Classification of the CIFAR-10 Dataset

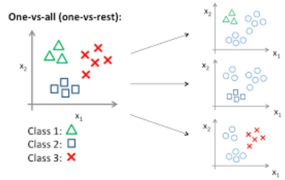
Raymond Cano, Xilin Liu
COMP 540 Machine Learning, Rice University

Abstract

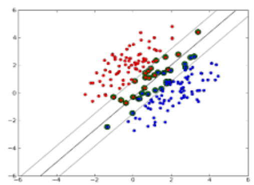
- *The CIFAR-10 dataset is a set of 60,000 images of objects. There are 10 classes of objects, ranging from cats to airplanes. The goal of this project was to create a machine learning model that could successfully classify these images. We ran multiple different classification methods on the dataset with varying success. At first, we tried using one vs. all (OVA) logistic regression. This yielded an accuracy of about 34%. Our next method was to use support vector machines (SVM) with a polynomial kernel, which yielded an accuracy of 54%. Finally, our best results came from running a convolutional neural network (CNN). Our current best accuracy is 79%.*

Types of Models

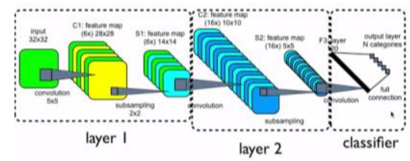
One vs. All Logistic Regression



Support Vector Machines



Convolutional Neural Networks



One vs. All Logistic Regression:
One vs All classifies comparing the probability of an item being one class against being all other classes.

Support Vector Machines:
Support Vector Machines are simple but powerful models that train on examples

Convolutional Neural Networks:
CNNs work to learn features of an image by convolving portions of the image through layers. Neurons are stacked in layers separated by nonlinearities and can peer into portions of their input image.

References

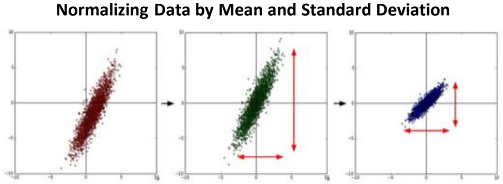
- Stanford University CS231n http://cs231n.github.io/convolutional-networks/#norm
- Kaggle CIFAR-10 Competition http://blog.kaggle.com/2015/01/02/cifar-10-competition-winners-interviews-with-dr-ben-graham-phil-culliton-zygmunt-zajac/
- He, et al. Residual Learning Network http://arxiv.org/pdf/1512.03385v1.pdf

Acknowledgement

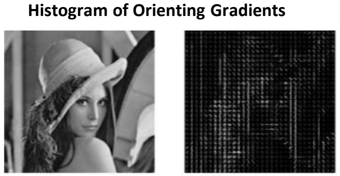
We would like to thank Professor Devika Subramanian for facilitating this project, as well as Michael Xie for great advice and motivation during the course of the project

Image Preprocessing

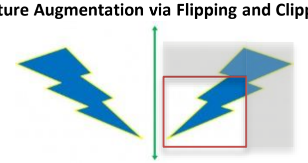
Normalizing Data by Mean and Standard Deviation



Histogram of Orienting Gradients

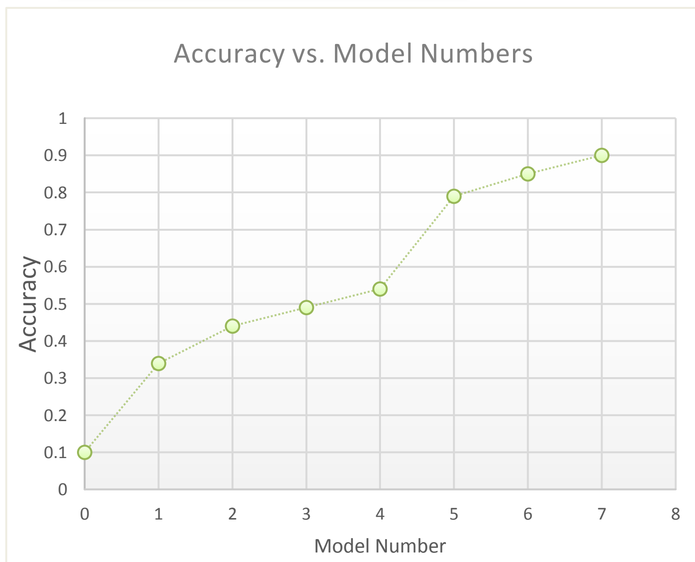


Feature Augmentation via Flipping and Clipping



One of the key factors in creating a reliable model is image preprocessing. Normalizing the data, also known as whitening, as well as flipping clipping into smaller images play huge rolls in normalizing a photo into the features a model should learn.

Results



Through a combination of discovering more complex models and preprocessing and augmenting features in a more sophisticated manner, we were able to increase our accuracy over time. Our biggest leaps in accuracy were when we changed the models we used, as seen between models 0 and 1, 1 and 2, and 4 and 5.

Our newest model, a convolutional neural network, is full of potential. Currently, without optimizing hyperparameters, we have achieved an accuracy of 79%. We hypothesize that it will reach an accuracy of 90% or more once we have maximized the model's potential through increased layering, use of dropout, and leaky ReLUs.

Model #	Date	Accuracy	Method
0	3/7/2016	0.1	Random Guess
1	3/10/2016	0.34	One Vs. All
2	3/12/2016	0.44	HOG + OVA
3	3/13/2016	0.49	HOG + SVM RBF Kernel
4	3/13/2016	0.54	HOG + SVM Polynomial Kernel
5	4/18/2016	0.79	Feed Forward Neural Net
6*	4/21/2016	0.85	Residual Learning Neural Net
7*	4/22/2016	0.9	Further Improvements