Association for Information Systems
AIS Electronic Library (AISeL)

Proceedings of the 2018 Pre-ICIS SIGDSA Symposium

Special Interest Group on Decision Support and Analytics (SIGDSA)

12-12-2018

A Perception Augmentation System for Autonomous Vehicles

Christian Kauten
Auburn University, kautenja@auburn.edu

Ashish Gupta
Auburn University, azg0074@auburn.edu

David Bevly
Auburn University, dmbevly@eng.auburn.edu

Xiao Qin
Auburn University, xqin@auburn.edu

Han Li
University of New Mexico, hanli@unm.edu

See next page for additional authors

Authors
Christian Kauten, Ashish Gupta, David Bevly, Xiao Qin, Han Li, and Alison Jenkins

A Perception Augmentation System for Autonomous Vehicles

Prototype Demonstration

Christian Kauten
Auburn University
kautenja@auburn.edu

Ashish Gupta
Auburn University
azg0074@auburn.edu

Xiao Qin
Auburn University
xqin@auburn.edu

Han Li
University of New Mexico
hanli@unm.edu

David Bevly
Auburn University
dmbevly@eng.auburn.edu

Alison Jenkins
Auburn University
alj0032@auburn.edu

Abstract

We describe a system prototype for perception augmentation in autonomous vehicles. The system is built using a fully convolutional deep encoder-decoder architecture to map pixels with depth measures to semantic class labels. Class labels recombine with depth measures to produce a 3-dimensional semantic map of the objects in front of the vehicle. The map, simplified to highlight areas of importance (e.g., other vehicles, pedestrians), is shown to the passenger using a novel user interface. The map is also analyzed for potential risks to queue alerts to the passenger. Alerts are both: (1) shown to the passenger using an addressable LED strip around the windshield, and (2) delivered to the passenger through a speaker.

Keywords

Convolutional Neural Networks, Perception Augmentation, Autonomous Vehicles

Introduction

Researchers and consumers alike show growing interest in autonomous features for automobiles. Indeed, such features provide a means of reducing the inherent risk of operating a personal vehicle; however, recent news suggests that there are flaws in Artificial Intelligence (AI) technology powering Autonomous Vehicles (AVs). These events decrease public trust in the AI technology, a trust that has been compromised by the discomfort of relinquishing control of a dangerous activity to a machine (McAllister et al., 2017). To bridge this deficit in trust, this work presents an AI system to augment human intelligence in AVs.

Semantic Segmentation Model

We apply a fully convolutional neural network based on the 56-layer Tiramisu to map pixels to semantic class labels (Jégou et al., 2017). This architecture features salient qualities including: (1) impressive quantitative and qualitative performance, and (2) low parameter count. To meet the real-time requirement of our system, we introduce a smaller, performance oriented Tiramisu with 45 layers. This 45-layer Tiramisu is identical to the 56-layer Tiramisu, but with one less convolutional layer per dense block and bottleneck block. Trained on the CamVid benchmark (Brostow et al., 2009), the network achieves 90% categorical accuracy, only 1.5% less than the higher capacity model with 103 layers. We train the 45-layer Tiramisu with a fourth input channel for depth, and concatenate the depth channel with the inferred targets at test time to produce 3-dimensional semantic maps. Figure 1 describes a way of passing an input tensor $\mathbf{X}$ through the model to produce an output tensor $\mathbf{M}$.

