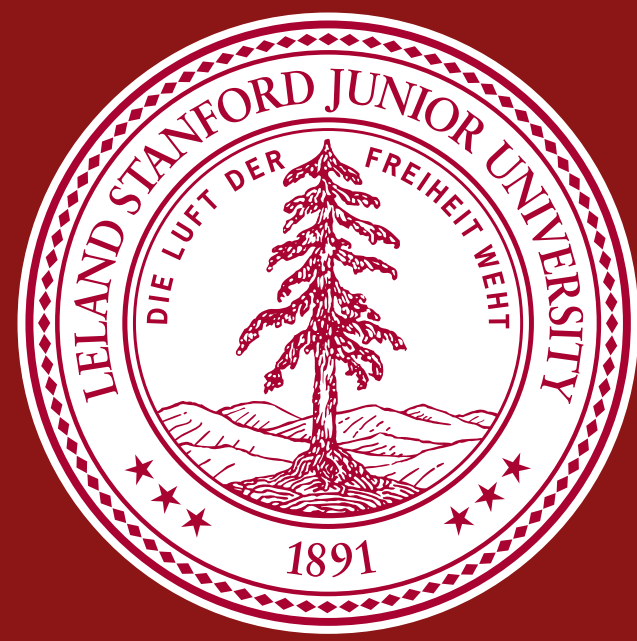




Strava Art Sketch Image Classification

Anthony Xie, Dylan Cunningham

Department of Computer Science
Stanford University

Stanford
Computer Science

Background

- 2.5 billion activities uploaded to Strava in past 18 months
- Strava art is a way for athletes to share creativity
- Sketch recognition can be viewed as a measure of general understanding of object recognition**
- Sketches are sparse, abstract, and lack color

Problem Statement

- Take 4 pre-trained ImageNet models and finetune them to perform well on our Strava Art dataset
- Input Strava art to the model and get ImageNet class output

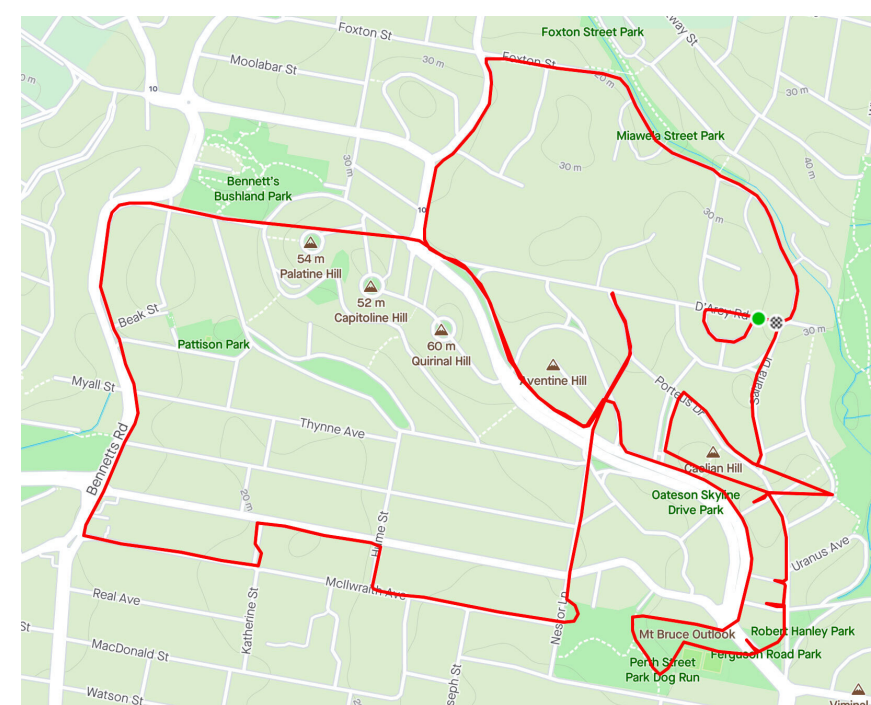


Figure 1: Elephant Strava art

Methods

- Fine tune 4 different types of models using ImageNet Sketch**

ResNet50

- Summation of previous layer to the next layer, allowing for reuse of learned weights
- Highly explored for transfer learning

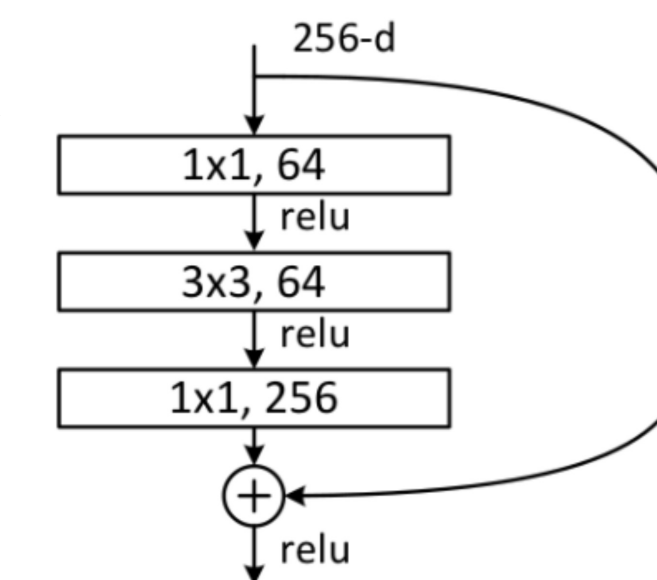


Figure 4: Residual Block¹

EfficientNet

- Use of compound scaling to find optimal width, depth, resolution for neural networks
- Higher accuracy w/ equal computing needs as ResNet50

