

# Proofpoint

## Logo Detection and Classification Models

11th February 2021

Duke

# Agenda

- Logistics
- Logo Extraction Results
- Pipeline Streamlining Strategies
- Logo Classification Model Update

# Faster RCNN Extraction

Alphabet

Investors



is for Google

As Sergey and I wrote in the original founders letter 11 years ago, "Google is not a conventional company. We do not intend to become one." [more](#)

*Larry Page*

Larry Page

Train Set Image



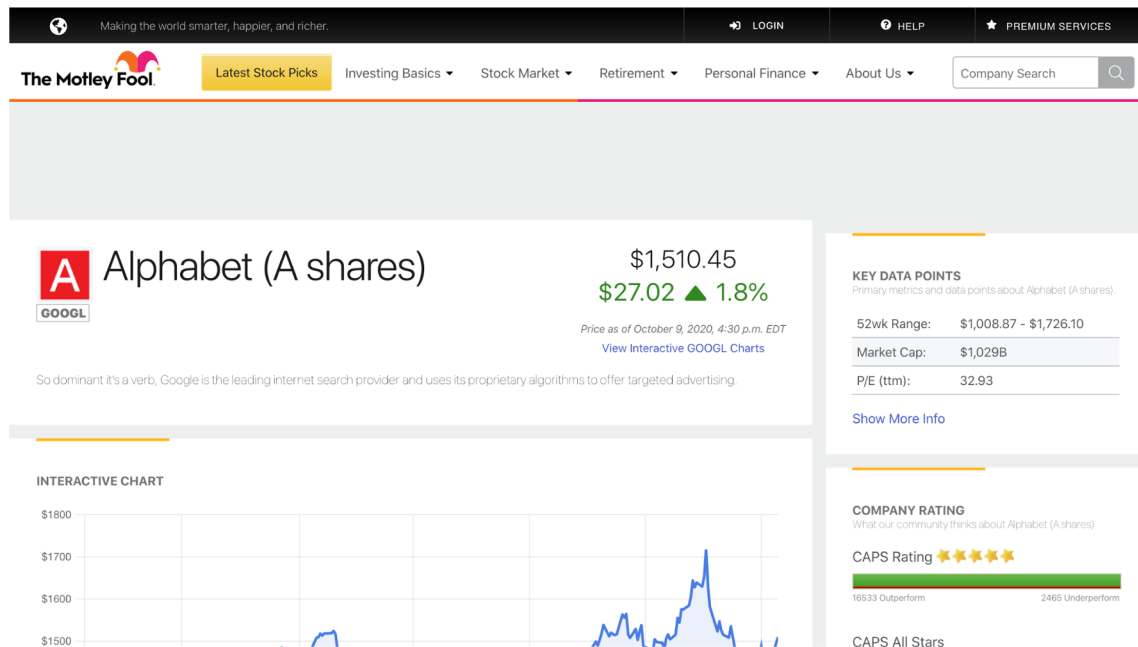
Alphabet



Our Annotations

Duke

# Faster RCNN Extraction



Test Set Image

The Motley Fool



Model Predicted Annotations

Duke

# A quick note....

- Decimal Prediction [34.267, 25.362, 894.213, 15.736]
- I had to round it up
- Cannot crop images with decimals

# YOLOv3-SPP Extraction



Test Set Image



Model Predicted Annotations

Duke

# Issues with Logo Extraction

- Time and space cost of image saving from object detection phase.
- Time cost of image loading and transformations in classification phase.
- More challenging image-label tracking.
- For YOLO, saving images takes more than double the time of inference.

