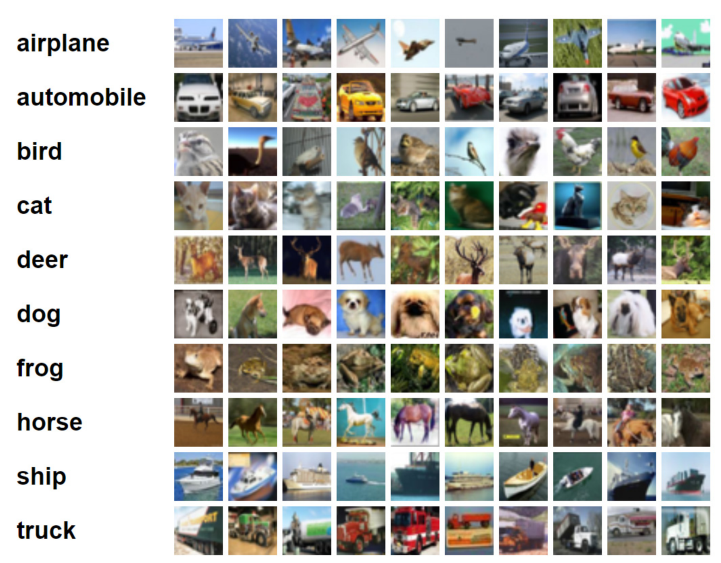
Exploring Image Recognition with Deep Neural Networks and the Effects of Normalization

Colin Jiang[1]; Luis G Sanchez Giraldo[1], Odelia Schwartz[1]
[1]Deparment of Computer Science, University of Miami

Introduction

In recent years, advances in machine learning, and particularly in deep convolutional neural networks (CNNs), have proven effective in a variety of traditionally difficult tasks including image recognition. However, due to the depth and complexity of these models, the training of these networks can be difficult, and classification results are dependent on tuning a large space of hyperparameters. There are also issues of generalization to image manipulations that were not in the training set. We explore the training of these networks with different forms of normalization, a nonlinear computation that is ubiquitous in neural models of visual processing in the brain, to improve performance. We also explore testing of networks for images with reduced contrast that were not in the training set. We find that normalization can particularly help in generalization. The training of these CNNs is done with the popular TensorFlow framework.

Fig 1. CIFAR10 Dataset



Training

CIFAR 10 dataset was used for training our convolutional neural networks. This contained 50,000 training images and 10,000 test images. Each image is size 32 x 32 x 3 and labeled out of 10 different classifications.

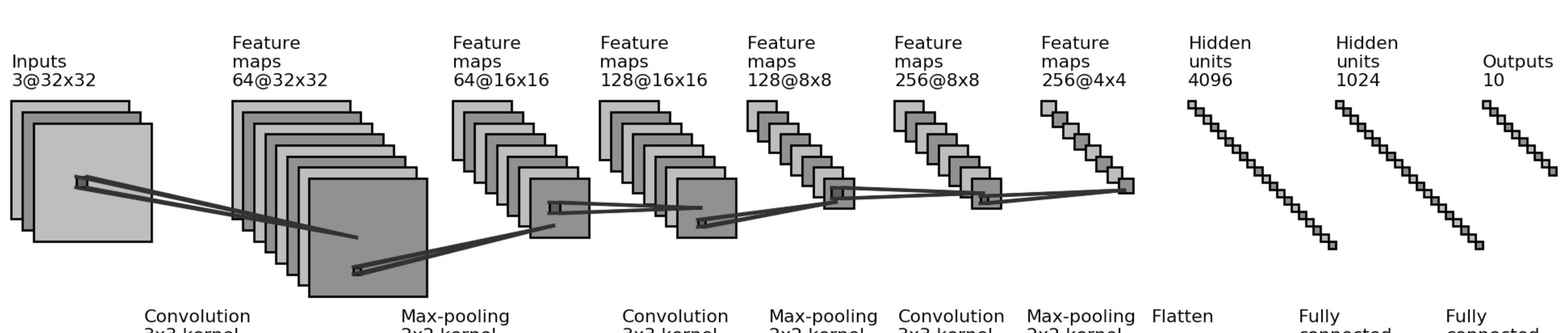


Fig 2. Convolutional network model

Convolutional network consisted of three convolutional layers, each followed by max pooling, and two fully connected layers which resulted in 10 outputs for each class.