Inception-ResNet

- Use of inception modules – reducing computation complexity by using multiple sized filters and concatenating
- Replacing the filter concatenation stage of standard inception modules

IG-ResNeXt

- ResNext layers – aggregate residual transformations, with a shared topology for all of its paths, similar to ResNet
- Use of cardinality – the measure of the size of the set of transformations
- Pre-Trained on both large Instagram dataset and ImageNet where the other models are just pretrained on ImageNet**

Results + Analysis

Model	No Fine-Tuning		With Fine-Tuning	
	Top-1 %	Top-5 %	Top-1 %	Top-5 %
ResNet50	0.00	1.59	5.55	16.66
tf-efficientnet-b5	0.79	3.18	7.93	18.25
Inception-ResNet-v2	1.59	3.97	7.14	19.84
ig-resnext101-32x8d	26.98	52.38	28.57	46.03
ig-resnext101-32x16d	28.57	54.76	23.01	34.92

Table 1: Model scores on Strava art test set

- Fine-tuning yielded increase in top-5 accuracy for 3/5 models, and increase in top-1 for 4/5 models**

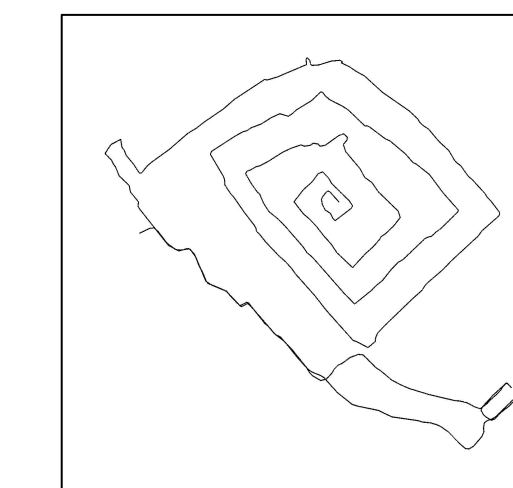


Figure 5. Common correctly labelled class, "snail"

- IG-ResNeXt models consistently predicted within the proper domain, and correctly labelled easily identifiable, unique classes (ex: snail)

- Most models inaccurately labelled abstract images and classes with visually similar classes (ex: labelling 'tricycle' for 'bike')

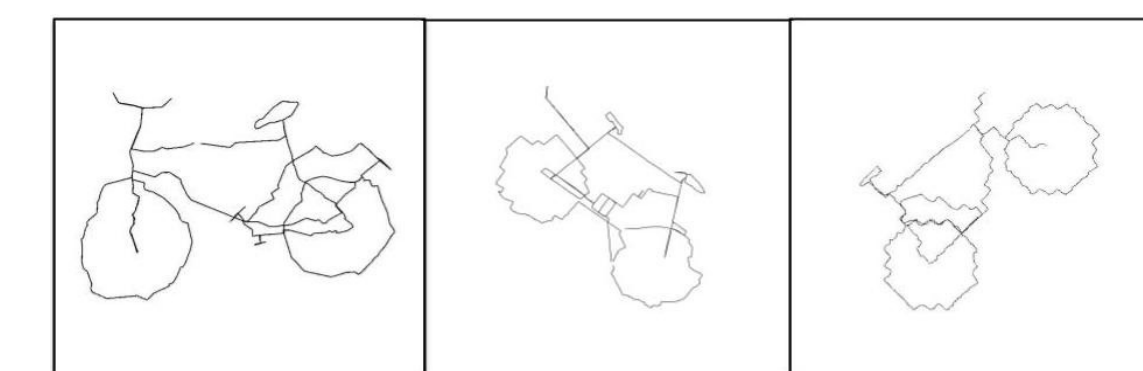


Figure 6. Common incorrectly labelled class, "bicycle"

- IG-ResNeXt models showed recognition intelligence by tending to predict animal classes for animals and object classes for objects, even if incorrect**

Conclusions

- Fine-tuning on ImageNet Sketch worked**, but we need more Strava training data and experiments with custom training routines and new hyperparameters to improve performance

