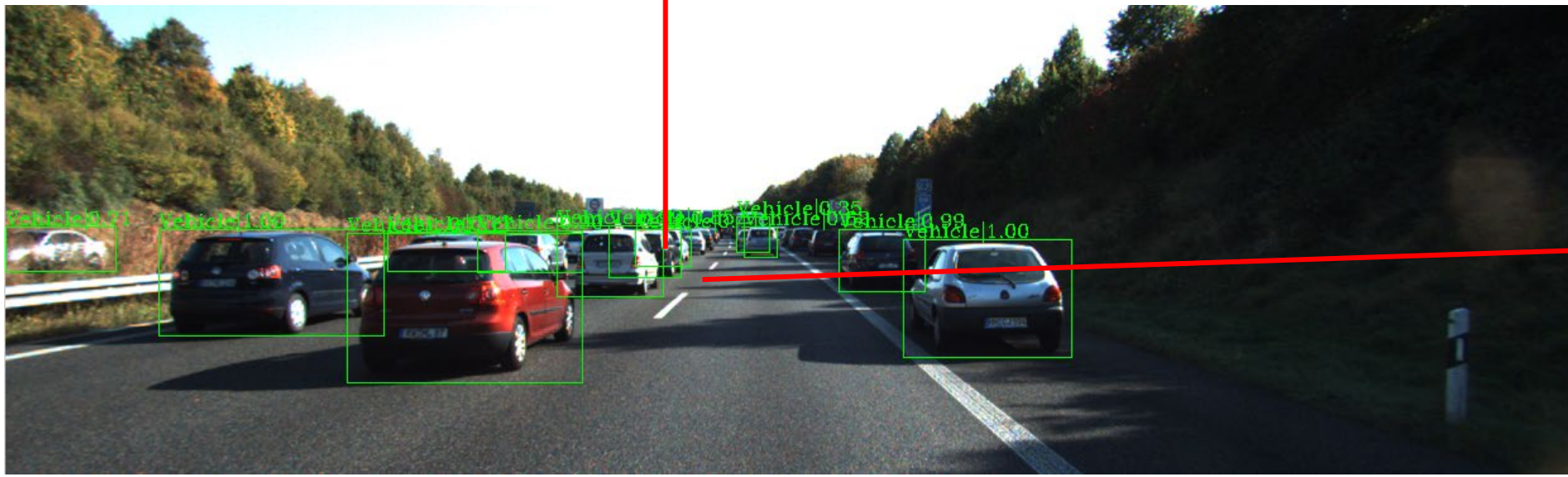


With video

Multi-scale spatial temporal sampling

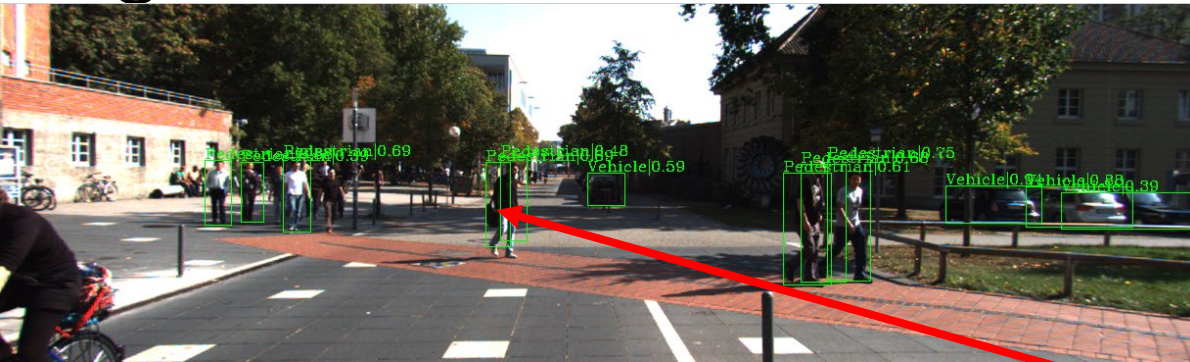


With video



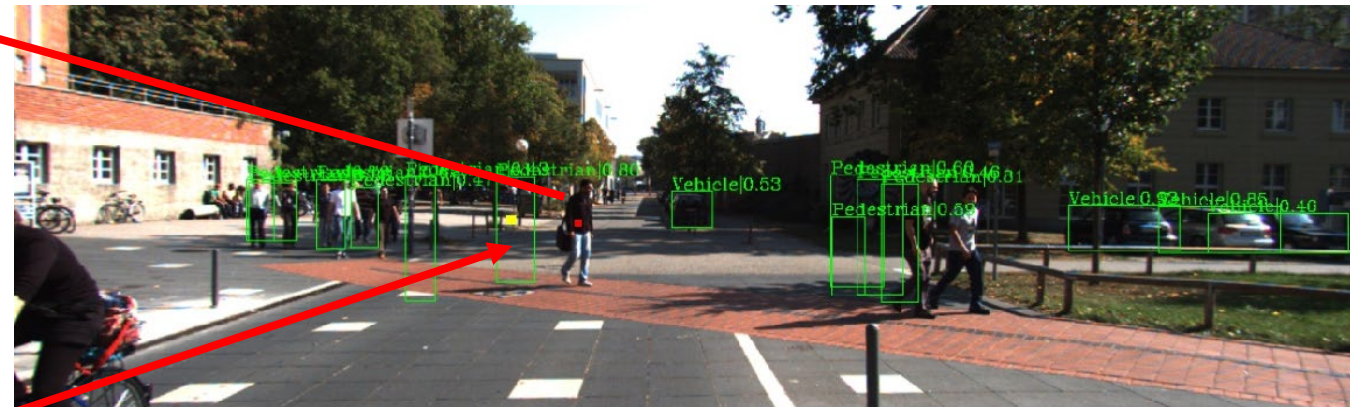
Without video

Using predicted offset to connect the bbox and stabilize regression

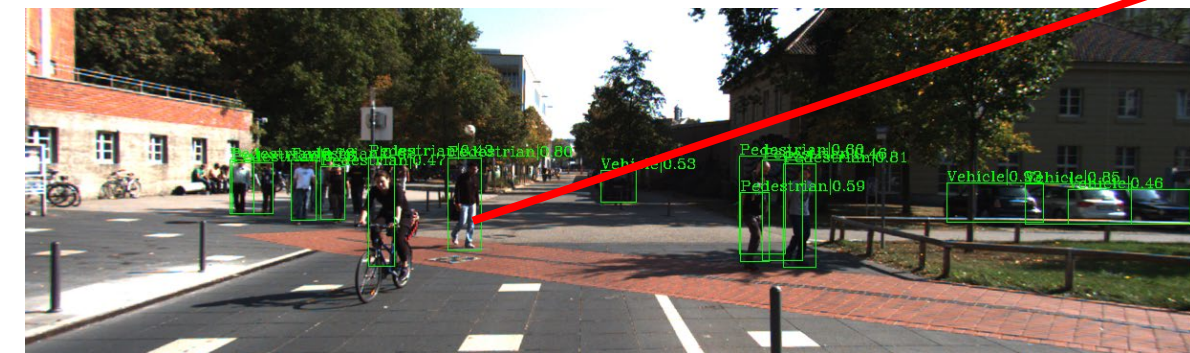


T=0

Use offset to aggregate feature and align bbox



Offset on T=0



T=10

Quantity result

```
writing results to /home/ld/RepPoints/out/reppoint_class3/test.pkl
```

class	gts	dets	recall	precision	ap
Vehicle	10747	21313	0.830	0.419	0.761
Pedestrian	9429	19405	0.732	0.356	0.645
Cyclist	509	3429	0.690	0.102	0.214
mAP					0.540

Without video

class	gts	dets	recall	precision	ap
Vehicle	10747	10484	0.734	0.752	0.703
Pedestrian	9429	6768	0.587	0.817	0.556
Cyclist	509	278	0.220	0.403	0.114
mAP					0.458

With video

Detection result:

Using temporal sampling bring the video information. It gives higher accuracy and less detection activation. But it also filtered too much redundant bbox, where we can see using video information only give $\frac{1}{2}$ detection results and reduce the recall.

