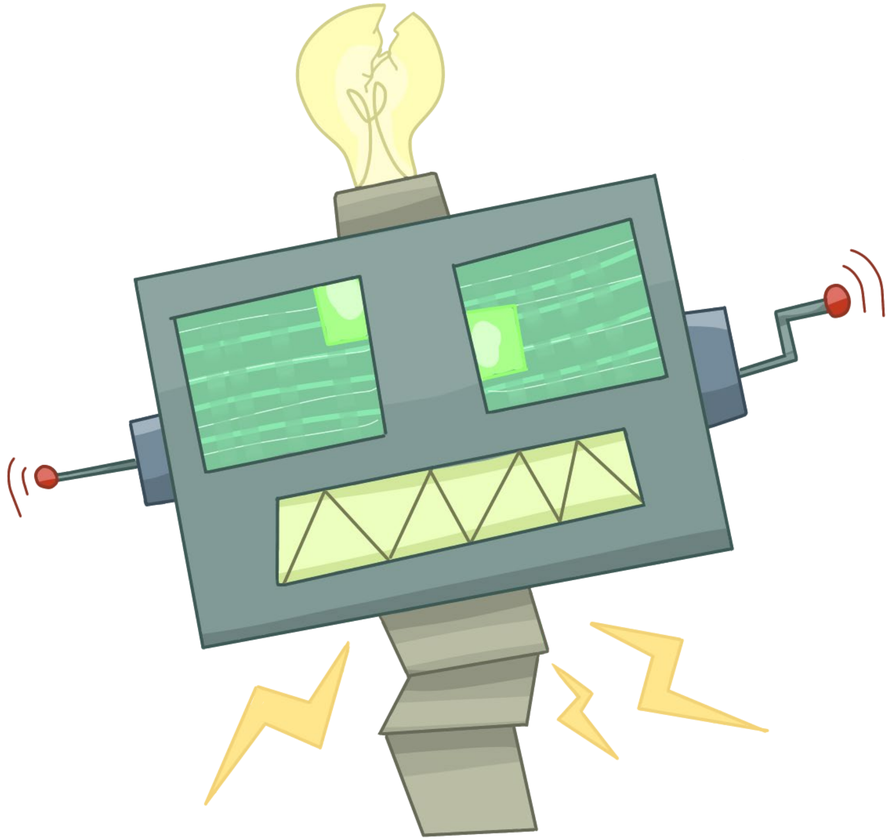


Show & Fail

An attempt at metadata enhancement through machine learning

Todd Peters
Jason Long

TCDL 2022 DIGITAL JUGGLING

TEXAS STATE UNIVERSITY

MEMBER THE TEXAS STATE UNIVERSITY SYSTEM

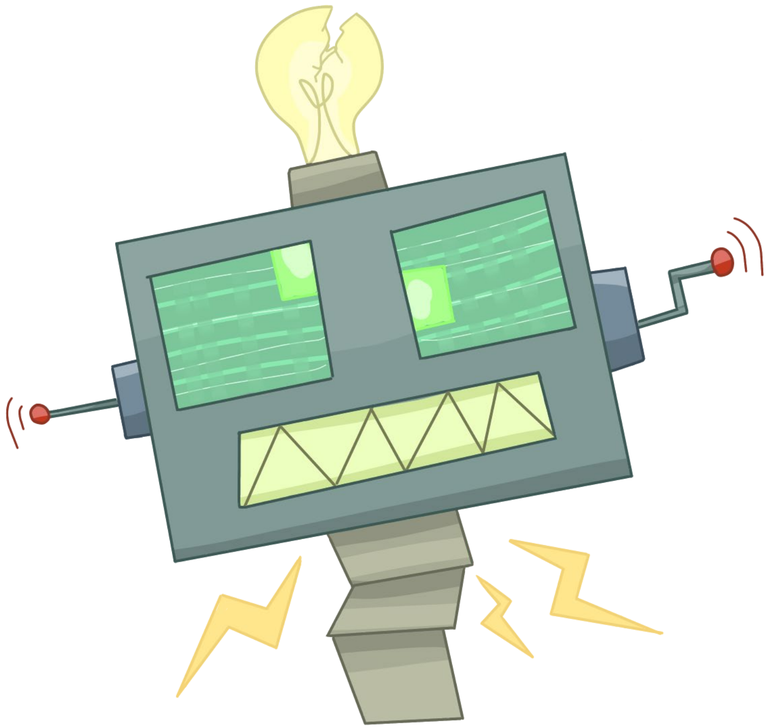
Background

- Inspired by a presentation at the 2019 Texas Conference on Digital Libraries by Patrice-Andre Prud'homme, Oklahoma State University

- Recent machine learning projects have created software and models intended to be more accurate and more easily used. The use of some of these tools does not require a deep understanding of artificial intelligence and advanced mathematics and puts them within reach of libraries without access to computer scientists.

