Hsin-Hung Wu: henrywu@bu.edu
Hanyu Chen: chenhy1@bu.edu
Israel Ramirez: israelr@bu.edu

# Parking Lots Project final report

## Goal

Our goal is to update the City of Boston's outdated municipal parking lots data set. Based on the specified tax exempt codes and the coordinates, we first have to collect the parcel images of the relevant years(2019- 2020) and run them through a parking lot classifier with an expected accuracy of at least 85%. The classifier will identify all the parking lots for us to analyze and compare.

## Data Collection(screenshotScript.py & index.html):

The data sources we have are Mass GIS, Parcels 2020 data/Property Assessment FY202 data/ Parking Meters data from data.boston.gov. These data contains the tax exempt codes and coordinates of the parcels and we are mainly focusing on tax exempt code 986, 987, and 962. After we obtained the coordinates (latitude/longitude), we wrote a script that utilizes Google StreetView API to get the aerial images and we then screenshotted and cropped them into proper sizes. Now we are able to obtain all the parcel images we need automatically.

## Data Preprocessing:

With all the parcel images, our goal is to figure out whether an image contains parking lots or not by feeding them into a CNN binary image classifier. We manually labelled some images to create a training data set, but the data set is too imbalanced with too few parking lot images, which is important for this binary classification task. We tried to get parking lot images from external sources, but most of the images are really different from what we obtained in terms of angle, views, and varieties. We balanced out the two classes and we have only around 230 training images and 70 testing images. The validation set are randomly split from 20% of the training set.

We first preprocess the images in order to maximize the performances and deal with the lack of training images:

1. Images rescale to $\frac{1}{255}$
2. Image resize to (224, 224) to fit model input size
3. Random image rotations, horizontal/vertical flip, and shifting
4. Feature wise center to transform data set to 0 mean
5. ZCA whitening to decorrelate features

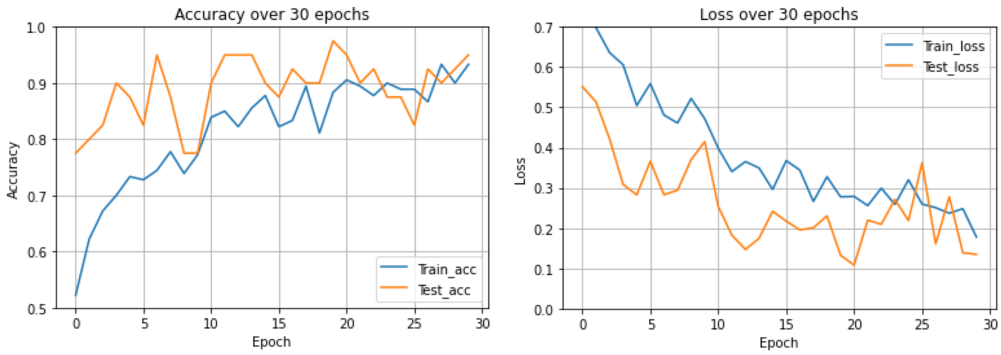
**Classifier Model (ParkingLotClassifier.ipynb):**

We decided to do transfer learning since training a CNN from scratch would take too much time. We use VGG16 with pre-trained weights from ImageNet, because this model deals with similar classification problem with our task. We dropped the top 3 fully connected layers and add our own layers as shown below:

$$\boldsymbol{Flatten \rightarrow Dropout(0.5) \rightarrow Dense(256, relu) \rightarrow Dropout(0.5) \rightarrow Dense(1, sigmoid)}$$

We also fine-tuned the top two layers of the VGG16. We trained with batch size of 10 with Adam optimizer with learning rate of 0.0001 and for 30 epochs.

The training accuracy and validation accuracy grows at around the same rate, which means there isn't much overfitting and the model performs quite well at around 90% training and validation accuracy. More detailed classifier report is shown in ParkingLotClassifier.ipynb.




Loss : 0.4689
Test accuracy is 86.36%

Test accuracy of 86% is descent but can be improved if we have more parking lot images.

**Analysis:**

After analyzing the classified images, we noticed that there is only small portion(15%) of images that contain parking lots. The number of parking lots are decreasing drastically overall time. Covid-19 also affect parking behaviors as less people are leaving home for work during confinement. Residential area has an increase in average occupancy from 84% to 93.4%. Office and cultural areas have decrease in occupancy levels from 49.3% to 26.6%.

## Bibliography

"How the Covid-19 Pandemic Affected Parking Behaviour." *IEM*, 1 July 2020, www.iemgroup.com/2020/07/01/how-the-covid-19-pandemic-affected-parking-behaviour/.