Identifying Perception Weaknesses and Training Candidates In Unlabeled Data

Auto.AI / February 2020
Overview

Perception Robustness
- Injecting noise into perception systems finds systemic problems

Where is perception just guessing?
- False confidence at the decision boundary

False negative prediction
- Weak detections followed by false negatives

Improving labeling effectiveness
- Automated edge case identification without human triage
What do we mean by a perception “edge case”? Autonomy can’t avoid what it doesn’t see...
Edge Cases include variations of everyday objects that are missing from training data.
Gaps in training data can lead to perception failure

- Safety needs to know: “Is that a person?”
- Machine learning provides: “Is that thing like the people in my training data?”

Edge Case are surprises

- Edge cases are the stuff you didn’t think of!
Adding noise causes objects to drop out

- Object list mismatches reveal perception triggering events

Simplified example: add light Gaussian Noise

Edge Cases:
- False negatives
- Novel objects
- False positives
- Combinations
Finding Edge Cases Experimentally

- Noise randomly perturbs data sample in decision space
  - Change in classification detects points near decision boundary
  - Many of these are Edge Cases (unknown unknowns)
Weak Detections Predict False Negatives

- Weak detections often correspond to false confidence
  - Perception says “90% confident” when it is just guessing
- In practice, we can use this to predict false negatives
  - Step 1: Perception drop-out on baseline images
  - Step 2: Near drop-out on augmented images
Problematic Flatbed Truck
A Few False Negatives Found:

“Camouflage”

“Children”

“Bare legs”

“Sun glare”?

“Red objects”

“Columns”

Baseline, un-augmented images with Mask-R CNN
// Your mileage may vary.
Many surprises aren’t obvious to humans
  ● Unlikely to be included in human-designed tests

Example:
  High visibility clothing missed by perception

Pilot study on real system:
  ● 82% recall of false negatives compared to ground truth
IDEA: Label & retrain on detected edge cases
- For common edge cases, no need for human analyst
Finding Perception Triggering Events

- Inject randomized noise into image
  - Differences in object lists happen at decision boundaries
  - Empirical field data “confidence” metric

- Manually examine issues found
  - False negatives predicted
  - Finds false positives
  - Finds defects in training data

- Use to automatically triage data for training/validation process