**Labeling Guidelines**

1. **Gồm 5 lớp:** motorcycle, car, bus, truck, container
2. **Ý nghĩa các lớp:**

* 0: motorcycle: xe máy, xe đạp, xe đạp điện, xe máy điện
* 1: car: xe hơi, xe taxi, xe bán tải
* 2: bus: xe buýt, xe khách
* 3: truck: xe tải, xe bồn, xe ben
* 4: container: xe container

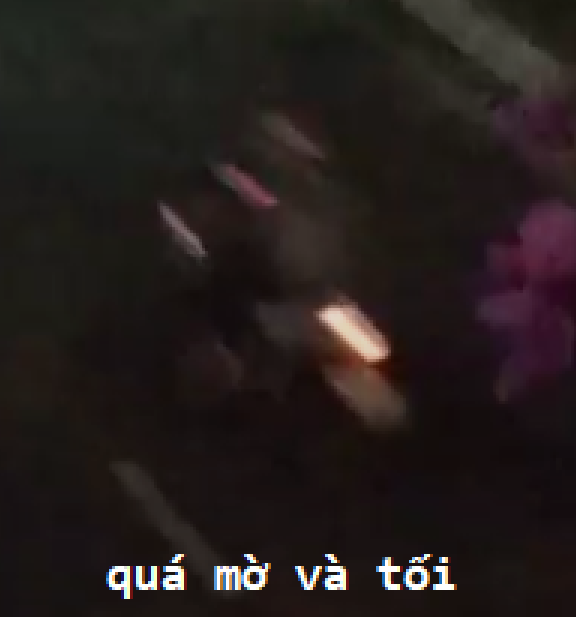
1. **Lưu ý khi gán nhãn:**
2. ***Đối với xe máy, chỉ gán nhãn xe không gán nhãn người***



1. ***Đối với tất cả xe, chỉ gán nhãn nếu xe đó lộ trên 70% nguyên hình***



1. ***Xe máy nếu bị che bởi một phần cơ thể con người hoặc áo mưa vẫn gán nhãn như thường***
2. ***Không gán nhãn đối với xe quá xa, hình ảnh quá mờ, quá tối, ví dụ:***



1. **Công cụ gán nhãn:**

1/ Dark Label

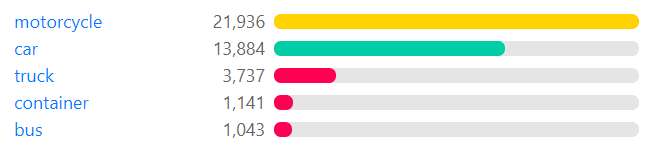
2/ CVAT (Computer Vision Annotation Tool)

1. **Model:**

* Chọn 1 model Object detection 5 năm trở lại đây
* Tải dataset là file true\_YOLO\_format.zip

Chi tiết dataset:

* 8000 hình (lấy từ 500 ảnh đầu của 16 video)
* Chỉ có 7241/8000 hình được gán nhãn
* 41741 xe được gán (trung bình là 5.2 xe mỗi hình)
* Phân bố 5 class:



Chi tiết các file trong dataset:

* Tên hình .jpg trùng với tên nhãn .txt
* Thứ tự các cột trong file txt: tên lớp

tỉ lệ vị trí center của bb / chiều rộng frame

tỉ lệ vị trí center của bb / chiều cao frame

tỉ lệ độ rộng bounding box / chiều rộng frame

tỉ lệ độ cao bounding box / chiều cao frame

* Tỷ lệ train 70% val 20% test 10% (chia dạng shuffle, nếu dùng train\_test\_split thì mặc định có rồi)
* Thống kê thời gian train model
* Độ đo đánh giá: IoU, mAP, AP, AR (đánh giá trên những ảnh tập test)
* Dùng GPU cả khi train và khi predict thì thời gian thống kê được mới khách quan
* Sau đó demo trên frame 500-1499 của bốn video sau: 3,5,10,22 (mỗi model đều dùng 4 đoạn này) (nhớ xuất video kèm bbox)

và thống kê thời gian predict 1 hình, tổng thời gian predict 4 đoạn video (quan trọng vì họ YOLO sẽ predict rất nhanh)