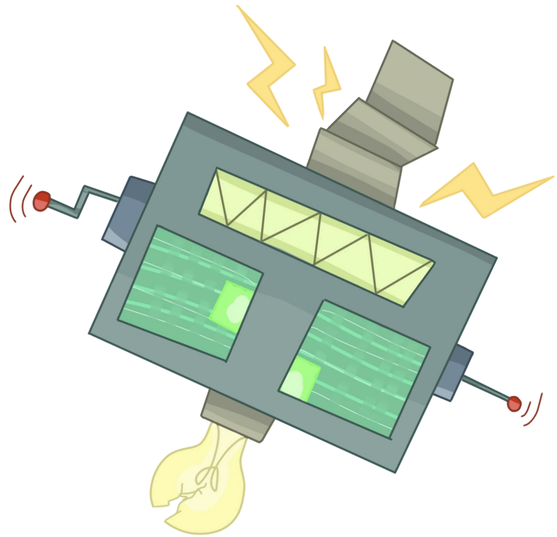
Question

With our limited knowledge, can we easily use neural networks on pre-trained models to classify objects in images and extract metadata that could be used for discovery, without extensive clean-up?



Answer

No



Steep Learning Curve – Hardware& Software

- Video Card
  - Compute Capability, minimum 3.0
  - https://developer.nvidia.com/cuda-gpus

- Compatibility between software components

  - NVIDIA CUDA toolit ver 9.0 or 10.1?
  - Visual C++ build tools ver 14 ?
  - Tensorflow version 1.14 or 1.15 or 2.0?

Steep Learning Curve – Trial and Error

Developed with Python in Jupyter Notebooks – can create different environments with different software versions

First experiments

Python

- Keras Retinanet – opensource Python code for machine learning
  - https://github.com/fizyr/keras-retinanet

Pre-trained Models – a neural network that was trained on a large dataset with powerful hardware

- Coco
- Inception
- ResNet
- Vgg19
- Xception

What is COCO?

COCO is a large-scale object detection, segmentation, and captioning dataset. COCO has several features:

- Object segmentation
- Recognition in context
- Superpixel stuff segmentation
- 330K images (>200K labeled)
- 1.5 million object instances
- 80 object categories
- 91 stuff categories
- 5 captions per image
- 250,000 people with keypoints

https://cocodataset.org/

Pretrained models are trained to recognize a limited number of categories, 80 in this case

Problems

Inaccuracy – error rate is too high to be used without extensive editing

Categories – categories with pre-trained models are too general

Modern skew - categories with pre-trained models trained with modern objects and color images