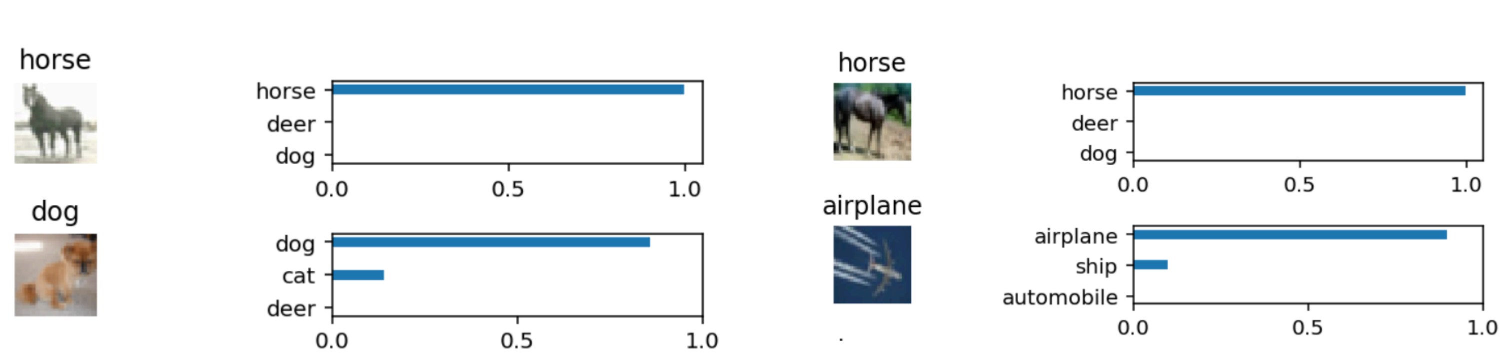


Fig 3. Sample classification

Normalization

Normalization is one computation that has gained some traction in deep learning, both for speeding up training and for improving object recognition [1,3]. Normalization is inspired by models of visual neurons in the brain [2]. In normalization, the response of a neural filter is divided by the rectified responses of other neural filter responses (figure 4). In the visual system in the brain, this helps in perceiving objects of different contrasts properly. More sophisticated forms of normalization have also been suggested that facilitate salience perception.