MOT Evaluation Matrices

IDs	num_switches	Total number of track switches.
FP	num_false_positives	Total number of false positives (false-alarms).
FN	num_misses	Total number of misses.
GT	num_unique_objects	Total number of unique object ids encountered.
MT	mostly_tracked	Number of objects tracked for at least 80 percent of lifespan.
PT	partially_tracked	Number of objects tracked between 20 and 80 percent of lifespan.
ML	mostly_lost	Number of objects tracked less than 20 percent of lifespan.
FM	num_fragmentations	Total number of switches from tracked to not tracked.
MOTP	motp	Multiple object tracker precision.
MOTA	mota	Multiple object tracker accuracy.
Prcn	precision	Number of detected objects over sum of detected and false positives.
Rcll	recall	Number of detections over number of objects.
IDP	idp	ID measures: global min-cost precision.
IDR	idr	ID measures: global min-cost recall.
ID F1	idf1	ID measures: global min-cost F1 score.

1. Detection + Single frame Flow -lidong,zhiyin

KITTI data occlusion:(of all frames): IOU>0: 57.15%, IOU>0.5:18.78%, IOU>0.5/IOU>0: 32.87%

Ground Detection with IOU align

	IDF1	IDP	IDR	RcII	Prcn	GT	MT	PT	ML	FP	FN	IDs	FM	MOTA	MOTP
0016	74.80%	74.80%	74.80%	100.00%	100.00%	28	28	0	0	0	0	163	0	94.80%	0.016
0017	83.90%	83.90%	83.90%	100.00%	100.00%	11	11	0	0	0	0	28	0	96.80%	0.004
0018	96.00%	96.00%	96.00%	100.00%	100.00%	21	21	0	0	0	0	38	0	97.30%	0
0019	79.50%	79.50%	79.50%	100.00%	100.00%	106	106	0	0	0	0	309	0	96.50%	0.007
0020	92.00%	92.00%	92.00%	100.00%	100.00%	134	134	0	0	0	0	98	0	98.60%	0
Overall	85.20%	85.20%	85.20%	100.00%	100.00%	60	60	0	0	0	0	127	0	96.80%	0.005

Ground Detection with Flow align

	IDF1	IDP	IDR	RcII	Prcn	GT	MT	PT	ML	FP	FN	IDs	FM	MOTA	MOTP
0016	83.50%	83.50%	83.50%	100.00%	100.00%	28	28	0	0	0	0	73	0	97.70%	0.011
0017	73.30%	73.30%	73.30%	100.00%	100.00%	11	11	0	0	0	0	52	0	94.10%	0.004
0018	97.40%	97.40%	97.40%	100.00%	100.00%	21	21	0	0	0	0	22	0	98.40%	0
0019	84.90%	84.90%	84.90%	100.00%	100.00%	106	106	0	0	0	0	146	0	98.30%	0.001
0020	91.90%	91.90%	91.90%	100.00%	100.00%	134	134	0	0	0	0	336	0	95.10%	0
Overall	86.20%	86.20%	86.20%	100.00%	100.00%	60	60	0	0	0	0	125	0	96.70%	0.003

RepPoint Detection

Reppoint detection with IOU Align

Name	IDF1	IDP	IDR	Rc11	Prcn	GT	MT	PT	ML	FP	FN	IDs	FM	MOTA	MOTP
0016	34.1%	42.5%	28.4%	55.0%	82.2%	28	6	16	6	372	1407	146	123	38.4%	0.265
0017	48.1%	50.9%	45.6%	67.9%	75.7%	11	4	5	2	192	283	33	23	42.3%	0.225
0018	79.2%	77.0%	81.5%	92.8%	87.8%	21	19	2	0	182	101	62	20	75.5%	0.156
0019	39.3%	48.7%	33.0%	56.3%	83.1%	106	25	60	21	1000	3826	486	364	39.4%	0.258
0020	61.2%	66.4%	56.7%	73.9%	86.6%	134	53	72	9	784	1783	292	124	58.2%	0.156
Overall	52.4%	57.1%	49.0%	69.2%	83.1%	60	21	31	7	506	1480	203	130	50.8%	0.212

