**Abstract**

**Authors:** Samuel Cowin, Jianghong Wan, Vivaswat Suresh

With the rise in Deep Learning and the success of Autonomous Vehicles both in utilizing this technology and commercially, investigations into this field for furthering these developments are crucial. One realm that has garnered much attention is the application of Deep Learning for Semantic Segmentation of scenes for understanding. Within this study, we will examine various approaches to Semantic Segmentation [2] drawing from implementations completed on the popular Cityscapes dataset. To do this, we start with popular CNN architectures (i.e., VGG, AlexNet, ResNet), tweaking and modifying them to change the output architecture from classification to do per-pixel identification as outlined in *Fully Convolutional Network (FCN)* by Long et al [1]. This will be done on the Kaggle version of the dataset taken from Berkeley [3], and, time permitting, extended to larger versions of the data for further understanding [4].

**References:**

1. J. Long, E. Shelhamer, and T. Darrell, “Fully convolutional net- works for semantic segmentation,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 3431–3440.
2. Alberto Garcia-Garcia, Sergio Orts-Escolano, Sergiu Oprea, Victor Villena-Martinez and Jose Garcia-Rodriguez. "A Review on Deep Learning Techniques Applied to Semantic Segmentation," 2017.
3. Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. "Image-to-Image Translation with Conditional Adversarial Networks," 2018.
4. M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, “The Cityscapes Dataset for Semantic Urban Scene Understanding,” in *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.