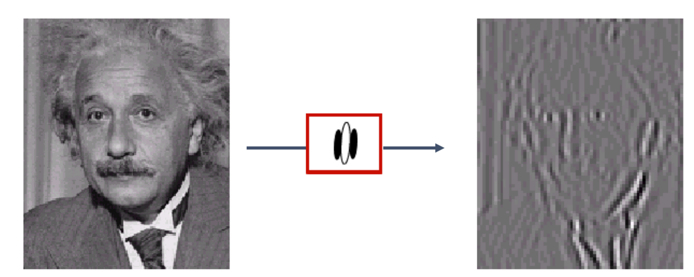


Fig 4. Image convolved by a filter

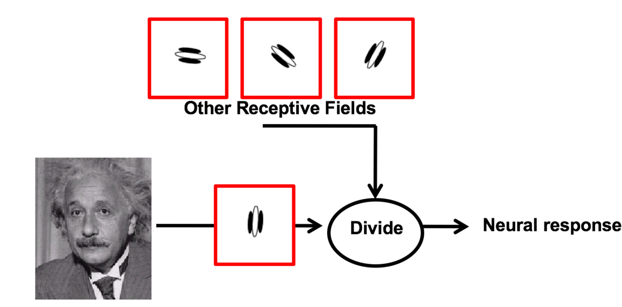


Fig 5. Local normalization

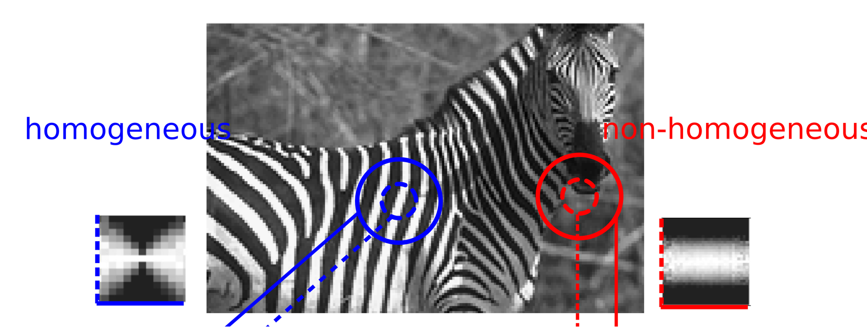


Fig 6. Flexible Normalization

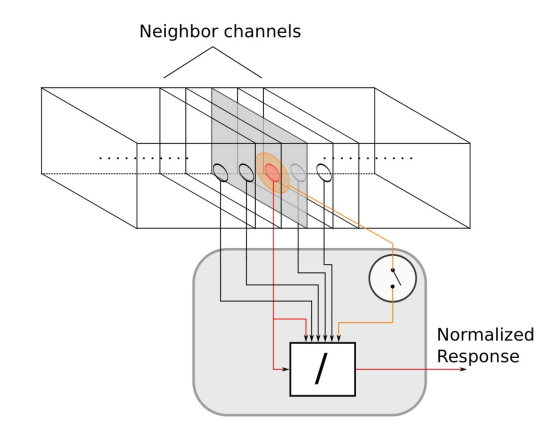


Fig 7. Normalization in CNNs

We explore two types of normalization:
(i) Local response normalization [3] dampens responses that are uniformly large in a local neighborhood. This is shown in the schematic of Figure 5.
(ii) Flexible normalization [2], which is a more accurate model of neurons in visual cortex. Flexible normalization considers whether the center and surround are statistically dependent, and if they are, then divisively normalizes by the surround. Figure 7 shows the incorporation of normalization in a deep convolutional neural network.