Reppoint detection with Flow Align

Name	IDF1	IDP	IDR	Rc11	Prcn	GT	MT	PT	ML	FP	FN	IDs	FM	MOTA	MOTP
0016	43.4%	54.2%	36.2%	55.4%	82.8%	28	6	16	6	360	1395	118	127	40.1%	0.266
0017	53.4%	56.5%	50.6%	68.1%	75.9%	11	4	5	2	190	281	32	24	42.9%	0.223
0018	79.8%	77.6%	82.0%	92.8%	87.8%	21	19	2	0	182	101	53	20	76.2%	0.156
0019	43.6%	54.0%	36.6%	56.4%	83.2%	106	25	60	21	997	3823	389	363	40.5%	0.258
0020	62.0%	67.3%	57.5%	74.0%	86.6%	134	52	73	9	783	1782	389	124	56.8%	0.156
Overall	56.4%	61.9%	52.6%	69.3%	83.3%	60	21	31	7	502	1476	196	131	51.3%	0.212

RetinaNet Detection

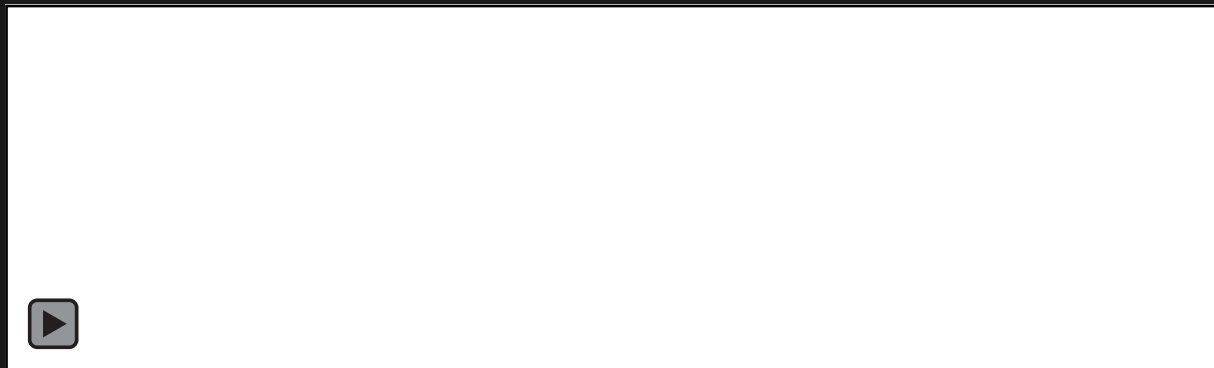
RetinaNet Detection with IOU Align

Name	IDF1	IDP	IDR	Rc11	Prcn	GT	MT	PT	ML	FP	FN	IDs	FM	MOTA	MOTP
0016	35.9%	43.6%	30.5%	58.8%	84.0%	28	9	14	5	349	1286	186	133	41.7%	0.263
0017	41.6%	39.3%	44.3%	69.5%	61.7%	11	4	5	2	380	269	58	26	19.8%	0.221
0018	82.8%	80.5%	85.3%	92.8%	87.6%	21	18	3	0	185	101	53	16	75.9%	0.163
0019	36.2%	40.9%	32.5%	60.5%	76.3%	106	32	59	15	1650	3458	641	371	34.4%	0.261
0020	55.1%	57.4%	53.0%	74.7%	80.8%	134	61	66	7	1211	1729	404	121	51.1%	0.164
Overall	50.3%	52.3%	49.1%	71.3%	78.1%	60	24	29	5	755	1368	268	133	44.6%	0.214