Datasets

1

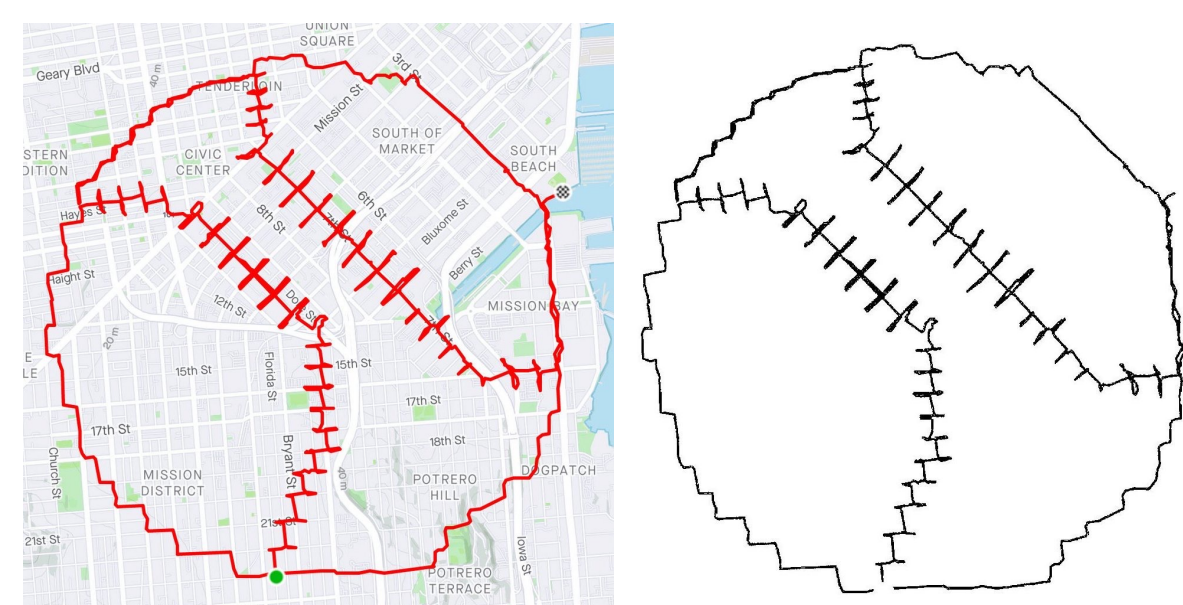


Figure 2: Strava dataset preprocessing

- 126 strava art images pulled from Strav.art
- Selected based on ImageNet classes
- Masked background and padded image
- Used as test dataset

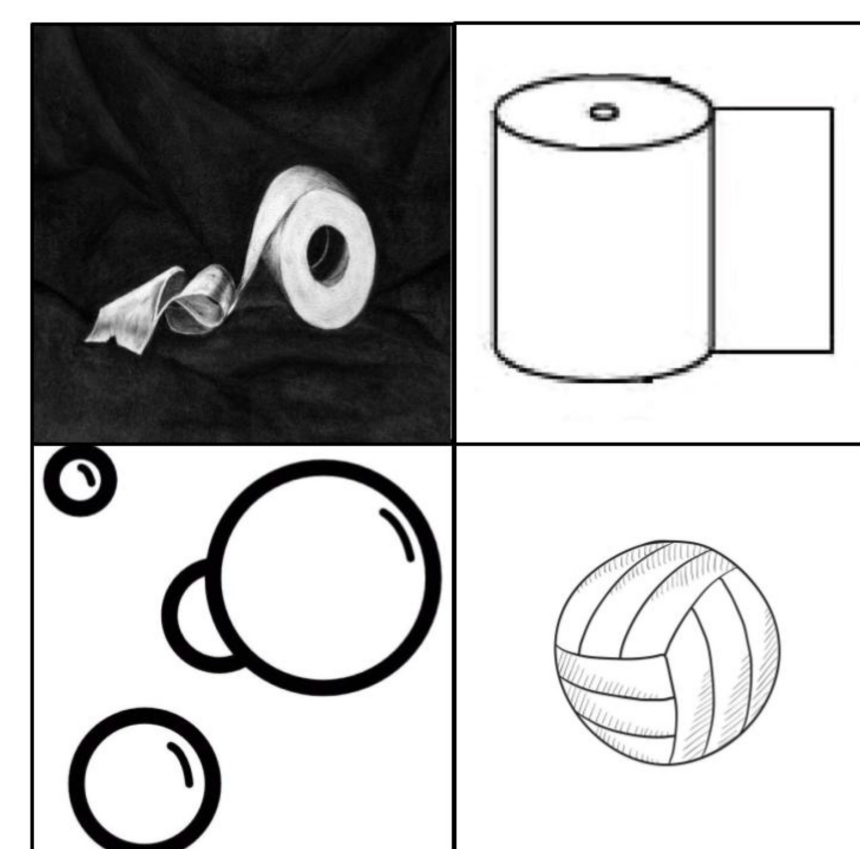


Figure 3: ImageNet Sketch

- 50 images for each of the 1000 ImageNet classes
- Pulled from Google Images
- Used as training and validation

1. Figure from Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition, 2015.