May Zisbrod
Deep Learning for Computer Vision
Prof. Peter Belhumeur
Columbia University
Fall 2024

Visual Insights in Manhattan: A Deep Learning Analysis of Rental Listings

- Participants - May Zisbrod (Individual project)
- Objectives - This project aims to uncover correlations between visual features of rental listings of studio and one-bedroom apartments and various listing attributes, such as price, neighbourhood, and amenities. Using StreetEasy dataset extracted using RapidAPI, which includes rental information and property images, the project will explore how visual elements—such as room brightness, decor style, and space utilization—might influence rental prices and appeal. The primary objectives are:
  1. Price Prediction from Visual Cues - Determine if image-based features can predict rental prices in Manhattan neighbourhoods.
  2. Amenity Detection - Assess if visual characteristics correlate with specific amenities or property types.
  3. Neighbourhood Style Analysis - Identify unique visual patterns in neighbourhoods and explore how aesthetics vary across regions.
- Plan for Completion
  - Training Data & Preprocessing - The dataset includes images and structured data (e.g., price, neighbourhood, amenities). Images will be preprocessed to standard dimensions and enhanced to highlight details for model training. Structured data will undergo feature engineering, creating visual and non-visual feature columns (e.g., year built and neighbourhood).
  - Deep Learning Framework – I'll use PyTorch because of its flexibility and support for CNN architectures. This choice enables the integration of pre-trained models and fine-tuning, essential for extracting visual features efficiently.
  - Network Architecture - A Convolutional Neural Network (CNN) will be employed, utilizing pre-trained layers to expedite feature extraction. A dense layer will be added to integrate visual embeddings with structured data to predict rental prices and amenities.

- Predicted Success - This model is expected to identify strong correlations between visual and non-visual features, achieving moderate accuracy in predicting price ranges and detecting neighborhood aesthetics.