Deep Architecture

Richard Aw
Mitch Veele
Wolfgang Herwald

Outline

1. Data Preprocessing
2. Data Augmentation
3. Handling Class Imbalance
4. Modeling
5. Model Performance and Results

Objective

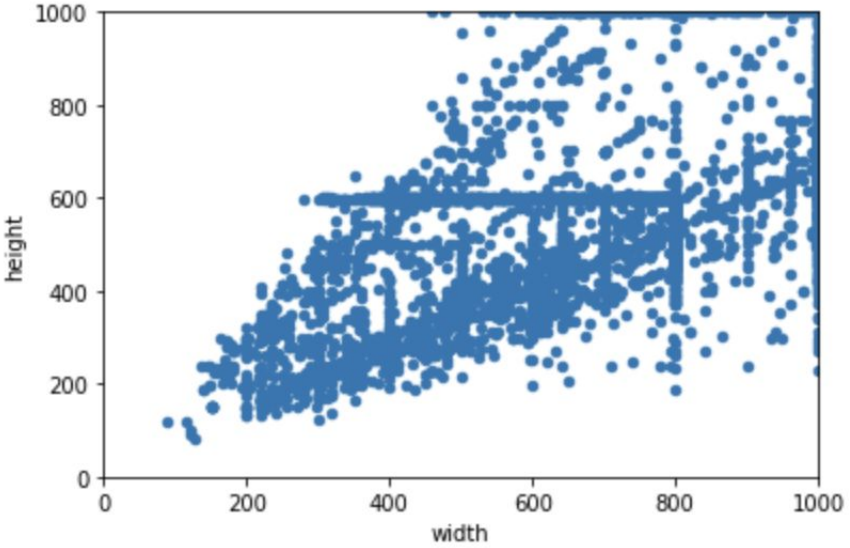
Given an image of a building, predict its **architectural style**

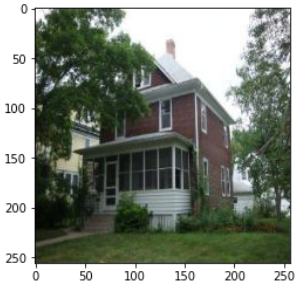
Data

Kaggle Architectural Styles: 10,113 Images of buildings. 25 class labels for style

Preprocessing

- Only true preprocessing was to scale all of the images to the same dimensions (256 x 256)
- 600 x 800 is most common and around the 50th percentile





Training Issues

Class Imbalance

- class_weights = tensor([1.2831, 1.2880, 0.8039, 1.1788, 0.7397])
- nn.CrossEntropyLoss(weight=class_weights)

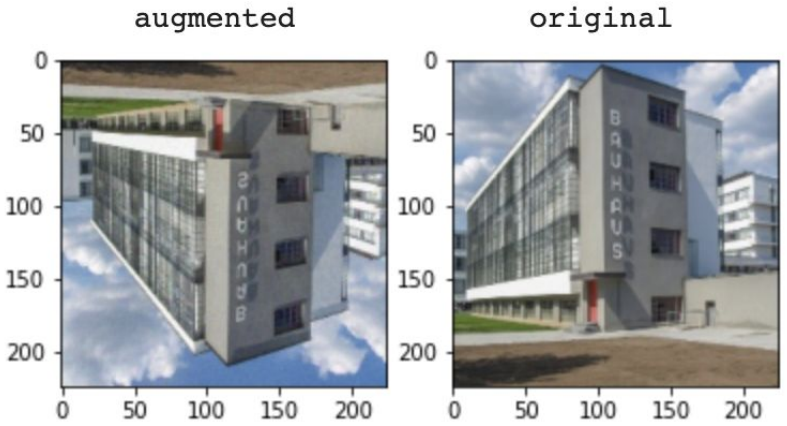
Number of Classes

- Started with full 25 classes but training took a long time since there's more data
- After a 5 epochs accuracy was ~ 50%

	class_index	class_label	N	Proportion
0	0	Gothic architecture	331	0.159
1	1	Byzantine architecture	313	0.150
2	2	Greek Revival architecture	523	0.251
3	3	Russian Revival architecture	352	0.169
4	4	Art Deco architecture	566	0.271

Data Augmentation

- RandomCrop (CenterCrop for validation)
  - Standard ResNet dims (224, 224)
- Flips - Vertical / Horizontal
- GaussNoise
- Normalize() (train and valid)
  - Standard ImageNet Normalization



