

# Presentation #2

...

# Data Collection

- Data from 28 cities (Indianapolis, Incheon, Indore, Islamabad, Jacksonville, Jaipur, Jeddah, Jerusalem, Johannesburg, Kabul, Kampala, Karachi, Kathmandu, Kawasaki, and 14 cities from the previous week)
- Ground Videos
  - 188 videos
- Aerial Videos
  - 234 videos
- Total Videos
  - Past 14 days
    - 422
  - Total
    - $390 + 422 = 812$  videos total

# Paper

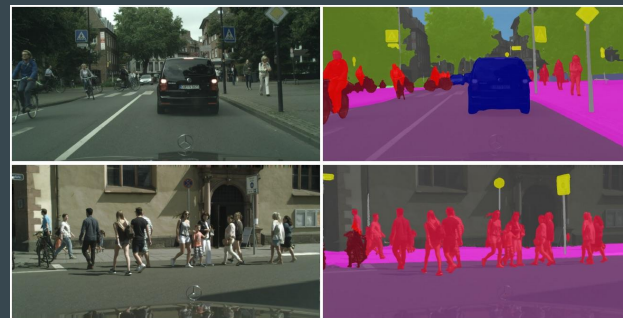
- BDD100K
  - Contains 100k videos with high resolution and diverse scenes, etc.
    - 10 tasks
    - 40 second videos
    - Image Tagging on 3 times of day, 6 scene types, 6 weather conditions
    - High Resolution (720p) images
    - High Frame Rate (30 fps)
    - GPS/IMU recordings to preserve driving trajectories
    - 70k training videos
    - 10k validation videos
    - 20k testing videos
      - Testing Dataset
        - Used to provide an unbiased evaluation of a final model fit on the training dataset

# Paper (Continued)

- Imitation Learning
  - GPS/IMU recordings to show human driver action given visual input/driver trajectories
    - Used for imitation learning algorithms and to measure similar driving behaviors in the validation/testing sets
- Object Detection
  - Bounding Box Annotations for Categories
- MOT Dataset
  - 2,000 videos 400k frames
  - Annotated at 5 frps = 200 frames per video
  - 3.3 million bounding boxes in dataset

# Paper (Continued)

- Dataset consists of driving videos under different
  - Weather conditions; time; scene types
  - Variety of Annotations
    - Scene Tagging
    - Object Bounding Box
    - Lane Marking
    - Driveable Area
    - Full Frame Semantic/Instance Segmentation
    - Multiple Object Tracking
    - Multiple Object Tracking with Segmentation
- Semantic Segmentation
  - Segmenting images into labels with a semantic meaning
    - Each pixel has a semantic meaning
    - Annotations are for pixels, not bounding boxes



# Paper (Continued)

