2021 embedded vision summit

VIRTUAL | MAY 25-27

Tools and Strategies for Quickly Building Effective Image Datasets

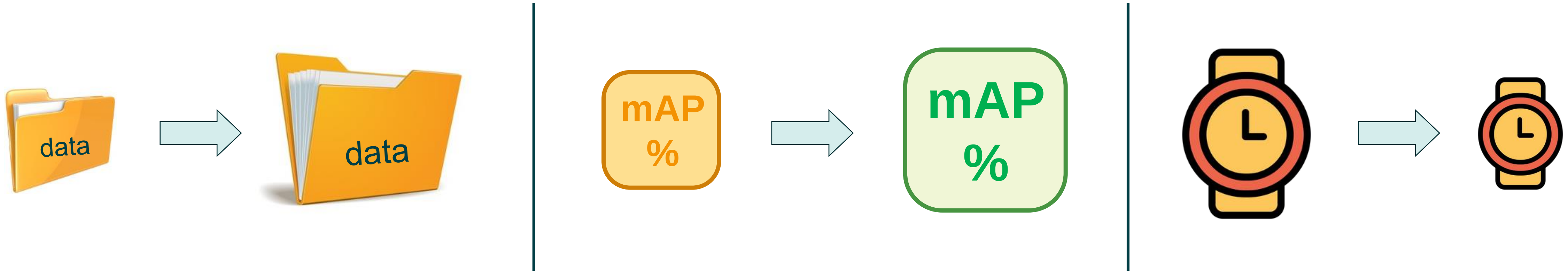
Evan Juras

Computer Vision Engineer

EJ TECHNOLOGY CONSULTANTS

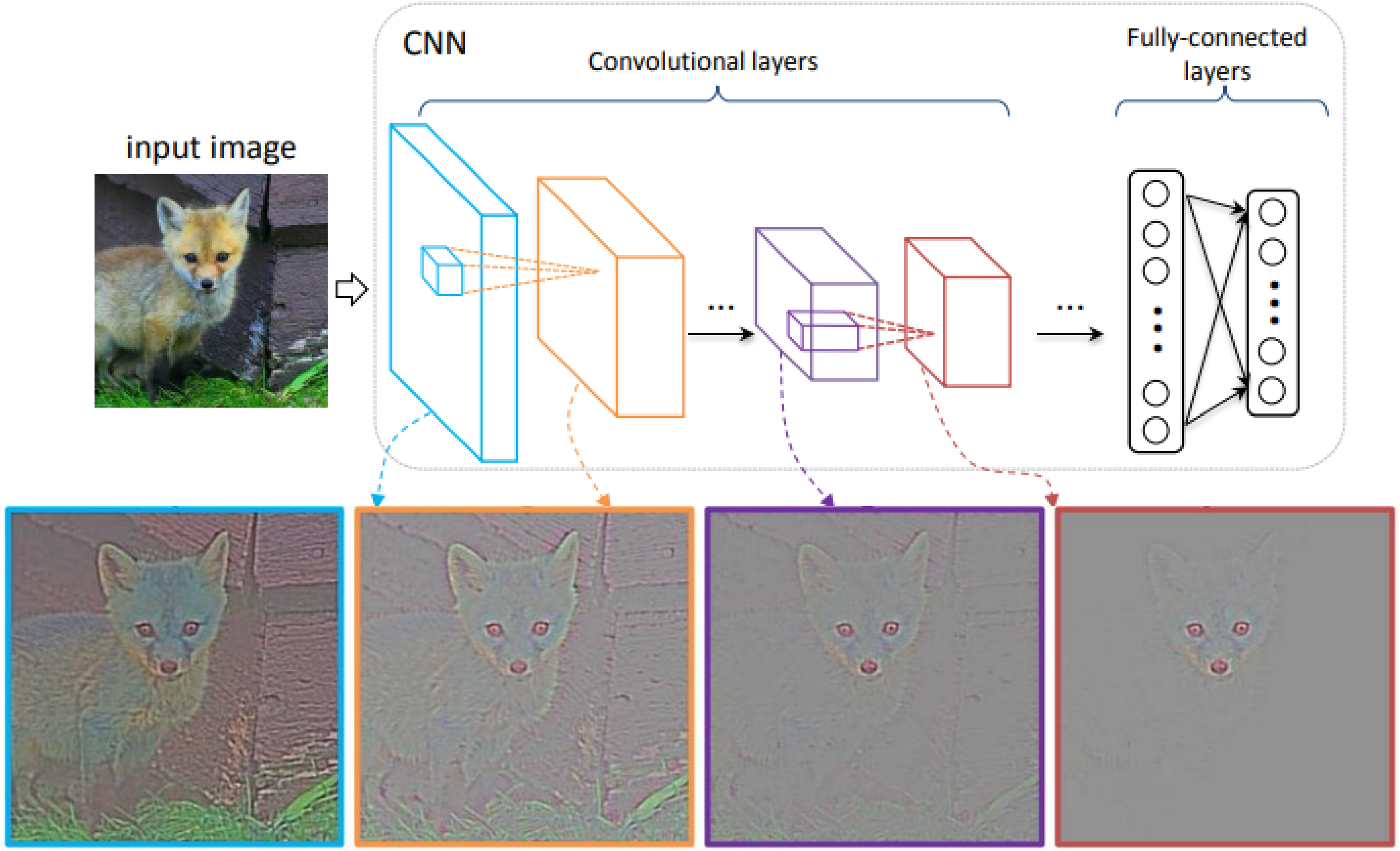
Purpose

- Demonstrate practical strategies and tools for quickly building image datasets
- Train higher-quality models with less time investment
- Focus on object detection models



Object Detection Neural Networks

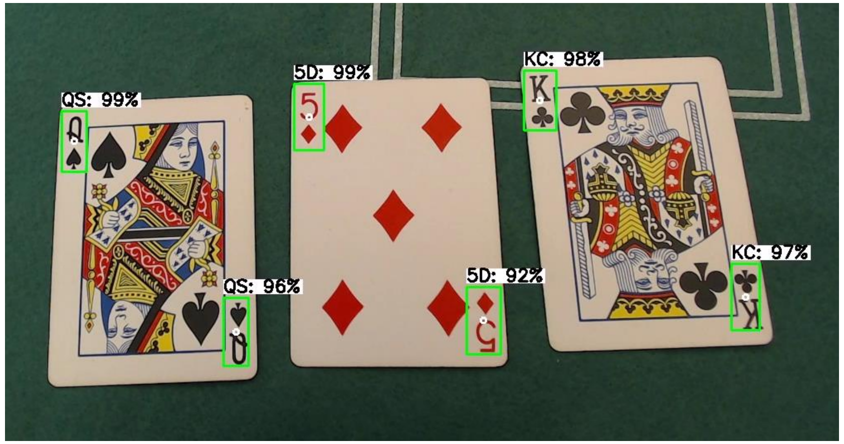
- Object detection models locate and identify objects in images
- Use CNNs to find features in images and correlate them to known objects they've been trained on
- For this presentation, "training" is transfer learning and fine-tuning



Visualizing and Comparing Convolutional Neural Networks
arXiv: 1412.6631v2

Applying Object Detection Models

- **Scope of application determines amount of training images needed**
- **Constrained applications with limited variety of visual conditions:**
  - Lower generalization of model is okay
  - Fewer training images needed
- **Applications with wide variety of visual conditions:**
  - Need to be accurate at different light levels, angles, coloration, distance from camera, etc.
  - Larger number of images are needed



Playing card detector – low variance

Self-driving car – high variance

How can we Quickly Create an Object Detection Dataset?