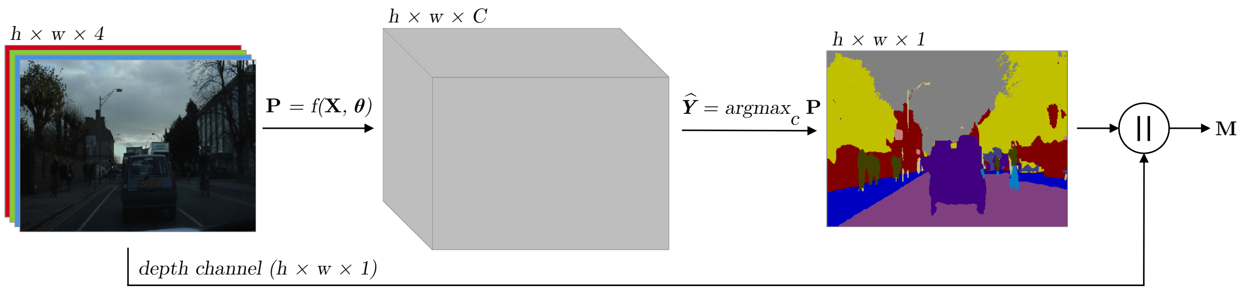

Figure 1. The perception task is modeled as a Convolutional Neural Network (CNN) $f(\boldsymbol{x}, \boldsymbol{\theta}) : \mathbb{R}^{h \times w \times 4} \to \mathbb{R}^{h \times w \times C}$, where C is the number of labels in the classification task. An RGB-D input tensor $\mathbf{X}$ passes through the CNN to generate a probability tensor $\mathbf{P} = f(\mathbf{X}, \boldsymbol{\theta})$. For each pixel, the label with the highest probability is taken as the predicted label, i.e., $\hat{Y} = \arg\max_{c \in C} \mathbf{P}_{:,:,c}$. The depth channel, $\mathbf{X}_{:,:,4}$, is concatenated with the predicted targets, $\hat{Y}$, to produce a 3-dimensional semantic map $\mathbf{M}$.

Perception Augmentation System

We build our perception augmentation system, depicted in Figure 2, to augment passenger perception of driving environments. The overall goals are to: (1) increase situation awareness, (2) reduce passenger anxiety, (3) improve immersion in the driving and monitoring tasks, and (4) increase trust in the AI technology in AVs. We pass the 3-dimensional semantic map $\mathbf{M}$ from our model through an alert controller to determine the instantaneous risk of the scene and both (1) control individual LEDs in real-time, and (2) enqueue sound events to a speaker. $\mathbf{M}$ also flows through a simple User Interface (UI) controller to be cleaned and presented to the passenger. For simplicity and demonstration purposes, we design use cases only for `Pedestrian` and `Car` classes from our datasets.

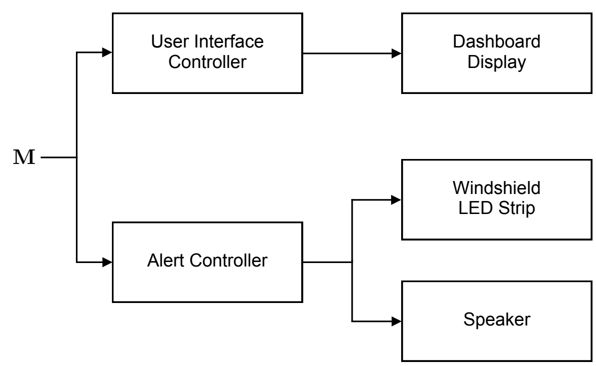

Figure 2. Perception Augmentation System

Alert Controller

To reduce passenger anxiety and improve immersion, we devise an alert controller to map $\mathbf{M}$ to an LED strip and auditory events. The LED strip mirrors the semantic information from $\mathbf{M}$ around the border of the windshield to draw passenger attention to classes in the scene using colors coded to semantic labels with brightness determined by depth. Similarly, the speaker announces classes in the scene that are close to the vehicle and recommended or planned actions (see, for example, *brake*).

User Interface Controller

To raise trust in autonomous operation and increase situation awareness, a novel UI renders data in $\mathbf{M}$ for passengers. The UI controller reduces the class masks in $\mathbf{M}$ into basic geometric shapes and displays the regions using a the same color coding as the LED strip. Further, the UI controller presents the distribution of classes and depths in $\mathbf{M}$ as a situation measurement.

Future Work

As a future research direction, we will investigate the efficacy of our perception augmentation system in terms of passenger behaviors like situation awareness, anxiety, immersion, and trust.

References

Brostow, G. J., Fauqueur, J., and Cipolla, R. 2009. "Semantic object classes in video: A high-definition ground truth database," *Pattern Recognition Letters* (30:2), pp. 88–97. Video-based Object and Event Analysis.

Jégou, S., Drozdzal, M., Vázquez, D., Romero, A., and Bengio, Y. 2017. "The One Hundred Layers Tiramisu: Fully Convolutional DenseNets for Semantic Segmentation," *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* pp. 1175–1183.

McAllister, R., Gal, Y., Kendall, A., Van Der Wilk, M., Shah, A., Cipolla, R., and Weller, A. 2017. "Concrete Problems for Autonomous Vehicle Safety: Advantages of Bayesian Deep Learning," in *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, .