![](_page_0_Picture_0.jpeg)

![](_page_0_Picture_1.jpeg)

![](_page_0_Picture_2.jpeg)

### VmAP: A Fair Metric for Video Object Detection Anupam Sobti, Vaibhav Mavi, M. Balakrishnan and Chetan Arora Indian Institute of Technology Delhi

### The Problem

• Video: Heavily correlated frames

![](_page_0_Picture_6.jpeg)

High correlation between an object being detected in nearby frames

# Claims

- VmAP is more sensitive to bias
- VmAP results in a better operating point (much lesser False Positives)
- VmAP provides better post-tracking performance
- AP treats all frames as Independently Identically Distributed (IID).

### Toy Example

![](_page_0_Figure_14.jpeg)

#### Bias Sensitivity : Synthetic Detectors

![](_page_0_Picture_17.jpeg)

#### **Operating Point : Maximise Set-level F1 Score**

	Precision-Recall curve for Person							
.0								
.8								
.6								
.4								
.2								

0.6

0.0

0.2

0.4

	Video (Set) Recall		False Positives	
	P-R OP	VP-VR OP	P-R OP	VP-VR OP
DFF	86.96	84.78	108	1
FGFA	89.13	82.61	105	1
RDN	91.3	86.96	258	3

93.48

109

2

- Three detectors (D1, D2, D3) with the same recall have different consequences
- D2 might lead to a vehicle colliding with Person A
- D3 would lead to collision in the later phase

![](_page_0_Picture_24.jpeg)

#### Correlation with post-tracker performance Thinking ahead

MEGA

95.65

- 9% increase in spearman correlation coefficient
- VR is a better measure than FR

![](_page_0_Figure_28.jpeg)

## Timeline plot: Understanding Video Object Detection in a Plot

![](_page_0_Figure_30.jpeg)

Sets with at least 1 Sets with no detections detection are marked as are marked as false a true positives (green) negatives (red)

![](_page_0_Figure_32.jpeg)