- **Problem: to train a new object detection model, a large image dataset is needed**
  - Manually gathering and labeling images is time consuming

- **Solution: Use these strategies to accelerate dataset creation!**
  - Using datasets already available online
  - Capturing images from video
  - Using sleek annotation and automated methods tools to quickly label data
  - Synthetic image generation and data augmentation

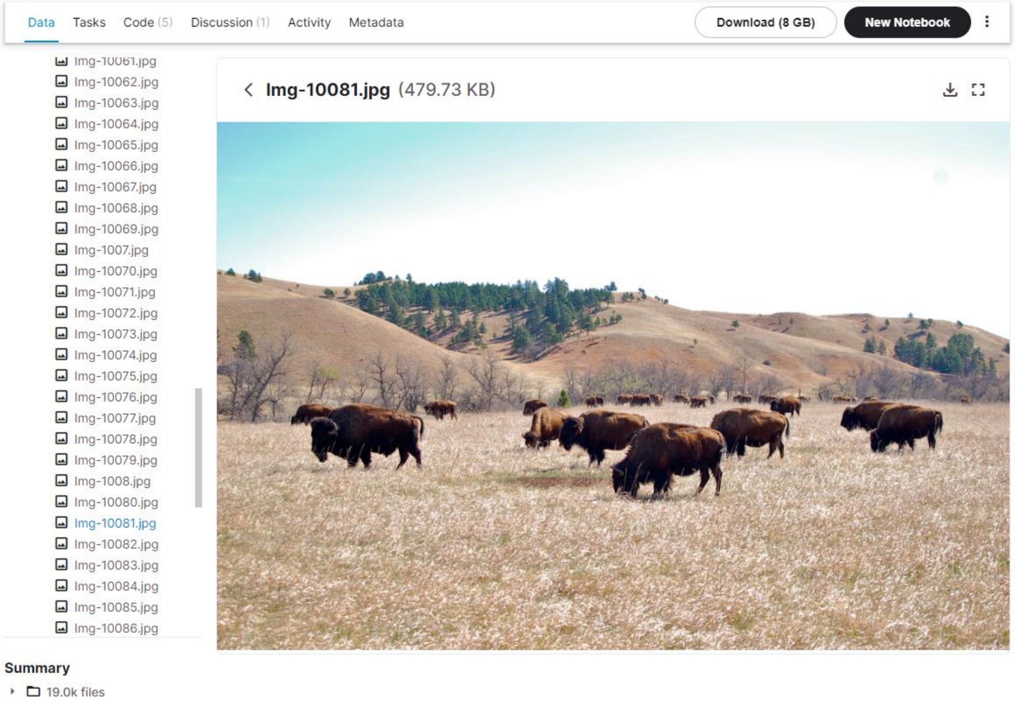
NO AUDIO
NBC NEWS NOW
CLOIE MUSUMEC
WOMAN PLAYS DEAD AFTER TRIPPING
RUNNING AWAY FROM BISON AT YELLOWSTONE

My example – Bison distance detector

Using Online Image Datasets

- **Datasets available online can be a good starting point**
  - Academic datasets such as Open Images Dataset, ImageNet, COCO, etc.
  - User-contributed datasets from TensorFlow, Kaggle, or other sources
- **Issues with online datasets**
  - Can be filled with poor quality images or label data
  - Datasets not be available or appropriate for your application
  - Copyright and licensing: need to be careful using copyrighted material for commercial purposes

kaggle



Img-10061.jpg
Img-10062.jpg
Img-10063.jpg
Img-10064.jpg
Img-10065.jpg
Img-10066.jpg
Img-10067.jpg
Img-10068.jpg
Img-10069.jpg
Img-1007.jpg
Img-10070.jpg
Img-10071.jpg
Img-10072.jpg
Img-10073.jpg
Img-10074.jpg
Img-10075.jpg
Img-10076.jpg
Img-10077.jpg
Img-10078.jpg
Img-10079.jpg
Img-1008.jpg
Img-10080.jpg
Img-10081.jpg
Img-10082.jpg
Img-10083.jpg
Img-10084.jpg
Img-10085.jpg
Img-10086.jpg

Img-10081.jpg (479.73 KB)

