Classifying U.S. Houses by Architectural Style

Author: Chris Pesto

Background

There are a variety of house styles used for homes across the United States, each with its own regional and historical significance.

This explores how well convolutional neural networks can identify house architectural styles in real home listings.

Dataset

N=2500 images handpicked from Zillow.com – essentially all real listings from various U.S. states

5 classes, 500 images each:
- Colonial
- Craftsman
- Mediterranean
- Ranch
- Victorian

Different sizes and aspect ratios
Multiple sub-types within each class

Data Processing

Resized proportionally to fit in 256x256 pixels

Placed on a black background to form a 256x256 square

Training Set Methodology

20% of each class randomly selected and set aside as a test set

20% of each class randomly selected as a validation set

Final 60% (300 images per class) used as training data

Flip-LR data augmentation used on training data, yielding 600 examples per class

Problem Statement

Two objectives:
- Classification
  - Classify the images by architectural style
  - Softmax cross entropy loss
- Localization
  - Identify the (primary) house in the image with a bounding box
  - Sum of the squared differences between each coordinate and ground truth

Methods

- Baseline
  - Convolutional layer: 2 x [Conv – BatchNorm – ReLU – MaxPool]
  - Classifier FC:.[Linear-128 – BatchNorm – ReLU – Linear-5]
  - Box FC: [Linear-1024 – BatchNorm – ReLU – Linear-4]
- ResNet-18 as fixed feature extractor
  - PyTorch implementation pretrained on ImageNet, with FC layer replaced
  - Classifier FC: [Linear-5]
  - Box FC: [Linear-4096 – BatchNorm – ReLU – Linear-1024 – BatchNorm – ReLU – Linear-4]

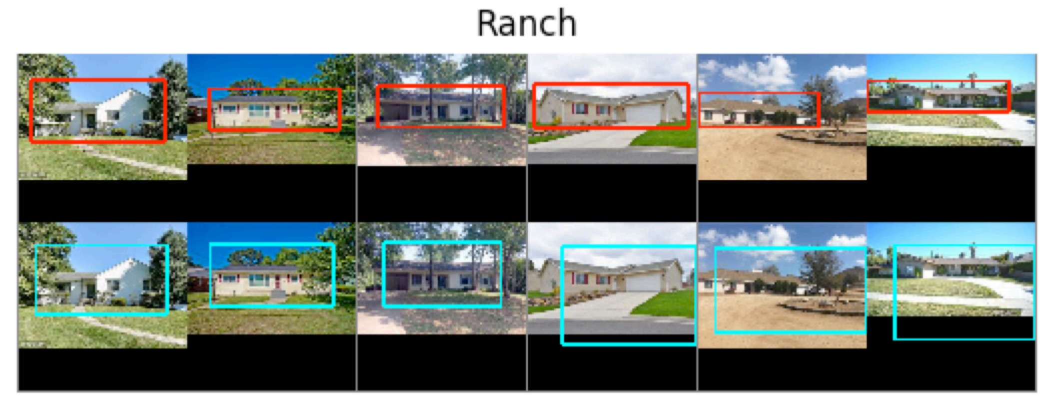
Results

Best classifier accuracies:
- Baseline: 60%
- ResNet-18 as feature extractor: 75%

Baseline model bounding boxes (3 lowest loss, 3 highest loss for each class):

Victorian

Ranch



Note: classification and bounding box results are from different training runs with data augmentation on and off. I'm working on getting these to work together.

Related Work

Obeso, et al in Dec. 2016 used CNNs to classify Mexican historical buildings.

Other work has applied machine vision to surveying and urban planning.

Future Directions

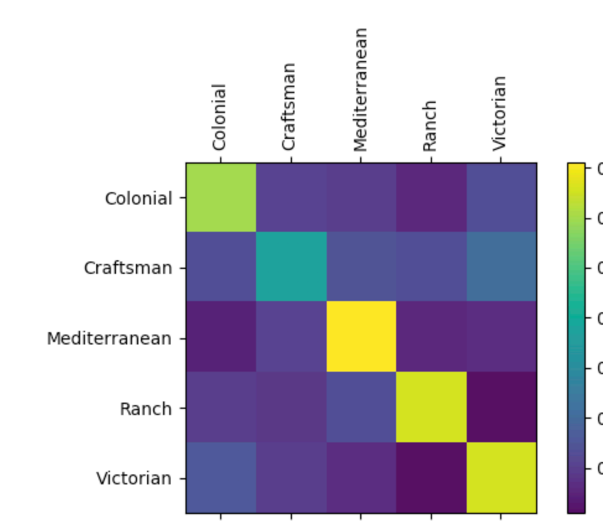
This demonstrates that CNNs can classify architectural style with reasonable accuracy.

Applications include surveying and better filtering on real estate searches.

The models train quickly – the largest impediment is data collection.

Collecting a large corpus with many styles could give interesting insights into the distribution of house styles in the U.S.

Baseline model confusion matrix:



ResNet-18 model confusion matrix: