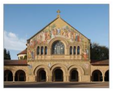
Deep Architectural Style Classification

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?



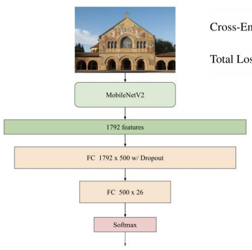
Byzantine?
Romanesque?
Colonial?

Dataset

(a) Byzantine (b) Romanesque (c) Tudor Revival

- 5053 images, 25 classes of architectural styles. Created by **Xu et al. (2014)** for their regression classifier.
- We added 26th class with 'no architecture' style
- **Split:** 70% train, 20% dev, 10% test, preserving class dists.
- **Data augmentation:** horizontal flipping, crops, rotations

Models & Methods



$$\text{Cross-Entropy Loss} = \frac{1}{|D|}\sum_{(x,y)\in D}\left[-f_y + \log\sum_j e^{f_j}\right]$$

$$\text{Total Loss} = \text{Cross-Entropy Loss} + \lambda_1 L_1(W) + \lambda_2 L_2(W)$$

- 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

Xu et al. (2014)	Baseline CNN	Human Expert	MobileNet Transfer
46.2%	55.4%	56.0%	**75.7%**

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.

Table 2. Best and worst per-class test F1 scores from MobileNet Transfer model

Class	F1(%)
No Architecture	**98.0**
Ancient Egyptian	**98.0**
Novelty	89.4
Russian Revival	84.2
International	53.3
Bauhaus	52.6
American Foursquare	**28.6**
Overall	**75.3**

(a) American Foursquare (b) American Craftsman

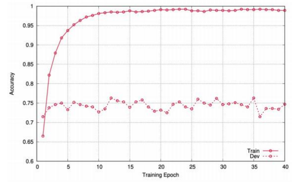
Figure 1. Misclassification

(a) Original
(b) Door (c) 'House' (d) Window

Figure 2. Fine-grained classification

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

Xu, Z., Tao, D., Zhang, Y., Wu, J., and Tsoi, A. C. Architectural style classification using multinomial latent logistic regression. In European Conference on Computer Vision, pp. 600–615. Springer, 2014.