Finding Protests in Social Media Data

Gaspar Garcia Jr. | Benjamin Zhou | Dylan Moore
Stanford University | CS231n | June 6, 2017

Abstract

Determining the stability of authoritarian regimes, among them China, Russia, Iran, and Saudi Arabia, is of central importance to policy makers, businesses, societal actors, and scholars around the world. For many authoritarian regimes, protests and visible civilian unrest pose a serious threat to the regime's stability.
Independent measures of protest activity in such countries would be highly valuable for scientific and public policy purposes. However, because repressive regimes actively remove evidence of protests, there are very few publicly available metrics for this issue. We are working with Professor Jennifer Pan of Stanford University to change that. In our project we created a model that will be able to automatically identify protest images as soon as they are posted.
We are using a labeled dataset of 500,000 images taken from Chinese social media posts and compare the performance of various classifiers on their ability to discriminate protests from non-protests in China. Our work includes a baseline Support Vector Machine classifier that achieves approximately 75% accuracy on our validation set, 3 layer Convolutional Neural Network with 2 affine layers that has approximately 85% accuracy on the validation set, a deeper (5 layer) CNN with an accuracy of 85%, and a transfer learning model with an accuracy of 86%.

Foundational Research

In approaching this problem, we considered the findings of several papers (of which, the ones of particular interest are listed below as references) that better informed our understanding of what existing work was out there. This research, particularly that of Lawrence, Steve, et al. and Wang, Chuan, et al. led us to the conclusion that CNNs would ultimately be the most effective classifiers for this task. For comparison, we were inspired by the Xia et al. paper to explore using SVMs as an alternative approach.

Dataset and Features

The dataset consists of 231618 protest images and 261516 non-protest images. Since this is a supervised learning problem, we designated labels for protest as 1 and non-protest as 0. The size of the training data is about 10 GB, and images had width and height rescaled to 100x100x3, where the last dimension corresponds to the RGB depth. The test data set has 4400 images, which were randomly selected from a target dataset of 4 million images from Weibo. In the test set only about 10% of the samples are protest related images. Below we present four sample images from our dataset.

Example images from dataset

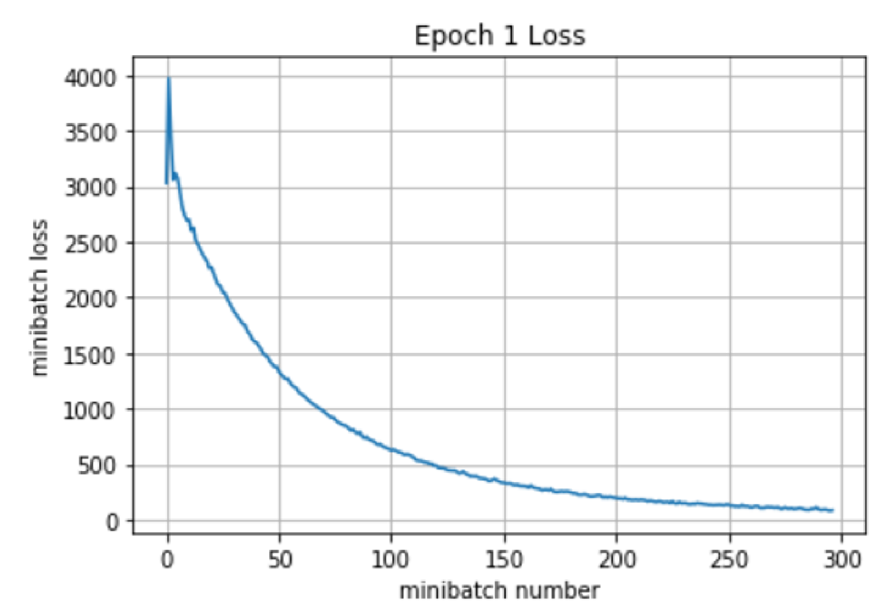
Methods

All our work was done on a Google cloud instance using Jupyter notebooks. We use Scipy for our shallow models and our deep models are all written in TensorFlow. For both models described below, we are experimenting with a mini-batch of 20,000 training images from our complete dataset. These samples were randomly chosen and have half protest images and half are non-protest images. Our validation set was a set of 2,000 images different images with the same 50/50 distribution of labels.

