

Training Deep Neural Networks on Occluded Datasets: Project Status Report

Luke Bun, Shao-Jung Kan, and Cameron McCarty

May 2023

1 Outline

In this study we will train a deep convolutional neural network to classify occluded images and see how its performance compares to a network trained solely on non-occluded images.

2 Experimental design

2.1 Dataset

For our dataset, we settled on CIFAR-10 as a base. CIFAR-10 images are sufficiently large (32-by-32 pixels) and diverse (50,000 training images over 10 classes) that they would be suitable for our study of occlusion. Other relatively small datasets, such as Tiny ImageNet, which uses 64-by-64 pixel images, were thought to be excessive for our needs and using them instead would come at the cost of computation time. We also opted to use CIFAR-10 over CIFAR-100, which has the same number of images with the same size but with more diverse classes, because we hypothesized that on a dataset where there are more classes, performance will be inherently worse. So much so, that when we occluded images, performance would be so degraded that the range of accuracies between models would be too small to draw strong conclusions.

We then modified the CIFAR-10 database by occluding different parts of each image. First, we divide each of the 32-by-32 pixel CIFAR-10 images into a 3-by-3 grid of super pixels, where the superpixel in the upper left corner is 10-by-10 pixels, while the other 8 superpixels are 11-by-11 pixels. Next, for each superpixel, we create a new image with one of two occlusion strategies. The first occlusion method replaces the superpixel with a black square of equal size. The second adds Gaussian noise over the superpixel region. Each occlusion method produces a dataset that is 9 times the size of the initial CIFAR-10 dataset.

The code to download and occlude the CIFAR-10 dataset has already been completed.

[INSERT EXAMPLES OF STOCK IMAGES AND OCCLUDED IMAGES]

2.2 Model

We decided to base our model on [INSERT MODEL HERE]. Compared to larger datasets, such as ImageNet, CIFAR-10 is relatively simple. Therefore, deeper networks were deemed excessive and would greatly increase computation time. This model was originally training on the ImageNet dataset, therefore the last layer will need to be slightly modified to have fewer output classes (10 rather than 1000).

2.3 Training

We will train 3 models: one on standard CIFAR-10, one on CIFAR-10 with black superpixels, and one on CIFAR-10 with Gaussian noise superpixels. Each model will have its training parameters (learning rate and weight decay) optimized. Data will first be randomly split into training (80%) and validation (20%) sets. Then we will use a greedy approach to first optimize learning rate, testing a range of 5 values logarithmically spaced between 1 and 1e-4, and then optimizing weight decay, testing a range of 5 values logarithmically spaced between 1e-1 and 1e-5. Performance will be measured by accuracy on the validation set.

2.4 Evaluation

After training, each model will be evaluated on the standard test set, a test set occluded with black squares, and a test set occluded with Gaussian noise. Performance for each model on each test set will then be compared to determine if training on occluded images significantly improves performance on occluded image classification and which occlusion strategy is more generalizable.