The City College Visual Computing Laboratory (CCVCL) serves as an experimental environment for both research and education in advanced computer vision and other media computing. The research activities in the CCVCL primarily focus on the understanding of 3D natural scenes and the analysis of events in these scenes via multiple sensor modalities, including visible cameras, thermal sensors, and acoustic sensors. The research seeks to improve state-of-the-art machine learning methods while simultaneously applying techniques to real-world applications. Our work is supported by NSF, DoD, DHS, and industry partners. Our lab’s graduates are now working at Amazon, BAE Systems, CUNY, Nvidia, and AI startups.
This project studies automated emotion recognition using four different modalities – audio, video, electromyography (EMG), and electroencephalography (EEG). We collected a dataset using the 4 modalities with 12 human subjects, each expressing 6 different emotions or maintaining a neutral expression. To improve the predictive power of our emotion recognition system, three different aspects of recognition have been investigated: model selection, feature selection, and data selection. This is a collaborative work with Prof. Tony Ro at CUNY Graduate Center.

Using CNNs for crowd analysis with minimal labeled data

In this work, we use a generative adversarial network (GAN) to train crowd counting networks using minimal data. Our study focuses on how GAN objectives can be modified to allow for the use of unlabeled data to benefit inference training via semi-supervised learning. More generally, we explain how these same methods can be used in more generic regression target semi-supervised learning, with crowd counting being a demonstrative example. We demonstrate that a convolutional neural network (CNN) trained as part of the semi-supervised GAN can significantly outperform an equivalent CNN trained with standard supervised learning techniques, even when the latter has access to substantially more labeled data.

Emotion analysis using audio/video, EMG & EEG

This project studies automated emotion recognition using four different modalities – audio, video, electromyography (EMG), and electroencephalography (EEG). We collected a dataset using the 4 modalities with 12 human subjects, each expressing 6 different emotions or maintaining a neutral expression. To improve the predictive power of our emotion recognition system, three different aspects of recognition have been investigated: model selection, feature selection, and data selection. This is a collaborative work with Prof. Tony Ro at CUNY Graduate Center.