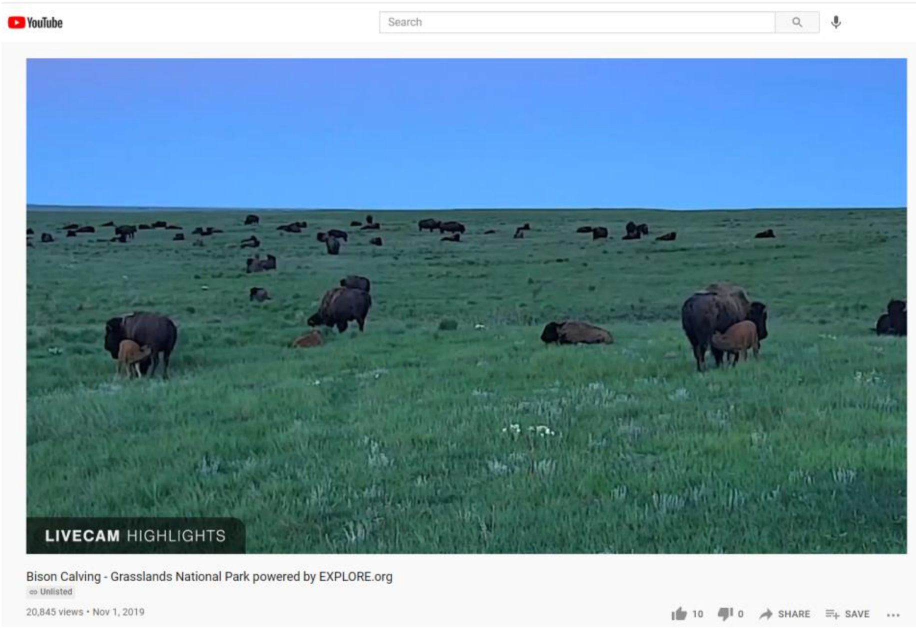
Summary

19.0k files

Capturing Images from Video

- **Gathering images from the camera used by the application significantly improves accuracy**
- **Various methods for capturing images specific to your application:**
  - Set up cameras in situations similar to actual in-situ application, record video, and grab frames from video
  - Use multiple cameras in multiple locations and aggregate videos to a central server
  - Use online video or live camera streams of objects you are interested in

