

News Image Classification

CS 640 Project 2 Topic 2

By: Maya Shen, Rongyu Wang, Xiankang Wu



Problem Definition

Nowadays journalists are not the first ones to record an emergency event but any civilians with phones, and they send to social media almost immediately when such event takes place in their proximity.

For this project, we build a system that looks for such event in a pool of images originated from social media feeds; the system consists of two parts:

- Deep Neural Network trained for object detection
- Expert System that yields an emergency level score based on the object(s) detected from images.

Background

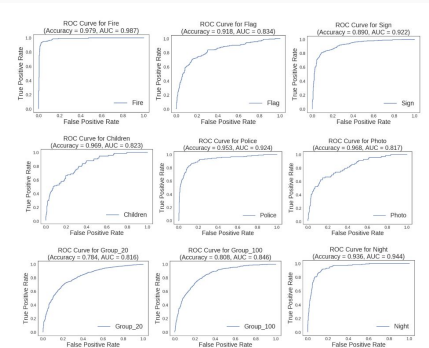
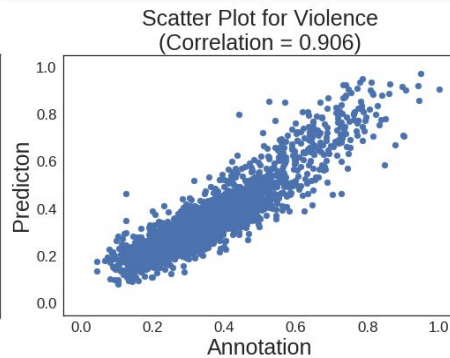
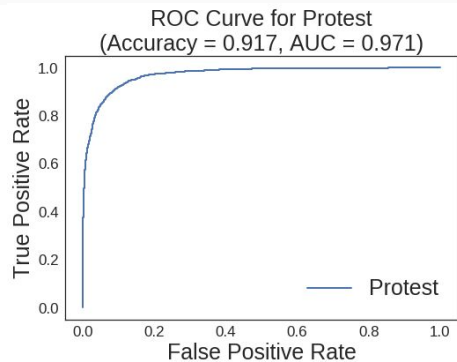
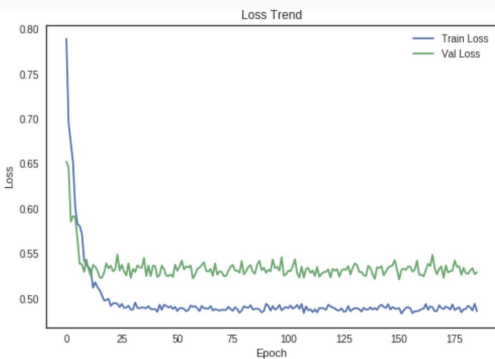
- Understanding issues such protest and violence has been a critical topic within research field crossing different academic disciplines.
- facial attribute classification
- Photograph & Portrait Analysis
- Social media presence examination
- And so on ...



Baseline Approaches

- Original Paper and Method Reproduction
- Transfer Learning on Single Class Classification
 - Fine Tuning
 - Fixed Feature Extractor
- Transfer Learning on Multi-label Detection

Original Paper Reproduction



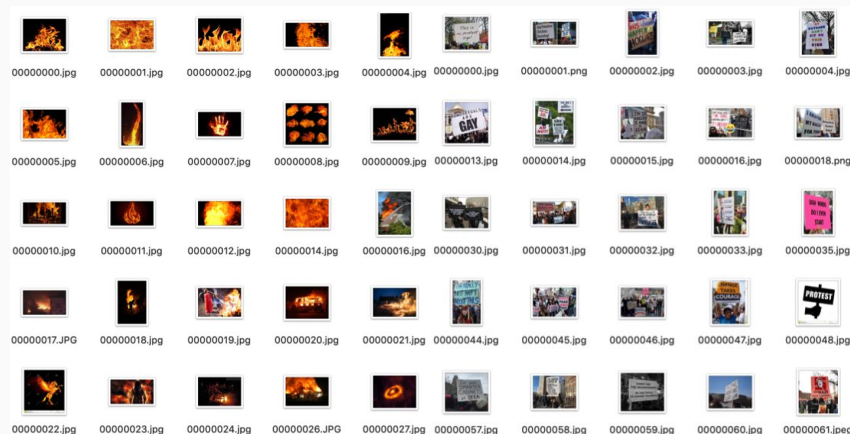
- Loss Trend keeps fluctuating after epoch = 50
- Average Accuracy/AUC/Correlation gets 85% and above.

Transfer Learning on Single Class (Protest)



- Fine Tuning (25 Epochs, best validation Acc: 0.8688)
- Fixed Feature Extractor (best validation Acc: 0.8375)

