# **Coyote:** A Dataset of Challenging Scenarios in Visual Perception for Autonomous Vehicles

Suruchi Gupta^, **Ihsan Ullah***, Michael G. Madden^
^School of Computer Science, National University of Ireland Galway, Galway, Ireland
*CeADAR Ireland's Center for Applied AI, University College Dublin, Dublin, Ireland

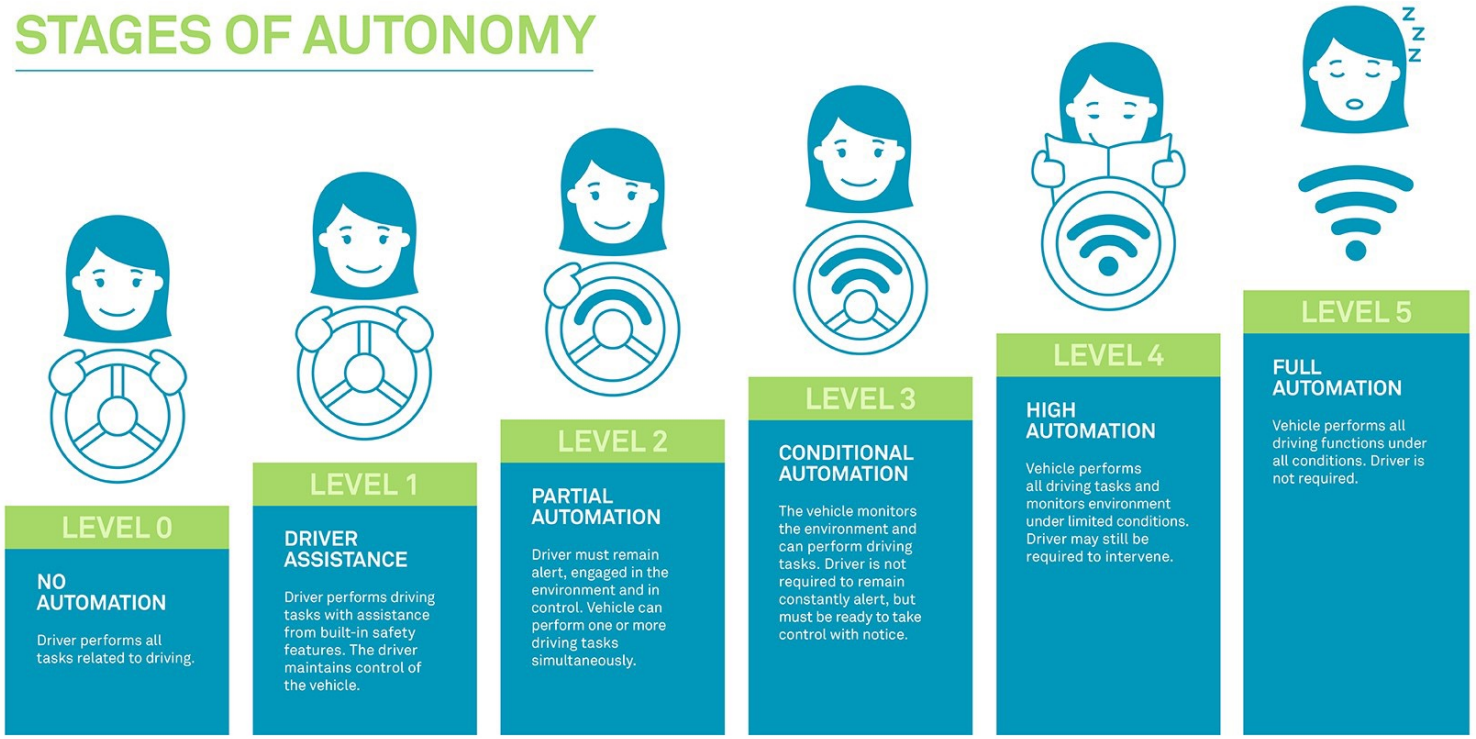
AISafety 2021
IJCAI-21 | Virtual | Aug 19-20, 2021

# Outline

- Introduction & Motivation
- Related Work
- Coyote Dataset
- Experiments
- Conclusions

# Introduction & Motivation

- Advancement in autonomous vehicles
  - Honda launched level 3 [1]
  - Tesla at level 2 [2]
  - Mercedes, Ford, and others are planning to launch later this year. [3,4]
  - Uses a combination of sensors, such as, LiDAR, Radar and camera to observe the surroundings
  - Video stream + computer vision





[1] https://japan-forward.com/honda-launches-worlds-first-level-3-self-driving-car/
[2] https://www.theverge.com/2021/5/7/22424592/tesla-elon-musk-autopilot-dmv-fsd-exaggeration
[3] https://www.cnet.com/roadshow/news/2021-mercedes-benz-s-class-level-3-autonomous-driving/
[4] https://techcrunch.com/2021/07/21/argo-ford-to-launch-self-driving-vehicles-on-lyfts-ride-hailing-app/