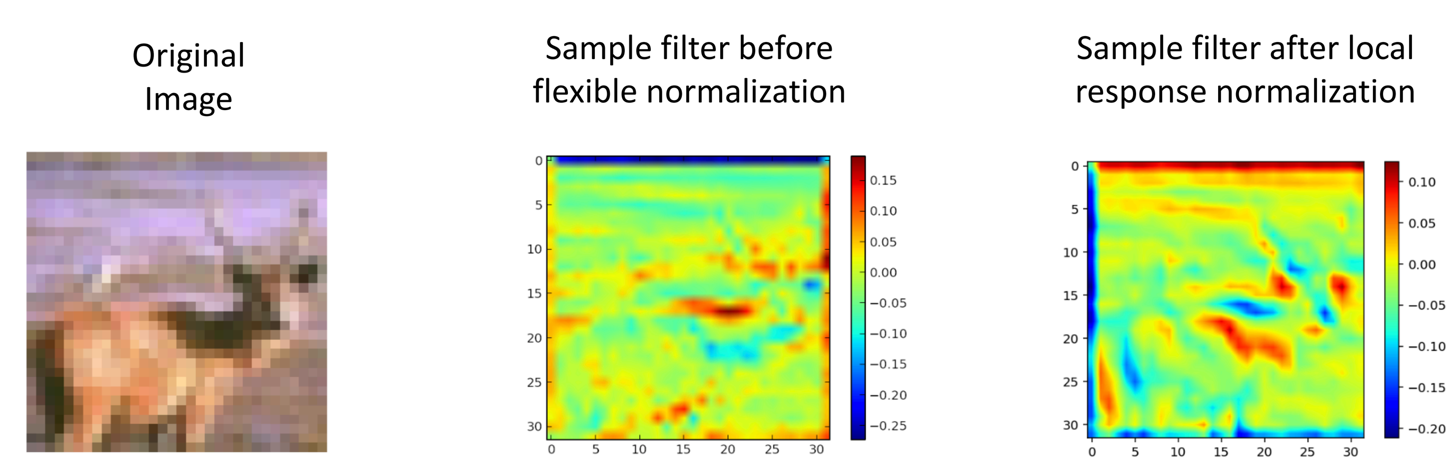


Fig 8. Local response normalization influence on filters

Results

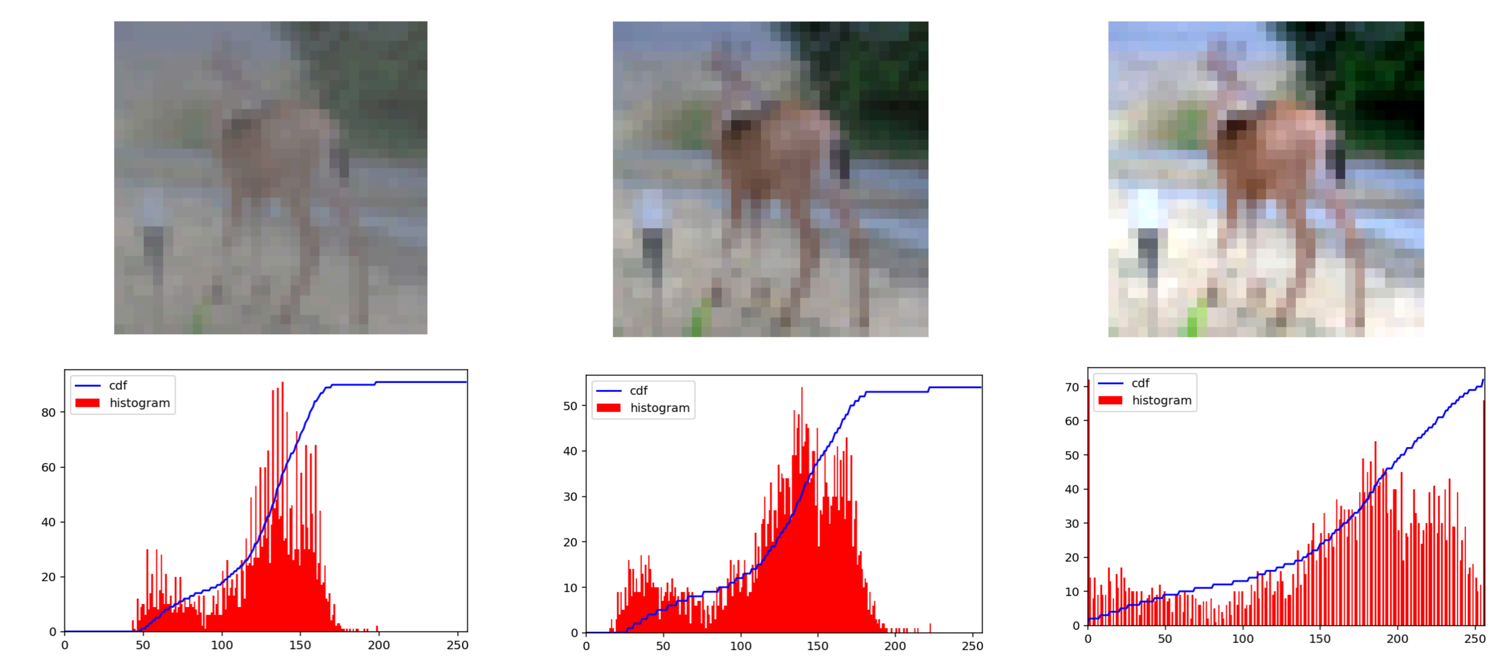
Local response normalization performed well, although marginally so. Flexible normalization did not perform as well, although there are further adjustments we need to make to properly integrate it with deep neural networks

Baseline:	80.67
Flexible normalization:	78.05
Local response normalization:	81.34

This material is based upon work supported by the National Science Foundation under Grant No. CNS-1659144, and by a NSF grant to Odelia Schwartz (1715475).

