Machine Learning Engineering Nanodegree Classification

Capstone Proposal

Wan Li

May 15, 2018

Architectural Picture Classification

Domain background

This project is aimed to analyze a picture of a building in terms of the style of buildings and the composition of the photo.

Though photos and renderings of buildings may have infinite possibilities, many of they can be classified into a limited number of popular types according to the composition, style of buildings, etc.

I would roughly classify exterior pictures of (mid-rise to high rise) buildings into a few categories according to the scene, point of view, including street view, far away view, mid-air view, top view and look up view. There are usually other objects like people, cars, boats and animals in the picture, which I will also identify.

The judgement of a style of architecture is kind of subjective without very concrete standards now, and a person need to learn a lot about architectural history and current famous architects to tell the possible style of a building. I want to design an algorithm to identify the possible style of building in a photo or rendering. For this project, I just focus on about 10 common architectural styles.

Based on this project, later I want to develop a project to help generate pictures of buildings based on input, view, objects and architectural style.

Previous Similar Research

There are some existing researches about classifying architectural styles and scenes. Most of them are about classifying traditional styles:

Classification of Architectural Heritage Images Using Deep Learning Techniques: https://www.researchgate.net/publication/320052364_Classification_of_Architectural_Heritage_Images_Using_Deep_Learning_Techniques

Classifying U.S. Houses by Architectural Style Using Convolutional Neural Networks:

http://cs231n.stanford.edu/reports/2017/pdfs/126.pdf

Automatic Architectural Style Recognition:

https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XXXVIII-5-W16/171/2011/isprsarchives-XXXVIII-5-W16-171-2011.pdf

Training an Architectural Classifier:

https://hackernoon.com/training-an-architectural-classifier-5f1b4f512368

Research on Classification of Architectural Style Image Based on Convolutional Neural Network

https://ieeexplore.ieee.org/document/8122517/metrics?part=1

Problem statement

Given a picture of a building, I want the algorithm to tell the view of the picture, the obvious objects like people and cars and the most possible style of the building. The view will be among a few classes I defined and the style will be among 10 styles that I choose for the project. The obvious objects will give a name of the object and its coordinates.

This is an image recognition and classification problem. These problems should be solved by machine learning algorithms such as SVM, CNN, etc.

Datasets and inputs

I would mainly use two data sets for this project: *arc_scene* and *arc_style*. The dataset *arc_scene* includes about 1000 pictures of buildings labeled according to its scene. The dataset *arc_style* includes 2000 pictures of buildings labeled according to its architectural style. I selected 12 styles that are far from inclusive, but they are very distinguishable and common in the architectural design industry. Each style has around 100 or 200 images.

Unlike traditional buildings, there is no clear criterion for classification of modern architecture. So I defined some styles based on knowledge of architecture. There are well known traditional architectural styles and the modern styles are mostly based on the work of a famous architecture with strong features. Here are some examples:

Frank Gehry:

Zaha Hadid:

Mies:

All the pictures have been manually collected from Google Images, Pinterest, Baidu Images and other websites and labelled by myself. These labelled pictures will be used as input for machine learning for classification.

The images have varying sizes bigger than 300*300, and will be resized as 256*256 when loaded.

Solution statement

Identifying the scene of image may be relatively simple. I will try SVM and CNN. To recognize objects like people and cars, I will use YOLO system to detect the objects in the picture. To predict the architectural style, I will use CNN with transfer learning.

Benchmark model

There may be two benchmark models, one model is to guess randomly. A random guess will provide a correct architectural roughly 1 in 10 times, corresponding to an accuracy rate of about 10%. Another benchmark model is a simple CNN or SVM model built by myself without transfer learning.

Evaluation metrics

I would use accuracy rate and a confusion matrix as evaluation metrics to evaluate the machine learning classification performance. Accuracy rate is the rate of correct predictions. Confusion matrix will give detailed information about the predictions.

For this project, I will simply use the YOLO system as it is without training or evaluating the model statistically because the dataset is not labelled in terms of objects. I will test it with a few test cases.

Project design

Step 1: Import and preprocess (resize, augment, etc.) the two datasets.

Step 2: Use SVM to classify scenes.

Step 2: Create a CNN to classify scenes.

Step 3: Write an algorithm with YOLO system to find the objects in the picture.

Step 4: Create a CNN from scratch to classify architectural styles.

Step 5: Use transfer learning to create a CNN to classify architectural styles.

Step 6: Write an algorithm to analyze a picture of a building.

Step 7: Test my algorithm.