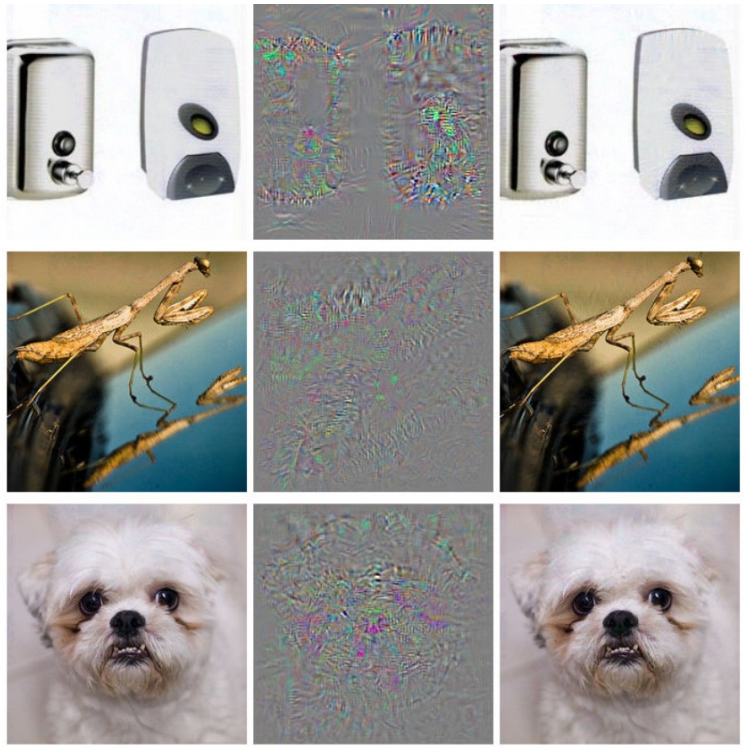
# Introduction & Motivation continued.



https://www.tesla.com/en_IE/autopilot

# Introduction & Motivation continued.

- Adversarial Examples
  - Natural or human-constructed images that cause the Machine Learning algorithms to make false predictions
  - Challenges the use of ML algorithms in safety-critical applications



Adversarial examples: (Left) is a correctly predicted sample, (centre) difference between correct image, and image predicted incorrectly (ostrich) magnified by 10x, (right) adversarial example.

[5] Szegedy, Christian, et al. “Intriguing properties of neural networks.” arXiv preprint arXiv:1312.6199 (2013)

# Existing Datasets

- For training algorithms for autonomous vehicles
  - KITTI [6], MS-COCO [7], etc.
- For training algorithms against adversarial datasets
  - WildDash [8], Adverse Conditions Dataset with Correspondences (ACDC) [9], FishyScapes (1030 images) [10]

FishyScapes: Place anomalous objects in front of the vehicle


[6] Menze, Moritz, and Andreas Geiger. "Object scene flow for autonomous vehicles." *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*. 2015
[7] Lin, Tsung-Yi, et al. "Microsoft coco: Common objects in context." *European conference on computer vision (ECCV)*. Springer, Cham, 2014
[8] Zendel, Oliver, et al. "Wilddash-creating hazard-aware benchmarks." Proceedings of the European Conference on Computer Vision (ECCV). 2018
[9] C. Sakaridis, D. Dai, and L. V. Gool. "ACDC: The Adverse Conditions Dataset with Correspondences for Semantic Driving Scene Understanding". arXiv e-prints, abs/2104.13395 , April 2021
[10] Blum, Hermann, et al. "The fishyscapes benchmark: Measuring blind spots in semantic segmentation." arXiv preprint arXiv:1904.03215 (2019)

# Coyote Dataset

- Aim:
  - Create a dataset of **naturally occurring scenarios**
    - potentially act as adversarial examples for Autonomous vehicles.
  - Evaluate the performance of the state-of-the-art object detection models on Coyote dataset
  - Identify the challenging scenarios for the autonomous vehicles and suggest mitigations.

# Coyote Dataset Continued.

- Overview of Dataset:
  - 894 images grouped into 7 broad categories
    - Object detection
      - Art in Surrounding and Murals
      - Vehicle Art and Textures
      - On-Road Scenarios
      - Street Signs
      - Parking Spaces
      - Advanced Scenarios
    - Semantic segmentation (19 images)
      - 3d Illusion
  - All the images in the dataset are:
    - In JPEG format
    - Publicly available
    - Unedited – except for 3 images that are cropped to reduce noise in the images

# Coyote Dataset: Art in Surrounding and Murals


Sample murals re-creating road scenarios

Sculptures that resemble road objects such as cars and trucks



Paintings on road depicting safety threat;