SVM and Baseline

Our shallow model fits an SVM over flattened images with a linear kernel. The SVM perfectly fits the 20,000 training images and achieves 75% accuracy over the 2,000 sample validation set. Previous work on this classification problem is by Han Zhang, a PhD student at Princeton. Han's work tried a two layer CNN classifier with a batch of 20,000 images. Han's CNN was evaluated on the test dataset and had precision of approximately 0.4 for protest samples.

3-Layer CNN



Our CNN model uses 3 convolutional layers followed by two affine layers. The first conv layer uses 32 4x4 filters with stride 2. The second conv layer uses 64 4x4 filters with stride 2. The third conv layer uses 64 4x4 filters with stride 1. Each convolutional layer uses relu activation and a batch normalization layer. The output from the last convolutional layer flows into a 2x2 max pool with stride 2. The max pool output is fed to an affine layer with 1048 outputs. This layer is followed by a dropout layer with dropout probability 0.5. The final layer is an affine layer with 2 outputs, one for each protest and non-protest class. We use an Adam optimizer and L2 regularization over the affine weights. The regularization rate is 5e-2 and the learning rate is 5e-4. We took inspiration from the architecture by AlexNet and VGG which use hidden affine layers with 4096 outputs preceded by 2x2 max pool.

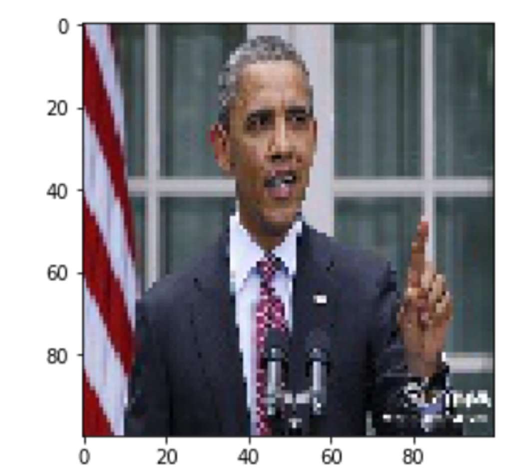
Hyperparameters and Regularization

We tuned hyperparameters over 3 epochs. The learn rate and regularization rate using random search on a logarithmic scale between 1e-2 and 1e-7. Dropout probability was tuned linearly between 0.1 and 0.5. The loss curve for the tuned model is shown above.

Deeper Neural Network

After implementing our 3-layer CNN we tried using a deeper network. We gave this model 5 convolutional layers and it achieved an accuracy of %85 on the validation set.

Transfer Learning



We implemented transfer learning by using the existing neural network SqueezeNet, trained on the ImageNet dataset, and fine tuned it with our protest dataset. Since the non-protest images could contain anything that was not protest, using a neural network that was trained on ImageNet, seemed like a good thing to try, since ImageNet contains many different categories. SqueezeNet finetuned with our protest data could differentiate non-protest from protest better. The test data set contained images that were drawn from a different distribution of images than our training data.
We added a modified version of our 3 layer neural network to the top of the 3rd layer (chosen because it achieves the highest accuracy) of SqueezeNet. Our model achieved %86 accuracy on the validation set, %89 on the training data, %62 on the test data, with a precision and recall of 0.24 and 0.72, respectively, on the test data.

Results

	Train Acc.	*Val. Acc.*	*Target Acc.*	*Recall*	*Precision*
SVM Baseline	1	0.75	-	-	-
3-Layer CNN	0.91	0.85	-	0.22	0.23
Deeper Neural Net	0.88	0.85	0.65	**0.55**	0.22
Transfer Learning	**0.89**	**0.86**	0.62	**0.72**	0.24

References

1. Wang, Chuan, et al. "Deep people counting in extremely dense crowds." 2015.
2. Lawrence, Steve, et al. "Face recognition: A convolutional neural-network approach."
3. Iandola, Forrest N., et al. "SqueezeNet"