Deep Architectural Style Classification
Paavani Dua, Alan Flores-Lopez, Alex Wade
{alanf94, paavanid, awade}@stanford.edu
Stanford University | CS230 | Winter 2019

Problem Statement
Given an image of a building facade, what is its architectural style?
- Byzantine?
- Romanesque?
- Colonial?

Dataset
- 5053 images, 25 classes of architectural styles. Created by Xu et al. (2014) for their regression classifier.
- We added 28th class with no architectural style
- Split: 70% train, 20% dev, 10% test, preserving class dists.
- Data augmentation: horizontal flipping, crops, rotations

Models & Methods
- Cross-Entropy Loss = −1/|D| ∑d∈D [ydlog(ˆyd) + (1−yd)log(1−ˆyd)]
- Total Loss = Cross-Entropy Loss + λ1L1(θ) + λ2L2(θ)
- Feature extractor + NN
- Tuned hyperparameters over five pretrained feature extractors: MobileNet V2, Inception V3, Inception ResNet V2, NASNet and ResNet V2
- MobileNet V2 achieved the highest dev accuracy along with a faster runtime
- Separately, fine-grained classification via object detection and classification

Results
Table 1. Overall Test Accuracy
<table>
<thead>
<tr>
<th>Xu et al. (2014)</th>
<th>Baseline CNN</th>
<th>Human Expert</th>
<th>MobileNet Transfer</th>
</tr>
</thead>
<tbody>
<tr>
<td>46.2%</td>
<td>55.4%</td>
<td>56.0%</td>
<td>75.7%</td>
</tr>
</tbody>
</table>

We evaluated our model against:
- Xu et al. (2014)’s results on the same task
- Our baseline CNN
- A human expert: trained architect

Figure 1. Illustrates the difficulty of the task and inconsistencies in the dataset itself: both houses could equally well be American Craftsman, but one was labeled as American Foursquare in the original data. Our classifier marked both as American Craftsman.

Figure 2. Depicts our fine-grained approach, analyzing a mixed-style house by classifying each smaller feature of that house. The overall image was classified as Queen Anne, the door Georgian, the house ‘Novelty,’ and the window American Craftsman.

Discussion
- Variance still a problem after L1/L2 dropout regularization and data aug
- Modern techniques achieve higher accuracy than previous work and human expert
- Existing dataset has flaws, as pointed out by human expert (e.g. pervasive mixing of classes and individual wrongly-classified images)
- Task has flaws: mixed styles
- Some architectural styles are considerably more difficult than others (Ancient Egyptian vs. International)

Future Directions
- Expert-curated dataset with all architectural features classified, not just entire image
- Saliency maps
- Classify architectures in real time

References