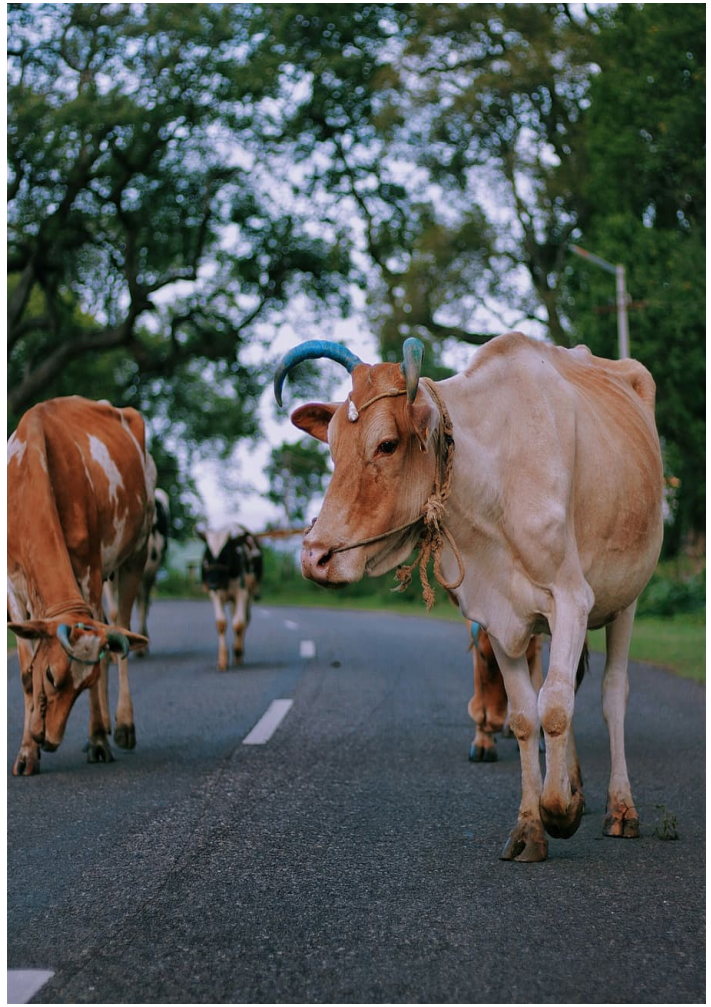
# Coyote Dataset: Vehicle Art and Textures


Motor vehicles camouflaged as other objects

Custom built motor vehicle

Vehicles with Textures

# Coyote Dataset: On-Road Scenarios



Photos of animals on roads


Birds on the Street

Challenging road scenarios


Road-side signs with pictures of humans

# Coyote Dataset: Street Signs

Art on Street Signs


Regional Variations of Street Signs

Custom traffic signals on the road

# Coyote Dataset: Parking Spaces

Unforeseen objects in Parking Spaces



Unconventional Parking Signs and Warnings

# Coyote Dataset: Advanced Scenarios


Additional artistic creations

Natural calamities


Examples of animal crossing signs

# Coyote Dataset: 3D illusions



3-dimensional illusions that may challenge semantic segmentation

# Experiments

- State-of-the-art models used for experimentation:
  - YOLO v3 [11] for Object detection
  - Faster R-CNN [12] for Object detection
  - DeepLab v3 [13] for Semantic Segmentation

- The models were pre-trained on MS COCO dataset

- To compare with the benchmark datasets, we used MS-COCO validation set and a random subset from KITTI dataset

[11] Redmon, Joseph, and Ali Farhadi. "Yolov3: An incremental improvement." *arXiv preprint arXiv:1804.02767* (2018)
[ 12] Ren, Shaoqing, et al. "Faster r-cnn: Towards real-time object detection with region proposal networks." *Advances in neural information processing systems* 28 (2015)
[13] Chen, Liang-Chieh, et al. "Rethinking atrous convolution for semantic image segmentation." *arXiv preprint arXiv:1706.05587* (2017)

# Experiments – YOLOv3


Successful classifications


Unsuccessful classifications

# Experiments – Faster R-CNN


Successful classifications


Unsuccessful classifications

# Experiments – DeepLabv3


Sample examples for Semantic segmentation

# Experiments – Cumulative Results

| Category Name | Model | Accuracy | Precision | Recall | F1-score |
|---|---|---|---|---|---|
| Coyote Total | YOLOv3 | 0.55 | 0.76 | 0.67 | 0.71 |
| | FRCNN | 0.52 | 0.60 | 0.79 | 0.68 |
| MS-COCO 2017 Val set | YOLOv3 | 0.46 | 0.79 | 0.53 | 0.63 |
| | FRCNN | 0.53 | **0.91** | 0.56 | 0.69 |
| KITTI Testing subset | YOLOv3 | 0.63 | 0.81 | 0.74 | 0.77 |
| | FRCNN | **0.76** | 0.86 | **0.87** | **0.86** |

# Conclusions

- A new publicly available test dataset
- Paintings in *Art in surrounding and Murals* category **confuse** the models the most
  - Both YOLOv3 and Faster R-CNN models perform worse on this category
- In *Advanced Scenarios* category with *3-D art on roads*,
  - DeepLab fails in some cases when 3-D representations of objects are painted on walls.
- Faster R-CNN model captures more details than the YOLOv3 model
- Future Directions:
  - Risks Mitigation using Sensor Fusion, Common Sense Reasoning, Better Treatment of Scale, or Spatio-temporal Reasoning
  - Extending Coyote dataset e.g. Adding new images, annotations, and other forms of data