features

October 25, 2019

1 Image features exercise

Complete and hand in this completed worksheet (including its outputs and any supporting code outside of the worksheet) with your assignment submission. For more details see the assignments page on the course website.

We have seen that we can achieve reasonable performance on an image classification task by training a linear classifier on the pixels of the input image. In this exercise we will show that we can improve our classification performance by training linear classifiers not on raw pixels but on features that are computed from the raw pixels.

All of your work for this exercise will be done in this notebook.

[21]:
```
from __future__ import print_function
import random
import numpy as np
from cs231n.data_utils import load_CIFAR10
import matplotlib.pyplot as plt

%matplotlib inline
plt.rcParams['figure.figsize'] = (10.0, 8.0) # set default size of plots
plt.rcParams['image.interpolation'] = 'nearest'
plt.rcParams['image.cmap'] = 'gray'

# for auto-reloading extenrnal modules
# see http://stackoverflow.com/questions/1907993/
↪autoreload-of-modules-in-ipython
%load_ext autoreload
%autoreload 2
```

```
The autoreload extension is already loaded. To reload it, use:
  %reload_ext autoreload
```

1.1 Load data

Similar to previous exercises, we will load CIFAR-10 data from disk.

[22]:
```
from cs231n.features import color_histogram_hsv, hog_feature
```

```python
def get_CIFAR10_data(num_training=49000, num_validation=1000, num_test=1000):
    # Load the raw CIFAR-10 data
    cifar10_dir = 'cs231n/datasets/cifar-10-batches-py'

    X_train, y_train, X_test, y_test = load_CIFAR10(cifar10_dir)

    # Subsample the data
    mask = list(range(num_training, num_training + num_validation))
    X_val = X_train[mask]
    y_val = y_train[mask]
    mask = list(range(num_training))
    X_train = X_train[mask]
    y_train = y_train[mask]
    mask = list(range(num_test))
    X_test = X_test[mask]
    y_test = y_test[mask]

    return X_train, y_train, X_val, y_val, X_test, y_test

# Cleaning up variables to prevent loading data multiple times (which may cause␣
 ↪memory issue)
try:
   del X_train, y_train
   del X_test, y_test
   print('Clear previously loaded data.')
except:
   pass

X_train, y_train, X_val, y_val, X_test, y_test = get_CIFAR10_data()
```

```
Clear previously loaded data.
```

1.2 Extract Features

For each image we will compute a Histogram of Oriented Gradients (HOG) as well as a color histogram using the hue channel in HSV color space. We form our final feature vector for each image by concatenating the HOG and color histogram feature vectors.

Roughly speaking, HOG should capture the texture of the image while ignoring color information, and the color histogram represents the color of the input image while ignoring texture. As a result, we expect that using both together ought to work better than using either alone. Verifying this assumption would be a good thing to try for your interests.

The `hog_feature` and `color_histogram_hsv` functions both operate on a single image and return a feature vector for that image. The extract_features function takes a set of images and a list of feature functions and evaluates each feature function on each image, storing the results in a matrix where each column is the concatenation of all feature vectors for a single image.

[23]:
```
from cs231n.features import *

num_color_bins = 10 # Number of bins in the color histogram
feature_fns = [hog_feature, lambda img: color_histogram_hsv(img,␣
↪nbin=num_color_bins)]
X_train_feats = extract_features(X_train, feature_fns, verbose=True)
X_val_feats = extract_features(X_val, feature_fns)
X_test_feats = extract_features(X_test, feature_fns)

# Preprocessing: Subtract the mean feature
mean_feat = np.mean(X_train_feats, axis=0, keepdims=True)
X_train_feats -= mean_feat
X_val_feats -= mean_feat
X_test_feats -= mean_feat

# Preprocessing: Divide by standard deviation. This ensures that each feature
# has roughly the same scale.
std_feat = np.std(X_train_feats, axis=0, keepdims=True)
X_train_feats /= std_feat
X_val_feats /= std_feat
X_test_feats /= std_feat

# Preprocessing: Add a bias dimension
X_train_feats = np.hstack([X_train_feats, np.ones((X_train_feats.shape[0], 1))])
X_val_feats = np.hstack([X_val_feats, np.ones((X_val_feats.shape[0], 1))])
X_test_feats = np.hstack([X_test_feats, np.ones((X_test_feats.shape[0], 1))])
```

```
Done extracting features for 1000 / 49000 images
Done extracting features for 2000 / 49000 images
Done extracting features for 3000 / 49000 images
Done extracting features for 4000 / 49000 images
Done extracting features for 5000 / 49000 images
Done extracting features for 6000 / 49000 images
Done extracting features for 7000 / 49000 images
Done extracting features for 8000 / 49000 images
Done extracting features for 9000 / 49000 images
Done extracting features for 10000 / 49000 images
Done extracting features for 11000 / 49000 images
Done extracting features for 12000 / 49000 images
Done extracting features for 13000 / 49000 images
Done extracting features for 14000 / 49000 images
Done extracting features for 15000 / 49000 images
Done extracting features for 16000 / 49000 images
Done extracting features for 17000 / 49000 images
Done extracting features for 18000 / 49000 images
Done extracting features for 19000 / 49000 images
Done extracting features for 20000 / 49000 images
```

```
Done extracting features for 21000 / 49000 images
Done extracting features for 22000 / 49000 images
Done extracting features for 23000 / 49000 images
Done extracting features for 24000 / 49000 images
Done extracting features for 25000 / 49000 images
Done extracting features for 26000 / 49000 images
Done extracting features for 27000 / 49000 images
Done extracting features for 28000 / 49000 images
Done extracting features for 29000 / 49000 images
Done extracting features for 30000 / 49000 images
Done extracting features for 31000 / 49000 images
Done extracting features for 32000 / 49000 images
Done extracting features for 33000 / 49000 images
Done extracting features for 34000 / 49000 images
Done extracting features for 35000 / 49000 images
Done extracting features for 36000 / 49000 images
Done extracting features for 37000 / 49000 images
Done extracting features for 38000 / 49000 images
Done extracting features for 39000 / 49000 images
Done extracting features for 40000 / 49000 images
Done extracting features for 41000 / 49000 images
Done extracting features for 42000 / 49000 images
Done extracting features for 43000 / 49000 images
Done extracting features for 44000 / 49000 images
Done extracting features for 45000 / 49000 images
Done extracting features for 46000 / 49000 images
Done extracting features for 47000 / 49000 images
Done extracting features for 48000 / 49000 images
```

