Milestone Report
Cheng Fei, Tianyu Lou, Wenjie Wang
10/31/2023

Introduction/Motivation

Have you ever strolled down a street or traveled to a new place, only to stumble upon a building of historical or unique significance? Did you find yourself curious about its architectural style and the stories it holds? Across time, countless renowned architects have left their mark with distinct styles. Yet, for humans, it can be a challenge to recognize them all. Instead, we're embarking on a journey to create an architecture style classification model, capable of discerning architectural styles based on their visual characteristics.

This model can be applicable in several scenarios; for instance, tourists can take photos of a building on the road and get its architectural style in real time. Specifically, this model is designed to recognize architectural styles from different historical eras or different architects' designs.

Dataset

Our dataset will consist of images showcasing various architectural styles. Unfortunately, there isn't an available dataset that enables us to initiate the data preprocessing and model fitting process. Consequently, we must generate the dataset from the ground up. Our approach involves extracting images from online sources like Google Images and assigning each image a corresponding label manually. In order to increase the quality of scraped datasets, we optimize the python crawler to filter out noisy datasets. Presently, we are aiming for 25 distinct styles, each represented by 400 images on average. Generally, these architectural styles are among the most common ones people can find in their daily lives, which ensures the applicability and feasibility of the model. Subsequently, we will undertake data preprocessing tasks, ensuring that all images are standardized in size without sacrificing any key features, and noisy information is removed. This phase typically consists of reorientation and rescaling. Furthermore, we will produce more data through data augmentation, such as adjusting the sharpness, color, brightness, and contrast, in order to optimize the model accuracy.

Model/Method

We use supervised learning with a convolutional neural network. We'll first determine iconic architectural styles in history and scrape the dataset accordingly from the internet. Then, we'll perform data preprocessing as well as data augmentation according to the scale and quality of our dataset. We will define our model class based on the number of classes and features,

choosing the most compatible convolutional neural network and defining our objective and optimizer according to the model we employed. Finally, we'll yield a predictive model.

To improve the performance of our model we plan to experiment with approaches such as creating multiple neural network models and select the one that has the best performance. Additionally, we may consider augmenting the complexity by adjusting the number of hidden layers and parameters. To rigorously evaluate the model's performance, we will employ cross-validation techniques.

Preliminary experiments

Currently, we have constructed a fully functional machine learning workflow using Google Collab, implementing dataset import, splitting into training set and validation set, performing data processing as well as augmentation, setting up a demo neural network, and finally performing a prediction from a standalone dataset using a single picture.

We have also established a dataset containing

Future Work

First, our priority is to generate a large dataset. We aim to incorporate the criteria outlined in the dataset section, encompassing 25 distinct styles, each supported by an average of 400 images. Presently, we possess a limited set of images sourced from an architecture website. However, we intend to develop a web script to automate the acquisition of the dataset, ensuring a targeted and representative data sample.

Second, we are to enhance the model's accuracy. We plan to experiment with approaches such as creating multiple neural network models and select the one that has the best performance. Additionally, we may consider augmenting the complexity by adjusting the number of hidden layers and parameters. To rigorously evaluate the model's performance, we will employ cross-validation techniques.

Third, considering the complexity of real-life datasets and the fact the architecture pictures online are well-polished and oriented, we will test our model with randomly taken photos on the road; for instance, there is a wide range of diverse architectural styles in Manhattan, we can take photos there to build our experimental datasets.

Finally, we plan to create either a mobile application or a web application, contingent on available time and resources. For the mobile application, we envision a real-time architecture classification feature, allowing users to effortlessly point their phone at a building to identify its

architectural style. Alternatively, the web application will enable users to upload images of buildings, providing an immediate output of the architectural style.