Capturing Images from Video, Continued



Bison Calving - Grasslands National Park powered by EXPLORE.org

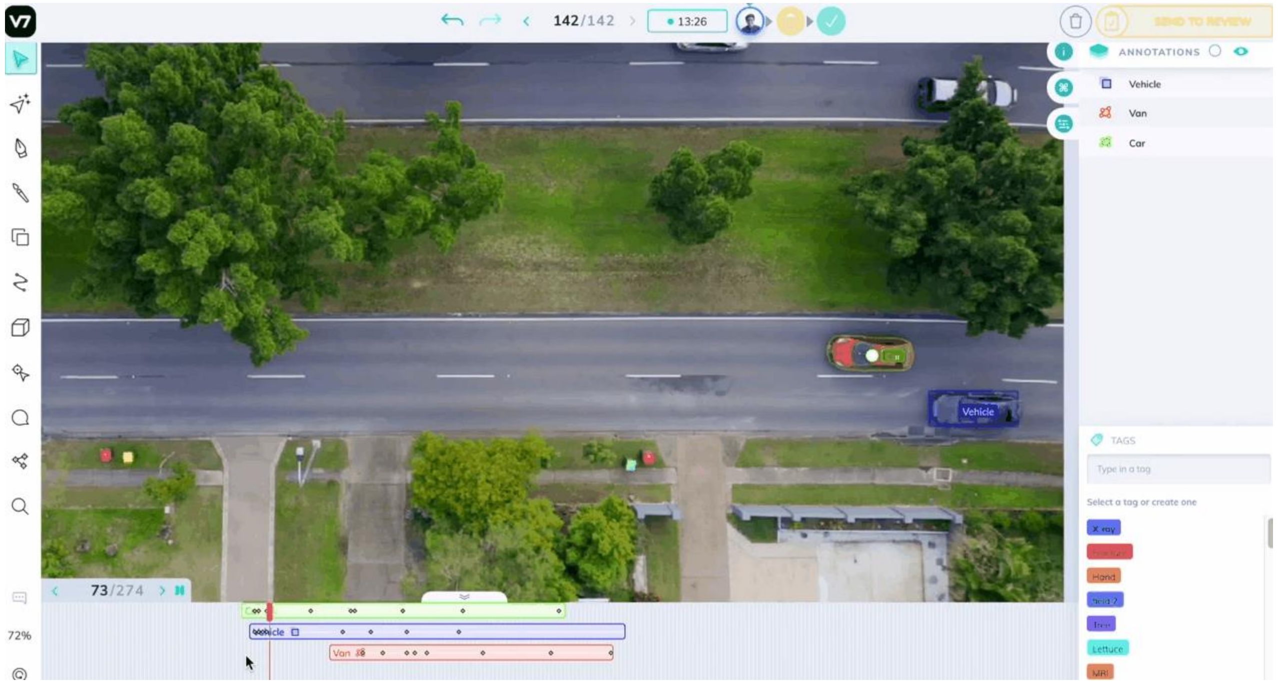
20,845 views • Nov 1, 2019

Python script to extract individual frames from a video:
github.com/EdjeElectronics/Image-Dataset-Tools#FrameGrabber

EJ TECHNOLOGY CONSULTANTS

Image Annotation Tools

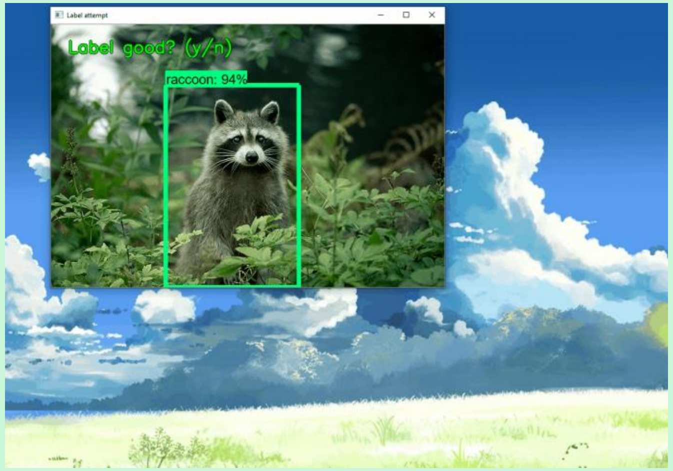
- **Many annotation tools are available to help accelerate the image labeling process**
- **Paid annotation tools:**
  - V7 Darwin
  - Supervisely
  - Hive – pay humans to annotate your data
  - Lionbridge – pay humans to annotate your data
- **Free annotation tools:**
  - CVAT
  - LabelImg



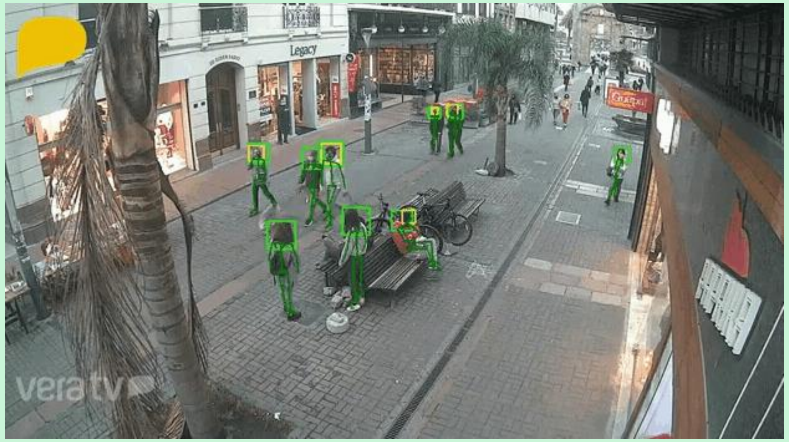
Video annotation tool from V7

Automated Labeling Methods

- **Supervised automated labeling**
  - Train a model off a small portion of the dataset
  - Use that model to automatically label the rest of the dataset, while manually rejecting poor labels and re-labeling them
  - github.com/EdjeElectronics/Image-Dataset-Tools#AutoLabeler
- **Other creative methods**
  - Tryolabs example uses OpenPose to automatically locate positions of heads in a frame



