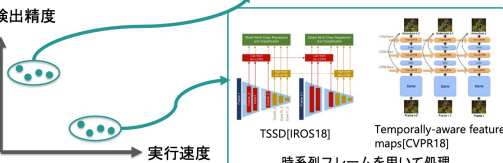
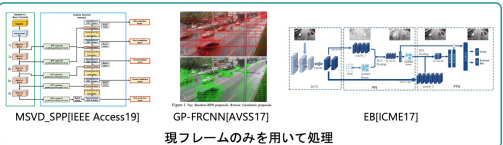
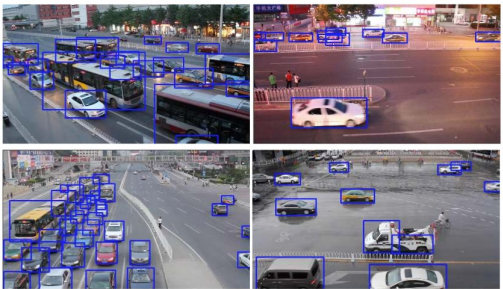


どんな研究？

道路状況の把握などのため、精度の高い車両検出技術が求められていますが、深層学習を持ちいた検出技術は精度が高い反面処理速度が遅いです。そこで精度と速度のトレードオフを改善するため、少量の過去フレームを活用し、効率よく重ね合わせることで、ブレなどに強い情報を作りリアルタイム処理可能な車両検出技術の研究を進めています。

車両検出とは？ 既存研究は？



車両検出は？

- 映像中の車両の領域を検出する技術
- 静止画に比べて、ボケや車両同士が重なってしまうため困難

既存研究の課題は？

- 高精度を追求するために、処理が遅くリアルタイム処理不可
- 処理が早い検出器は精度が低い

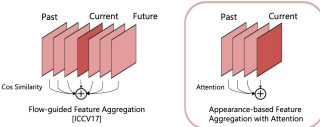
何がわかる？

- 映像中の物体検出とは？
- 何が優れている？
- 問題点は？
- 解決のアプローチは？

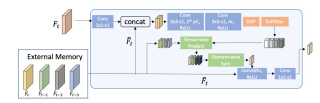
研究内容

解決のアプローチ

- 処理速度の早い特徴抽出器の軽い特徴マップを時間方向に集約し、検出精度を向上
- 先行研究の各フレームの類似度から結合の重みを決定から、アテンションにより一括算出



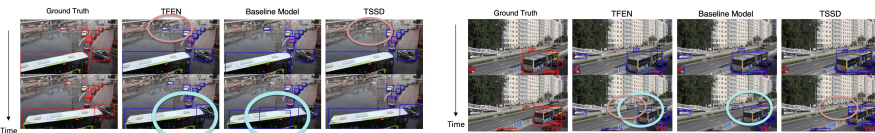
Temporal Attention Decoder & External Memory



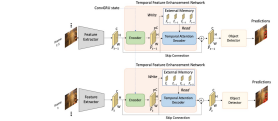
- 役割**
 - F_t と時系列特徴マップ F_{t-1} から現在・過去フレームに対する重み付けを算出
 - External Memory内の特徴マップと重み付けの積の重ね合わせにより特徴マップの生成

新規性 アテンションによる過去情報を一括集約した特徴量の生成

検出結果の比較



提案手法全体図



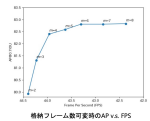
- 既存の静止画検出器にTFEN(Temporal Feature Enhancement Network)を導入
- 特徴量抽出器からの特徴マップを改善することで精度を向上

新規性 アテンションベースによる時系列特徴マップの結合

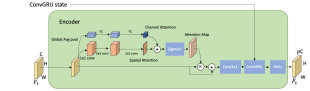
格納フレーム数が精度と速度に与える影響

検証 最大収容過去フレーム m を可変とし、評価時の観測 (訓練・評価にはFP16を使用、AP50Overallを指標)

- 結果**
 - 格納フレーム数を増加させることで精度が向上傾向
 - 精度が飽和し始めるのは $m=6$
 - 精度・速度のトレードオフ点は $m=4-6$ にあると推測



Spatiotemporal Encoder



- 役割**
 - ConvGRUにより時系列情報を組み込んだ特徴マップの生成
 - 空間 & チャネルアテンションにより物体領域を強調
 - 計算量削減のため、チャンネルを圧縮

新規性 時系列情報に空間 & チャネルアテンションの導入

関連手法との比較

Model	Channel	Step	Median	Clarity	Height	Memory	FPS	AP50
SPNetv2	2076	34x2	20.28	17.62	24.70	30.91	23.05	21.77
ASRCNN	4032	34x2	13.12	10.01	24.91	29.98	37.98	26.68
MEMNet	4032	38x1	16.26	16.47	18.73	29.28	29.06	27.32
ConvGRU	1512	16x1	18.71	18.16	14.53	16.17	18.17	23.52
Flow-GRU4Rec	564	16x1	12.05	14.25	12.24	16.29	14.16	18.85
GRU4Rec	564	16x1	10.05	10.05	10.11	10.11	10.11	10.11
TSST	674	16x1	12.12	12.44	12.42	12.42	12.42	11.78
MSVD_SPP	8128	30x2	16.42	16.59	18.81	16.11	17.96	16.11
YOLOv3-SPP	3048	30x2	16.88	17.24	18.42	18.11	17.46	18.42
FB-MNets	3048	30x2	15.01	15.78	15.26	15.42	15.02	16.16
TSST	1716	16x1	12.07	12.14	12.51	12.47	14.48	12.51
TSST	812	16x1	10.88	10.78	11.54	12.41	12.22	12.41

- 各研究(MSVD_SPP, YOLOv3-SPP)と比較し、同等の精度を保持しながら、実行速度は約倍
- 時系列情報を利用したTSSTは実行速度には追いつかないものの、精度面にてパフォーマンスが高い