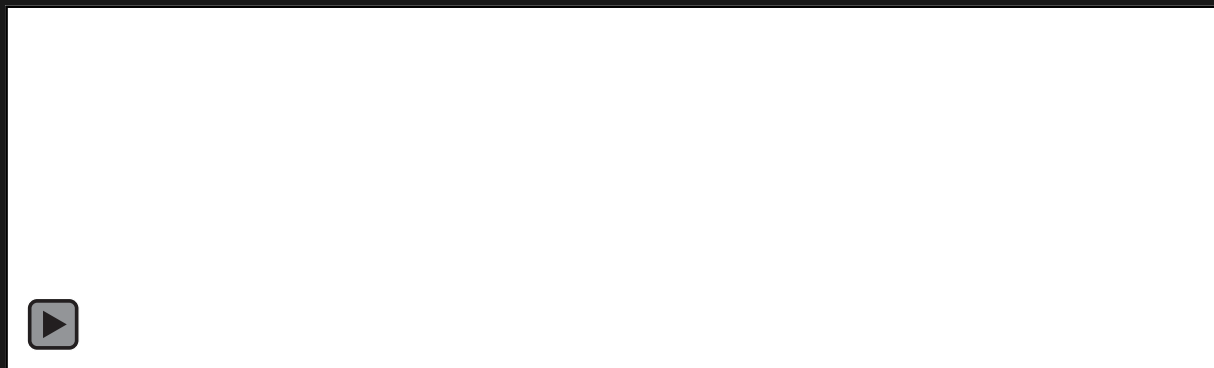
RetinaNet Detection with Flow Align

	IDF1	IDP	IDR	Rc11	Prcn	GT	MT	PT	ML	FP	FN	IDs	FM	MOTA	MOTP
0016	46.4%	56.3%	39.4%	58.9%	84.2%	28	8	15	5	346	1283	147	140	43.2%	0.266
0017	45.2%	42.6%	48.0%	69.7%	61.9%	11	5	4	2	378	267	44	27	21.8%	0.221
0018	82.8%	80.5%	85.3%	92.8%	87.6%	21	18	3	0	185	101	47	16	76.4%	0.163
0019	41.1%	46.4%	36.8%	60.5%	76.2%	106	31	61	14	1651	3459	496	363	36.0%	0.260
0020	58.0%	60.4%	55.8%	74.8%	80.9%	134	61	66	7	1206	1724	461	120	50.4%	0.163
Overall	54.7%	57.3%	53.1%	71.4%	78.2%	60	24	29	5	753	1366	239	133	45.5%	0.215

