UNIVERSAL OPEN-IMAGE PREDICTOR

Group 11

Aadish Joshi | asj170430

Piyush Mahatkar | pkm170230

Idea & Inspiration

- GoogleAI: Inclusive Images Challenge

- The motto: To make products that work for people all over the globe. In the field of classification, this means developing models that work well for regions all over the world.

- A system trained on a dataset that doesn't represent a broad range of localities could perform worse on images drawn from geographic regions underrepresented in the training data.

- For example.

Data Generated Locations

OpenImages Distribution (See Shankar et al., 2017)

Challenge Stage 1 Distribution (Illustrative)

Challenge Stage 2 Distribution (Illustrative)

WHAT IS THIS?

Sponge ?

WHAT IS
THIS?

Audi ?

Dataset

- Data size: ~16GB.

- # validation images: ~40K.

- # test images: ~15K.

- # label classes : 19998

- # human label: 803K[Confidence 1]

- # Machine labels : 15M [Confidence 0]

Starting point

- We knew that images would come from diverse regions. We used

- Data augmentation and ensembling.

- https://arxiv.org/pdf/1506.07224.pdf

- Some approaches for diversity:

- 1) Different CNN architectures

- 2) Different Image sizes

- 3) Train with/without augmentation

- 4) Different augmentation schedules

Hardware Utilized

- Processor: Intel(R) Core(TM) i5-5200U CPU @ 2.20GHz, 2195 Mhz, 2 Core(s), 4 Logical Processor(s)

- GPU: 2x NVIDIA GEFORCE 1080

- RAM: 8 GB

- HDD: 1x 1TB

- In memory: 8GB

Failed / Over Time Experiments

- ResNet50, 60epochs, 224X224, adam optimizer 0.001, batch 64, increasing augmentation schedule

    ~ 4+ hours / epoch

- InceptionV3, 30epochs, 400X400, adam optimizer 0.001, batch 32, increasing augmentation schedule

    – ~ 4+ hours /epoch

So we COMPROMIZED.

Dataset under consideration

- Data size: ~16GB.

- # validation images: ~10K.

- # test images: ~1K.

- # label classes : 20

- # human label: 551K [Confidence 1]

- # Machine labels : 512K [Confidence 0]

Data Augmentations used

- A very useful library: **albumentations**

- Horizontal flip

- One of the noises: Additive Gaussian Noise / Gauss Noise

- One of Motion Blur / Median Blur

- Shift scale rotate

- One of Optical Distortion, Grid distortion, Piecewise Affine

- One of CLAHE / Sharpen / Emboss / Random Contrast / Brightness

- HueSaturation

- Gray scale

- Jpeg Compression

Data Ensemble

- We trained data separately on

- InceptionV3

- ResNet50

- InceptionResnet V2.

- Our Module

- Weighted Avg of prediction.

Model for 20 classes
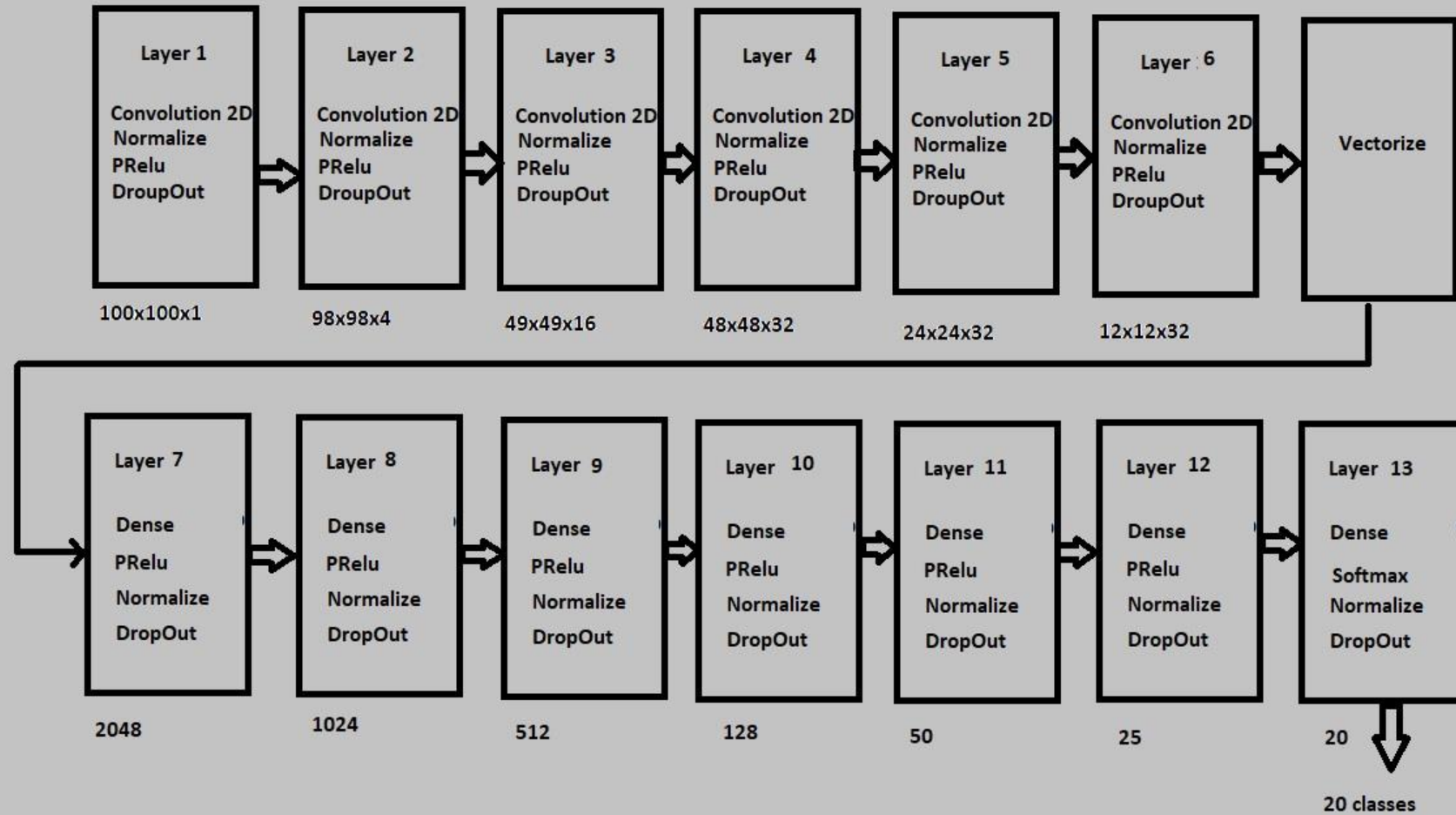
Training process

- Train/test split : 9:1 [9K : 1K]

- Epochs: 10

- Entropy loss: categorical_crossentropy.

- Optimizer: RMSProps lr 0.001

References

- albumentations

- https://github.com/albu/albumentations

- Squeeze-and-Excitation Networks

- https://arxiv.org/pdf/1709.01507.pdf

- Deep CNN Ensemble with Data Augmentation for Object Detection

- https://arxiv.org/pdf/1506.07224.pdf

- DEEP NEURAL NETWORKS UNDER STRESS

- https://arxiv.org/pdf/1605.03498.pdf