# Pipeline Streamlining / Unifying

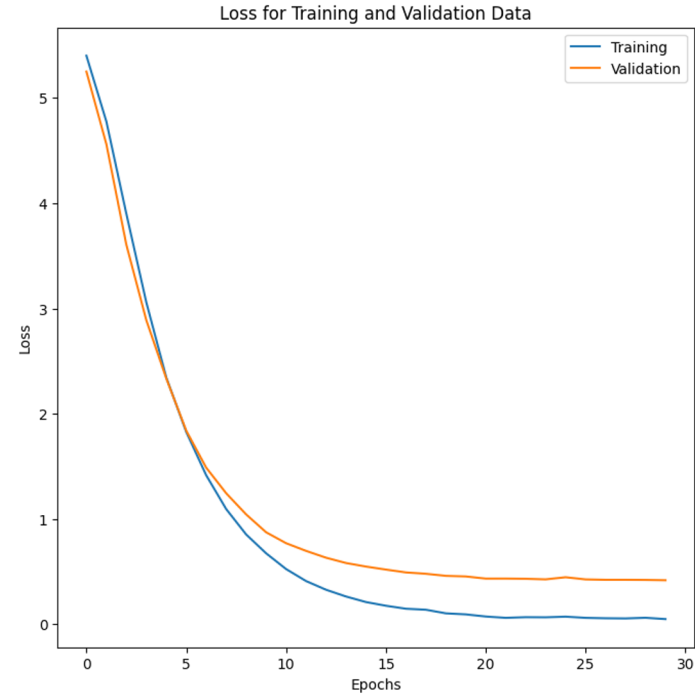
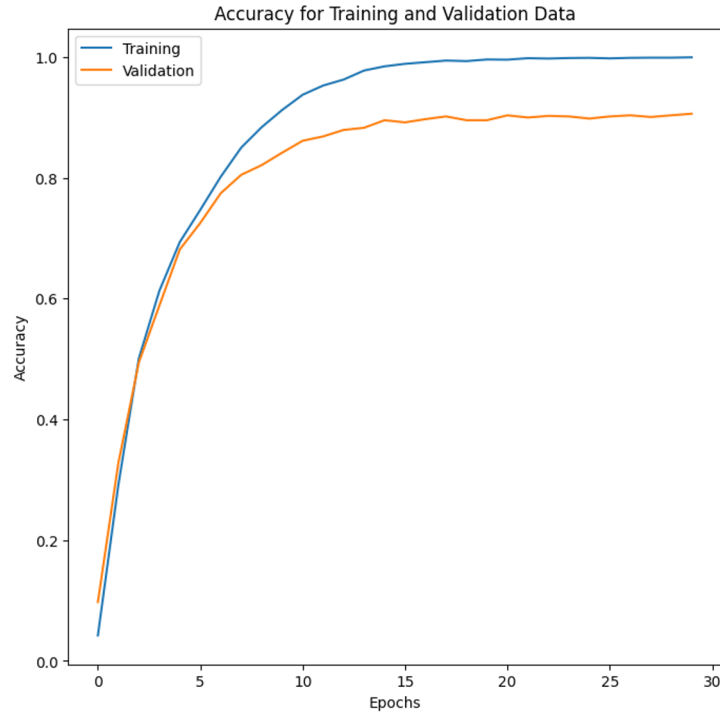
- All steps contained within the same python file - one command line argument.
- Faster pipeline with no unnecessary data reading and saving.
- Easier tracking of images and labels.
- No space cost.



# Logo Classification Model Updates

- Cleaning the dataset of misrepresentative or non-logo images
  - Before: 3822 training, 1273 test (261 classes)
  - After: 3338 training, 1117 test (251 classes)
- Types of images removed
  - Non-logos
  - Two or more logos of different brands in the same image
- Classes removed
  - Contained almost no representative logos of the brand
  - Duplicates (PwC and PricewaterhouseCoopers)
- Retrained model with almost same settings as before
  - 99.6% on training and 90.6% on validation

# Updated Loss and Accuracy Curves



# Notes on Updated Model

- Classes are imbalanced
  - Twitter, Google, Facebook have many sample images
  - Smaller companies may only have a handful
  - May be representative of few-shot/single-shot
- Will be tested on extracted images
- Will continue to be tweaked (different backbones, hyperparameters)

# Sample Removed Logos



secured by  
**sage**pay



**MasterCard**. Verified by  
**SecureCode**. **VISA**

