Object Counting on Low Quality Images
A Case Study of Near Real-Time Traffic Monitoring
Jean-François Rajotte\textsuperscript{1,} Martin Sotier\textsuperscript{1,} Cédric Nioseux\textsuperscript{1,} Louis-Philippe Noé\textsuperscript{1} and Thomas Bertrien\textsuperscript{2,3}
\textsuperscript{1} jean-francois.rajotte@crim.ca, \textsuperscript{2} martin.sotier@crim.ca, \textsuperscript{3}cedric.nioseux@crim.ca, \textsuperscript{4}louisphilippe.noee@gmail.com and \textsuperscript{5}thomas.bertrien@gmail.com

**Key Insights**
- Low-quality urban road traffic images from city cameras can be used as a non-intrusive traffic monitoring and historical exploration system.
- A pre-trained object detection neural network and a linear fit calibration is sufficient to obtain a good approximation of the true vehicle count.
- The proposed approach is simple but accurate, giving a counting error of about one vehicle.

**Motivations**
Knowledge of traffic flow variations and vehicle-related statistical analysis are key components of smart cities. Active technologies (RADAR, LIDAR, Radio-Based) for traffic monitoring require new hardware to be built, purchased, installed and maintained. There are significant costs of installation, maintenance and all the implications on acquiring the electronics.

Passive technologies can be used for traffic monitoring without requiring excessive additional costs.

**Use Case**
We propose a simple traffic monitoring system which leverages infrastructures (hundreds of cameras) that are already installed in many cities. As a proof of concept, we developed a counting method based on deep-learning object detection algorithms from camera images.

On a single image, the resulting vehicle count can be interpreted as a random sample of the road occupancy at the image collection time and aggregation statistics are calculated within two hours windows to smoothen the curve.

**Data**
- **Cameras**: 3 (#455, #430, #34)
- **Images**: Every 5-6 minutes, over several days
- **Format**: jpeg
- **Size**: 704x480, 704x280, 352x240
- **Quality**: Sporadically unavailable due to maintenance issues, incident resolutions or traffic studies
- **Orientation changes perturbing the location of the ROI**

**Model and Method**
The method consists of counting the number of vehicle annotations from an object detection model, in a region of interest (ROI), and to correct this raw counting (automatic annotations) based on a comparison with a ground truth (manual annotations).

**Limitations**
- **Manual (true)**: Local and Amazon Mechanical Turk (MTurk) campaigns. Quality of MTurk annotators evaluated by comparing them with local annotators.
- **Automatic (raw)**: You-only-look-once (YOLOv1) detector, a type of convolutional neural network which associates class probabilities and spatially separated bounding boxes, pretrained on COCO dataset.

**Calibration**

\[ N_{\text{true}} \approx N = C \cdot N_{\text{raw}} + C_0 \]

**Architecture Schema**

**Results**

- A naive exploration of adding the sun elevation angle feature as a proxy for luminosity was not conclusive.
- Adding the positions of the vehicles, based on the intuition that the object detection performs better with bigger/closer vehicles, could improve the model.
- We suggest to set the cameras to have a default point of view to avoid the correction of parameters defining the ROI in case of orientation changes. Otherwise, implement tools to detect angle changes, recognize the cardinal point of view or quantify an absolute/relative angle change.

**Conclusions**
We presented an object-detection-based traffic monitoring system using a single factor correction, per camera, to calibrate the neural network annotations in order to compensate the algorithm inefficiency.

A simple linear fit allows us to get a counting error of about one vehicle; with typically 4 vehicles per image (in daytime), this error can reduce to less than 10% of the counting range if we average the values over an hour. Although some improvements were evaluated and suggested, the less-than-a-car MAE already achieved by this simple method makes it appealing for many applications.