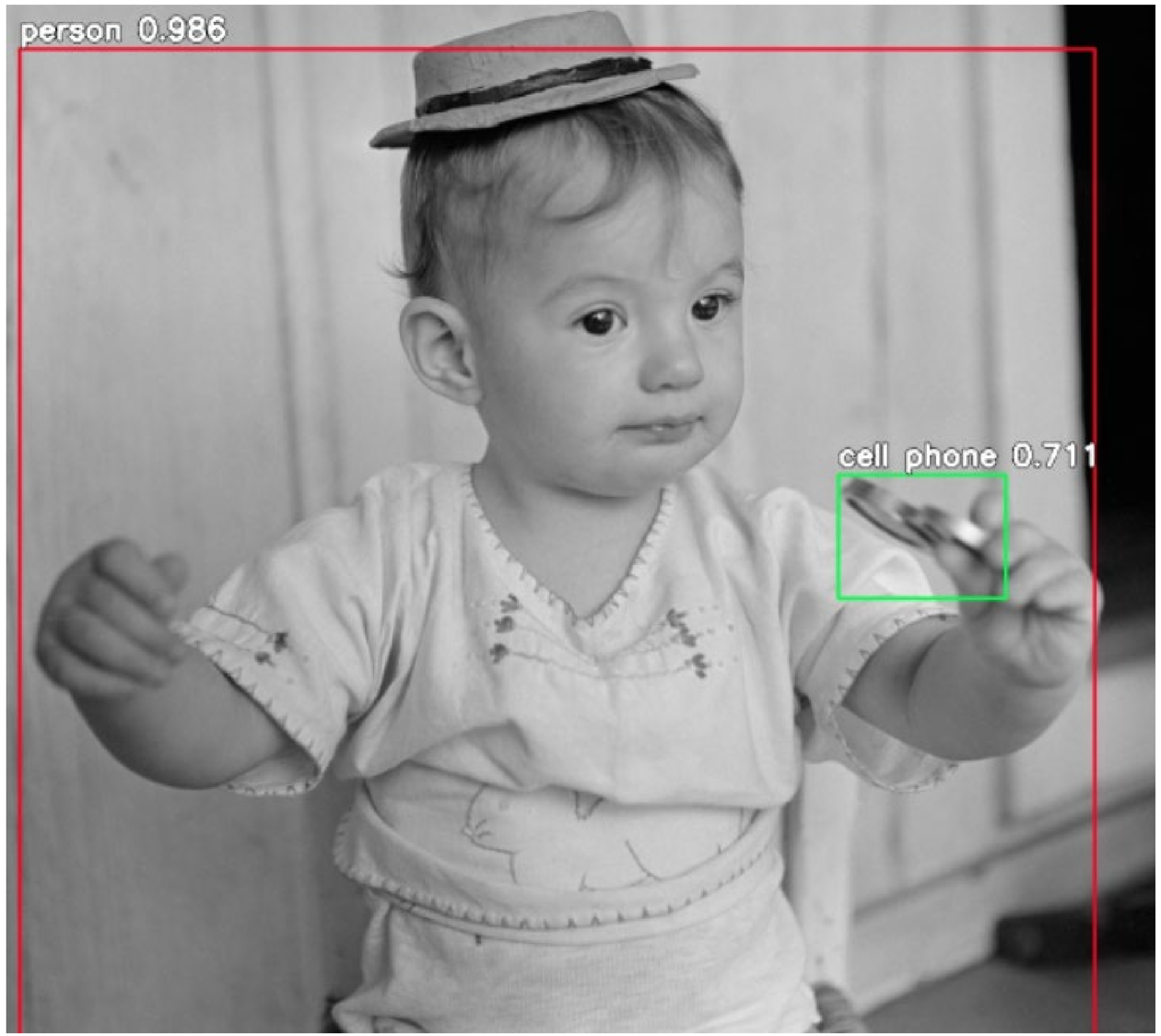
Classification was accurate for a number of images

Overall, metadata generated was not accurate enough to be directly usable

Metadata categories were

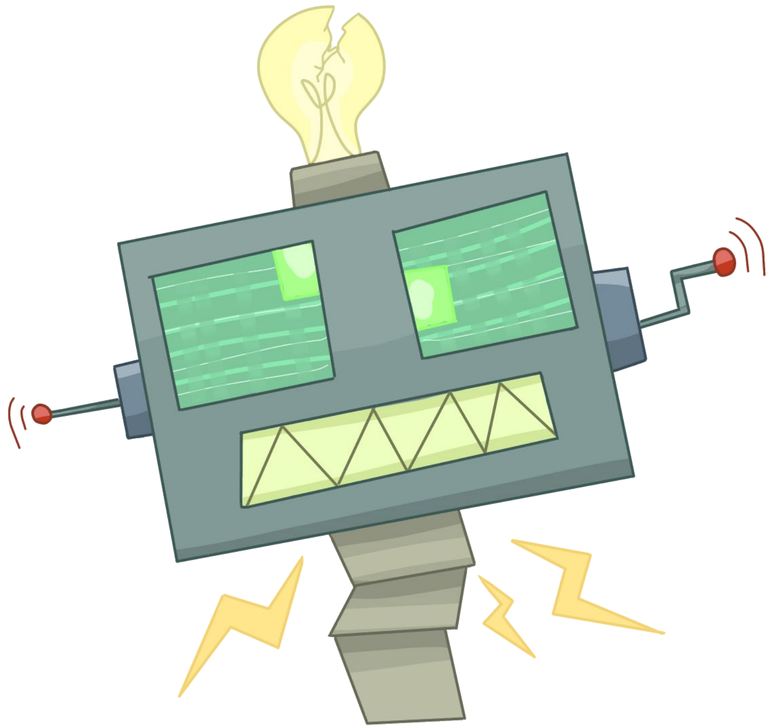
often too general to be useful

Metadata generated from modern categories applied to vintage images



Conclusions

We learned a lot about machine learning which will allow us to identify projects that can benefit from neural networks and to speak intelligently with faculty in Computer Science and Engineering that could be partners.



Future Directions

- Recruit/Partner with Grad students
- Transfer Learning – train existing models to classify new categories, such as local buildings or people
- Cloud computing - **AWS(Amazon Web Services), Azure (Microsoft), Google Cloud, and IBM Cloud**