Preeti Iyer
COS 429 – Project Summaries

Project 1 :: Using Deep Learning on Google Street View Images to Predict Poverty

This project explored the extraction of economic and development indicators using feature extraction from Street view Images, with the hopes that using the spatial information found in these images they could improve the ability to predict economic consumption of a certain geolocation. They iterated on the Places CNN (a CNN trained for scene recognition) to apply it to Ethiopian street view images, extract feature maps, then feed those into a ridge regression model to output an economic consumption rate. They also found that the existing VCG model worked well for some locations with good feature extraction but not others, in particular it worked well for Tanzania but poor for Ethiopia because of the presence of the features (nightlights) that the existing model uses to predict economic consumption. Overall building a model that can predict economic factors like consumption is very useful for governments to have appropriate knowledge and enact policy, so I found their iteration on exist work in this field to be very interesting.

Project 2:: How Generalizable is Human Movement?

This project aimed to build off of an urban studies project to incorporate computer vision/graphic to predict and simulate the movement of people in a given image of any landscape. The implementation this project sought to build off of currently only trains their model based off of a certain type of images; Urvashi sought to change the input data to the codebase to see how the different data would impact the weights of the features and thus the interactions of humans with a given space. The actual process for her implementation involved capturing 3 different sources of intermediary data to feed into the IDC model and return reward weights of predicted trajectories. Her implementation had higher response weights for grassier spaces, showing that with the additional training of new scenes/spaces, the predicted human paths changes and thus human movement cannot be generalized from just a particular set of scenes.

Project 3:: Convolutional Neural Networks for Infant Attention Prediction

In development psychology research it's essential to understand if you have the attention of an infant while understand how the infant is responding to a certain visual stimulus. However right now the *Princeton Baby Lab,* which does a lot of this research, does not automate this process and manually flags each frame of an experiment. This project sought to automate this process; their implementation took in an input image, used OpenCV Multiscale face detection to detect faces, used a Pytorch based CNN to indicate whether the face was a baby face or a parent face (parents often do these experiments with their children), then using the baby faces ultimately used a constructed CNN to indicate whether an infant face is looking straight ahead – based on front face detection. They validated their final CNN using 10 fold cross validations and test various structures to increase accuracy to an ultimate accuracy of 95.71%. The overall accuracy of the system, when taking into account all 3 detection/classification steps was 83.91%; they ultimately factored this into their acceptance threshold for "infant attention" to create a useful system that can be used by the Baby Lab to speed up experimental trails.