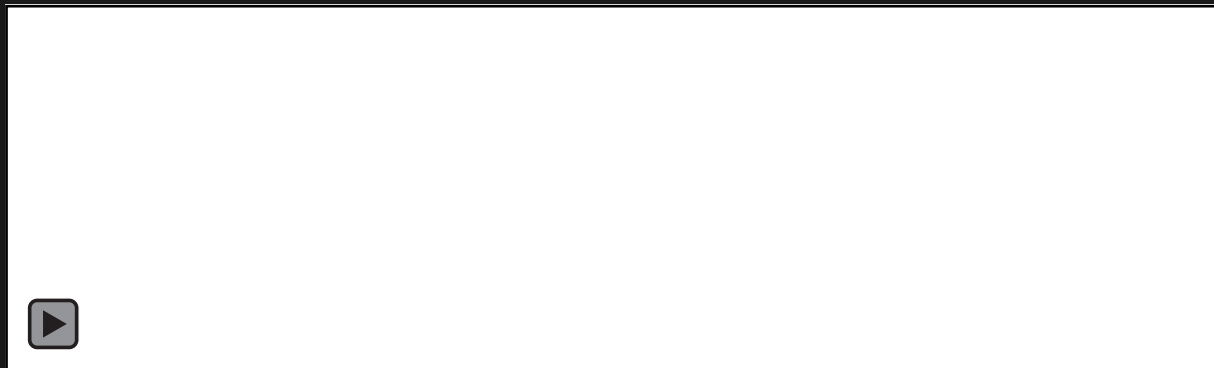
Align with Ground Truth



Ground Truth



Align with IOU

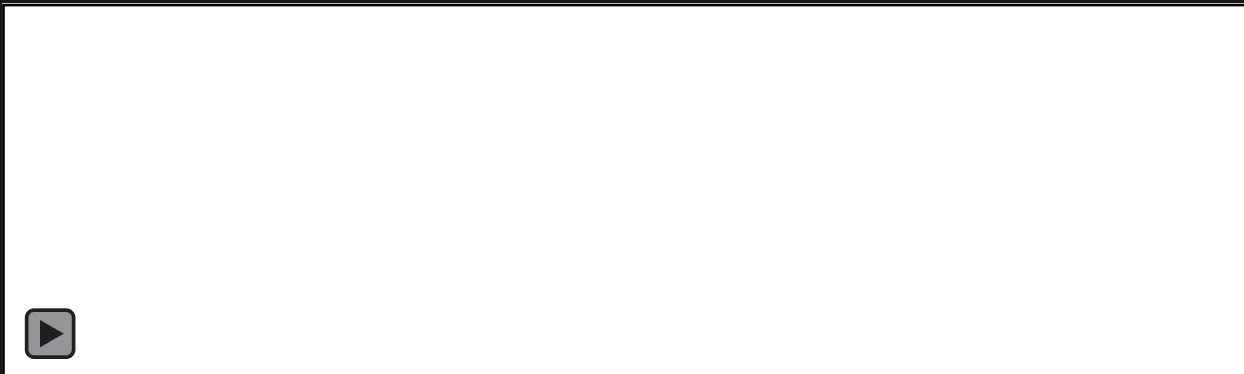


Align with Flow

Need Analyze

IOU Align with RepPoint and RetinaNet

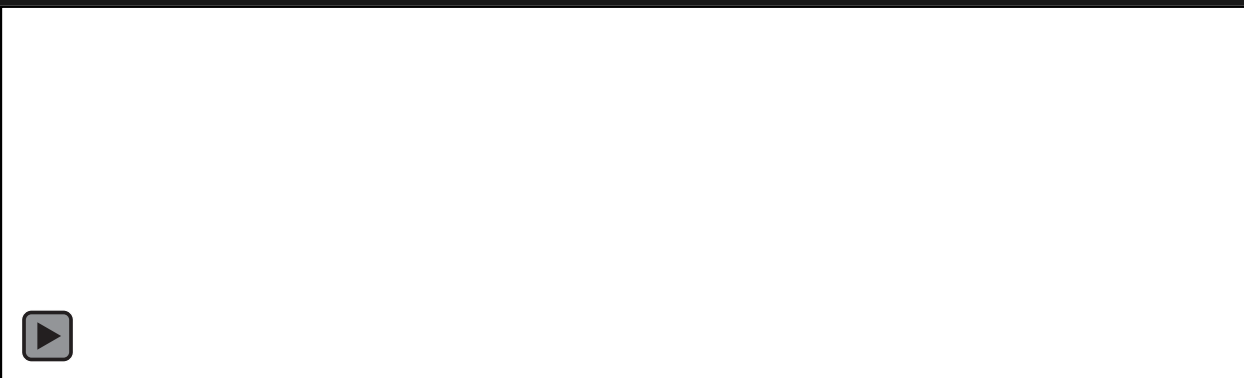
Need Analyze



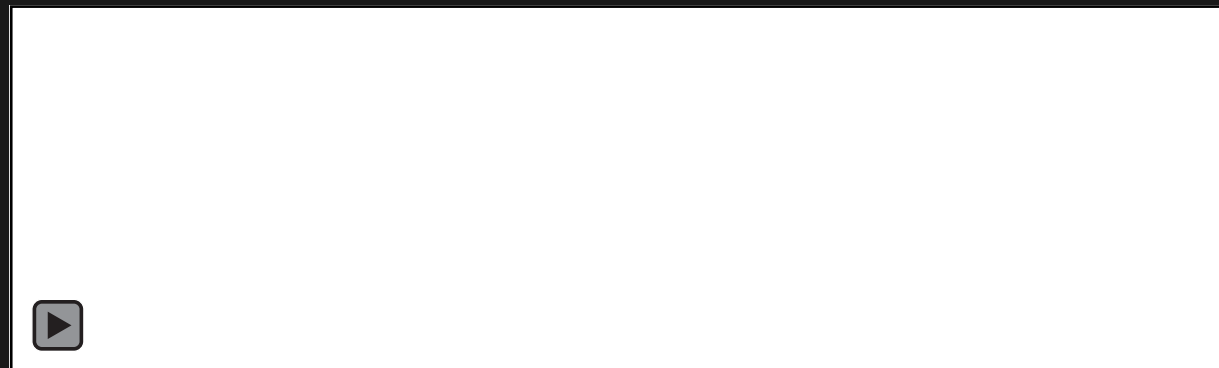
RepPoint



RetinaNet



RepPoint(highlight occlusion)



RetinaNet(highlight occlusion)