Manipulating Contrast

Although the gains in accuracy from normalization were small, we also wanted to compare the flexibility of classification of normalized models. So we manipulated the contrast of test images to see how well models with and without normalization would respond. In the figures below, the center image is a sample from the original dataset, while the left is the same image but with reduced contrast and the right is the image with its contrast increased by contrast stretching.



Results of Contrast change

The model with local response normalization performed significantly better on images with reduced contrast, even performing better than images with higher contrast, and having no significant loss in accuracy.

Baseline reduced contrast:	67.68
Baseline increased contrast:	78.49
Local response normalization reduced contrast:	80.59
Local response normalization increased contrast:	78.84

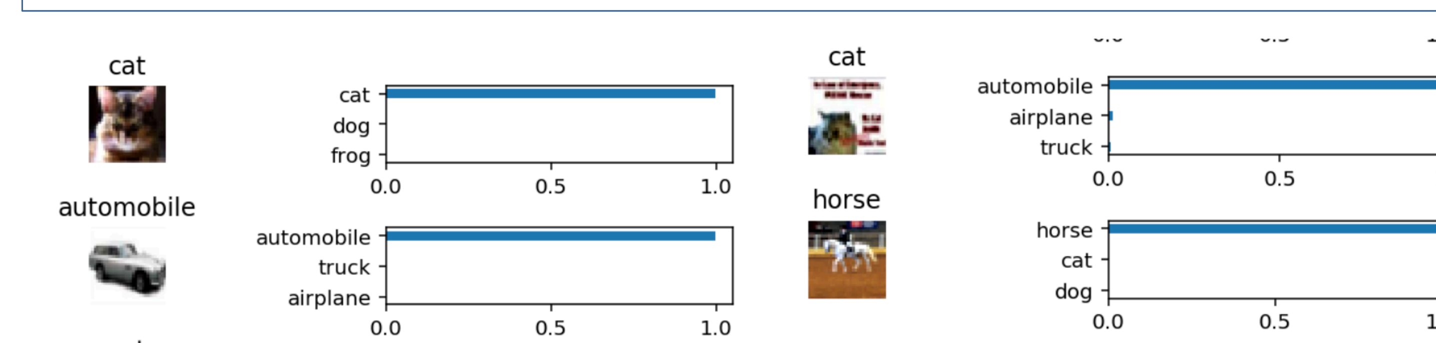


Fig 9. High contrast classification

Conclusions and Future work

Although initial results for both flexible and local response normalization was not promising, local response normalization in particular has shown a much greater ability to adapt to contrast changes and more versatile in its classification. In the future, it would be interesting to see if normalization is better adapted to other test images beyond contrast manipulation, such as background clutter and possibly adversarial examples.

Further tuning of the flexible normalization and its integration with deep neural networks is needed. Initial additions in convolution layers also led to increases in accuracy by ~2%

References

1. Ren, Mengye, et al. "Normalizing the normalizers: Comparing and extending network normalization schemes." *arXiv preprint arXiv:1611.04520* (2016).
2. R C Cagli, A Kohn, O Schwartz, Flexible Gating of Contextual Modulation During Natural Vision. Nature Neuroscience 8(11):1648-55, 2015.
3. A. Krizhevsky, I. Sutskever, G. E. Hinton. "ImageNet classification with deep convolutional nueral networks," in Advances in neural information processing systems (NIPS), 2012.