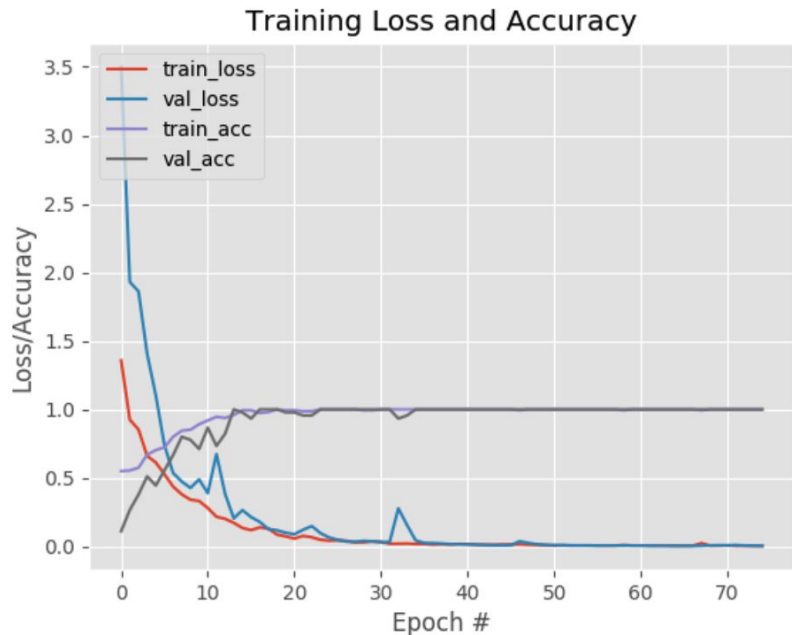
Transfer Learning on Multi-label objects



searching Bing API for 'fire'
250 total results for 'fire'
making request for group 0-50 of 250...
saving images for group 0-50 of 250...
fetching: <https://daily.jstor.org/wp-content/uploads/2015/08/Fire-1050x699.jpg>

- Initial idea: avoid using bounding boxes.
- To train our own classifier, first we need to find good resource of datasets.
- Bing Image Search API is a RESTful service, ideal for fetching images we desired with simple HTTP request

Transfer Learning on Multi-label objects



- With Epoch = 75, Our model performs well for a single object detection.
- Not ideal for multi-label objects detection when several objects are presented.
- Bounding box data is still required for training complicated classifier.

Improvement

- **Before Training:** Dataset Quality
- **During Training:** Parameter Tuning
- **After Training:** Extra Functionalities
 - Simple Expert System
 - Crowd Detection

Dataset: UCLA_PROTEST_IMAGE_DATASET

Fields	Protest	Sign	Photo	Fire	Police	Children	Group>20	Group>100	Flag	Night	Shouting
# of Images	11,659	9,669	428	667	792	347	8,510	2,939	970	987	548
Positive Rate	0.286	0.829	0.037	0.057	0.068	0.030	0.730	0.252	0.083	0.085	0.047

of images: 40,764

of protest images: 11,659

Training set/ Test set = 8:2

Speed and Learning Rate Tuning

- **Speed:** running time comparison for 25 Epochs, CUP took 42m11s to complete while GPU took 53s.
- **Learning Rate:** fine-tuned models should adapt relatively small LR as the network weights are at a good level to start with, a large LR will distort the model too quickly and too much. LR originally sets at 0.001, and was tuned towards smaller end, but results are less than desirable in terms of training speed and cost function.

Simple Expert System

```
import pandas as pd
import sys

# read in input .csv
path = sys.argv[1]
df = pd.read_csv(path+'/result.csv')

# locate violence column entry
var = df.loc[df.violence]

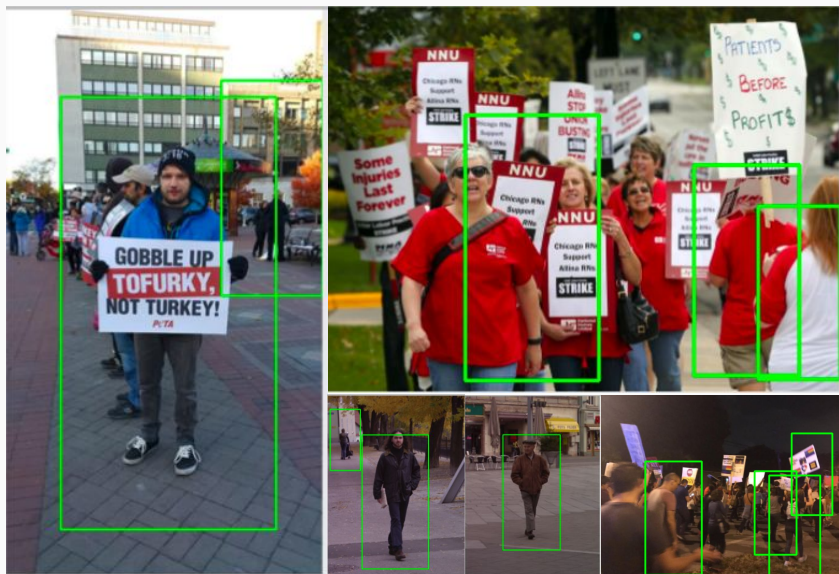
# generate simple responses
if (var < 0) or (var > 1) or (var is None):
    print("Invalid Violence Ratio")
elif var>=0 or var<= 0.3:
    print("Low Emergency Level")
elif var>0.3 or var < 0.7:
    print("Medium Emergency Level")
else:
    print("High Emergency Level")
```

A very simple 'expert system' was build to illustrate our idea of generating an Emergency Level Rating after an image's passed through the Neural Network.

Potential Improvements:

- Involving more input measures, now we're only looking at violence, could include measures for presence of fire, children
- More detailed categorization for emergency level

Crowd Detection



- Potential interest points: crowd count during a protest.
- Explored and used OpenCV built-in HOG descriptor and SVM detector.
- Draw bounding boxes around people detected and counting them.
- Applied Non-maxima suppression to suppress overlapping bounding boxes to achieve the correct final detection.
- Not ideal for dense, overlaid crowd scene.
- Great for Pedestrians when people are fully included.

Thanks!

