02514 Project 1.1 & 1.2

Dimosthenis Angelis(s212893) Dimitrios Bokos(S213233) Panagiotis Grigoriadis(s203269) Nikolaos Karavasilis(s213685)

HOTDOG

Objective: classify images into hotdogs and not hotdogs by designing and training a CNN.

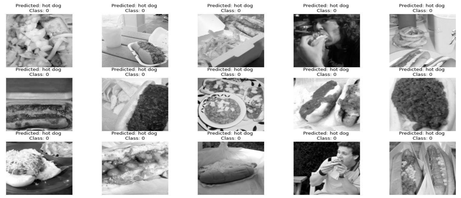
Overview

Model

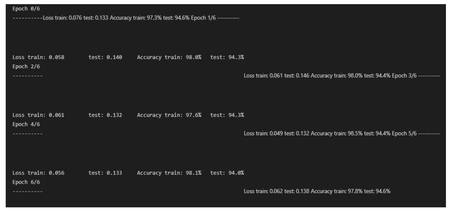
- CNN net: 3 convolutional layers, 3 max poolings followed by 2 fully connected layers.
- CrossEntropy loss paired with the Softmax function and the
- Adam optimizer.

Tasks

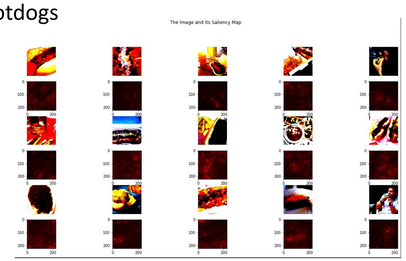
1) Accuracy of the model and visualized samples of pictures and predictions



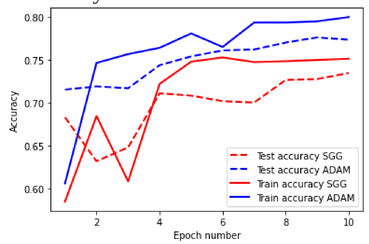
2) Transfer learning with the resnet_152 model..



3) Saliency maps and visualized images of the hotdogs



4) Data augmentation show that ADAM was the superior optimizer in both validation and train accuracy



Results

- CrossEntropy loss combined with a LogSoftmax as the last layer more optimal than the BCE loss with the Sigmoid last layer.
- Model accuracy was decent for its complexity with training accuracy around **77%** and validation accuracy around 7**5%.**
- **Transfer learning** significantly improved performance
- **ADAM** was the superior optimizer in both validation and train accuracy

Conclusions

When using the optimal parameters for our models, the accuracy improves significantly, yet naturally lacks the complexity and accuracy of more sophisticated models like ResNet.

TACO

TACO dataset:

- **1500** images
- **28** super classes
- **60** classes

Figure 1: Examples from the TACO dataset.

Preprocessing

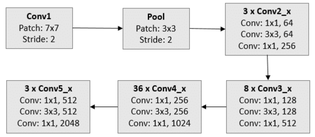
Steps:

1. Image normalization and corresponding ground truth format change
2. Split data in train (**70%**) and validation (**20%**)
3. Classification in supercategories (**28+1** for background)

Overview

Model

Finetuned **resnet_152** for image classification and bounding box regression

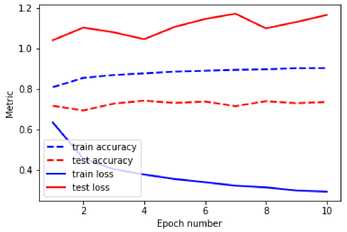


Tasks

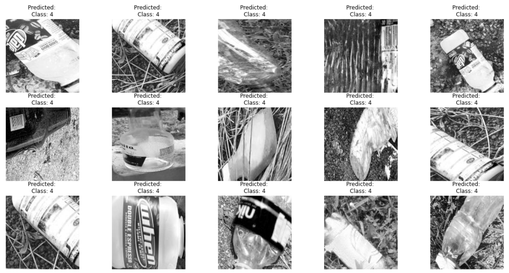
1. Preprocessing and saving of the new images and begin extracting the proposals.
2. For the extraction of the object proposals, we ended up utilizing the Edge Boxes (2000 proposals) as the Selective Search proved to be slow and inefficient.
3. Applied Intersection over Union (**IoU**)

Results

Accuracy and loss



Visualization of Classification



Conclusions

- Dataset manipulation proved challenging
- Unreliable data because:
  - Unbalanced number of images between classes
  - Small number of epochs