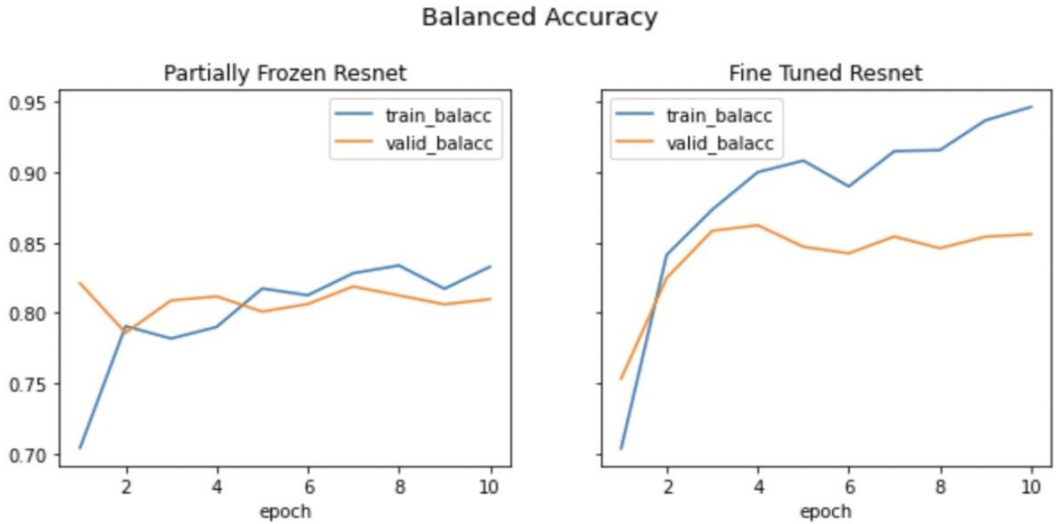
Models

- No “from-scratch” approach
  - Time-Consuming
  - Unlikely to perform better than fine-tuned approach

- ResNet-18 CNN
  - Pretrained on ImageNet
  - Tried partially frozen - just updating the last layer
  - Fine-tuning - update all parameters.

Performance - Accuracy

Accuracy vs epoch count



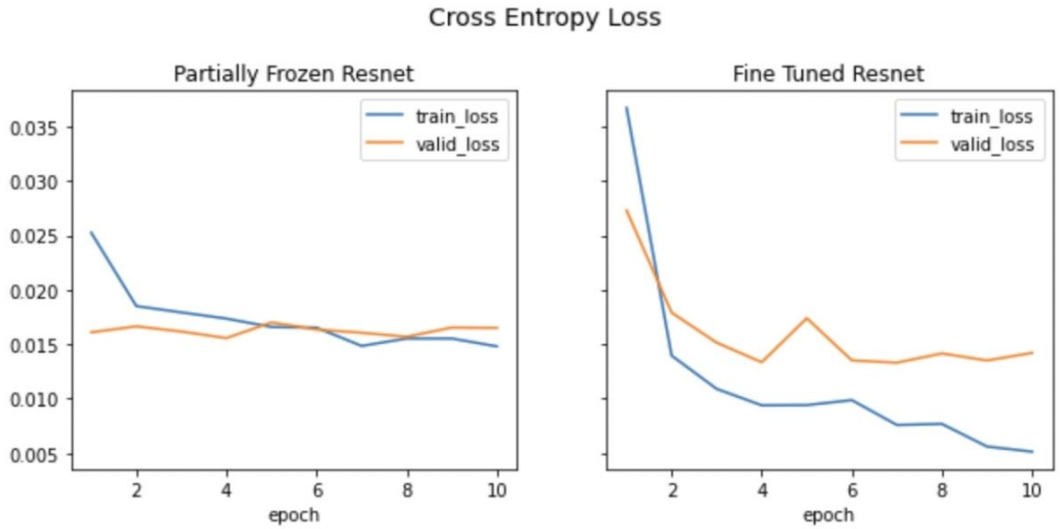
Partially frozen - 82%

Fine Tune - 86%

Excellent accuracy for small scale multi-class classification problem. Showing potential for larger scale problem.

Performance - Cross Entropy

Training Loss vs epoch count



Partially frozen - 0.016

Fine Tune - 0.013

Excellent loss for small scale multi-class classification problem. Showing potential for larger scale problem.

Conclusion

- Best model has 86% validation accuracy on a subset of the classes
- With more time and more complicated architectures, we could likely achieve better results on the full 25 classes
- Experimentation
  - More data
  - Augmentations
  - Image sizing/scaling
  - Different Architectures

Thanks for listening!