Automatically labeling raccoon images



Face mask detection in street camera video streams using AI: behind the curtain

(Courtesy of Tryolabs)

Synthetic Image Generation and Data Augmentation

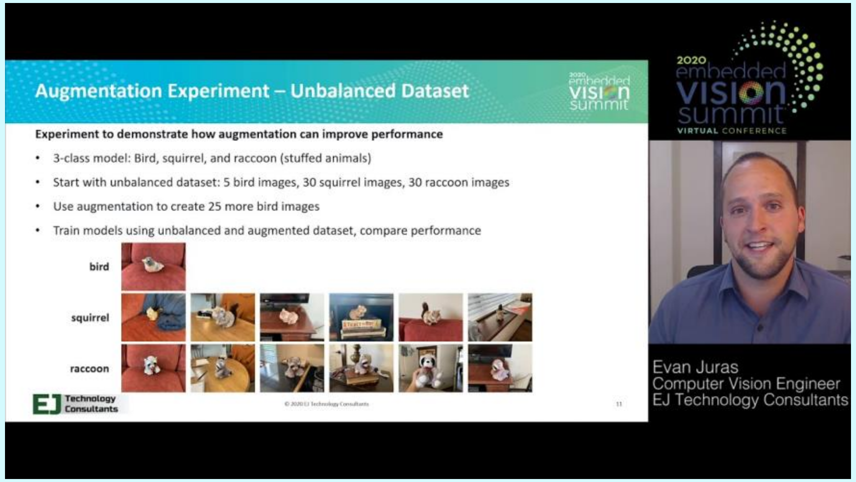
Synthetic Image Generation

- **Tools are available to synthetically generate images**
  - Mindtech
  - KineticVision

Synthetic scenario video from Mindtech

Data Augmentation

- **Generate new images from existing ones**
  - Increase visual variety of images
  - Resolve class imbalances
  - Great for 2D objects and perspectives



Augmentation Experiment – Unbalanced Dataset

Experiment to demonstrate how augmentation can improve performance

- 3-class model: Bird, squirrel, and raccoon (stuffed animals)
- Start with unbalanced dataset: 5 bird images, 30 squirrel images, 30 raccoon images
- Use augmentation to create 25 more bird images
- Train models using unbalanced and augmented dataset, compare performance

bird
squirrel
raccoon
© 2020 EJ Technology Consultants
11
2020 embedded vision summit VIRTUAL CONFERENCE

Evan Juras
Computer Vision Engineer
EJ Technology Consultants

Check out my data augmentation talk from EVS2020!

Limitations of These Methods

- **Still need some manual groundwork involved in collecting images**
  - Searching for datasets, setting up video recording, auditing labels for accuracy
- **Other limitations**
  - Difficult to use for uncommon objects or applications
  - Won’t bring model up to highest accuracy possible – for best performance, need to carefully curate and refine dataset, which takes time

Conclusions

- **There are various methods for quickly building an object detection dataset**
  - Online datasets
  - Fast image capture
  - Sleek annotation tools and automated labeling
  - Synthetic image generation and data augmentation
- **Depending on application, may still need to manually curate dataset for best accuracy**
- **Same concepts can be applied to image classification models**

Example of Resource Slide

Resource Links

Browser-based free annotation tool (CVAT):
https://cvat.org

Using TensorFlow's built-in datasets:
https://www.tensorflow.org/datasets/overview

Handy scripts for working with image datasets:
https://github.com/EdjeElectronics/Image-Dataset-Tools

Contact Information

Website: www.ejtech.io

Email: evan.juras@ejtech.io

References

[1]: M. Zeiler, R. Fergus (2013). Visualizing and Understanding Convolutional Networks. arXiv:1311.2901

[2]: W. Yu, K. Yang, Y. Bai, H. Yao, Y. Rui (2014). Visualizing and Comparing Convolutional Neural Networks. arXiv:1412.6631v2