1.3 Train SVM on features

Using the multiclass SVM code developed earlier in the assignment, train SVMs on top of the features extracted above; this should achieve better results than training SVMs directly on top of raw pixels.

[24]:
```python
# Use the validation set to tune the learning rate and regularization strength

from cs231n.classifiers.linear_classifier import LinearSVM

learning_rates = [1e-8, 1e-7]
regularization_strengths = [3e4, 5e5, 5e6]

results = {}
best_val = -1
best_svm = None

################################################################################
# TODO:                                                                        #
```

```
# Use the validation set to set the learning rate and regularization strength. #
# This should be identical to the validation that you did for the SVM; save    #
# the best trained classifer in best_svm. You might also want to play          #
# with different numbers of bins in the color histogram. If you are careful    #
# you should be able to get accuracy of near 0.44 on the validation set.       #
################################################################################
print("X_train shape: {}".format(X_train.shape))
print("X_train_feats shape: {}\n".format(X_train_feats.shape))

num_iters = 2000
num_combs = len(learning_rates) * len(regularization_strengths)
progress = 0
print("Training:")
for lr in learning_rates:
    for reg in regularization_strengths:
        print("Progress: {}/{} (LR: {:.3e}, reg: {})".format(progress,␣
 ↪num_combs, lr, int(reg)))
        # Train
        svm = LinearSVM()
        loss_hist = svm.train(X_train_feats, y_train, learning_rate=lr, reg=reg,
                              num_iters=num_iters, verbose=False)
        # Predict
        y_train_pred = svm.predict(X_train_feats)
        y_val_pred = svm.predict(X_val_feats)
        # Evaluate
        acc_train = np.mean(y_train == y_train_pred)
        acc_val = np.mean(y_val == y_val_pred)
        # Best value and SVM
        if acc_val > best_val:
            best_val = acc_val
            best_svm = svm
        # Save results
        results[(lr, reg)] = (acc_train, acc_val)
        progress += 1
print("Done training.")
################################################################################
#                              END OF YOUR CODE                                #
################################################################################

# Print out results.
for lr, reg in sorted(results):
    train_accuracy, val_accuracy = results[(lr, reg)]
    print('lr %e reg %e train accuracy: %f val accuracy: %f' % (
                lr, reg, train_accuracy, val_accuracy))

print('best validation accuracy achieved during cross-validation: %f' %␣
 ↪best_val)
```

```
X_train shape: (49000, 32, 32, 3)
X_train_feats shape: (49000, 155)

Training:
Progress: 0/6 (LR: 1.000e-08, reg: 30000)
Progress: 1/6 (LR: 1.000e-08, reg: 500000)
Progress: 2/6 (LR: 1.000e-08, reg: 5000000)
Progress: 3/6 (LR: 1.000e-07, reg: 30000)
Progress: 4/6 (LR: 1.000e-07, reg: 500000)
Progress: 5/6 (LR: 1.000e-07, reg: 5000000)
Done training.
lr 1.000000e-08 reg 3.000000e+04 train accuracy: 0.107857 val accuracy: 0.081000
lr 1.000000e-08 reg 5.000000e+05 train accuracy: 0.413041 val accuracy: 0.417000
lr 1.000000e-08 reg 5.000000e+06 train accuracy: 0.411633 val accuracy: 0.407000
lr 1.000000e-07 reg 3.000000e+04 train accuracy: 0.410184 val accuracy: 0.402000
lr 1.000000e-07 reg 5.000000e+05 train accuracy: 0.410980 val accuracy: 0.419000
lr 1.000000e-07 reg 5.000000e+06 train accuracy: 0.379755 val accuracy: 0.366000
best validation accuracy achieved during cross-validation: 0.419000
```

[25]:
```
# Evaluate your trained SVM on the test set
y_test_pred = best_svm.predict(X_test_feats)
test_accuracy = np.mean(y_test == y_test_pred)
print(test_accuracy)
```

```
0.414
```

[26]:
```
# An important way to gain intuition about how an algorithm works is to
# visualize the mistakes that it makes. In this visualization, we show examples
# of images that are misclassified by our current system. The first column
# shows images that our system labeled as "plane" but whose true label is
# something other than "plane".

examples_per_class = 8
classes = ['plane', 'car', 'bird', 'cat', 'deer', 'dog', 'frog', 'horse',␣
 ↪'ship', 'truck']
for cls, cls_name in enumerate(classes):
    idxs = np.where((y_test != cls) & (y_test_pred == cls))[0]
    idxs = np.random.choice(idxs, examples_per_class, replace=False)
    for i, idx in enumerate(idxs):
        plt.subplot(examples_per_class, len(classes), i * len(classes) + cls +␣
 ↪1)
        plt.imshow(X_test[idx].astype('uint8'))
        plt.axis('off')
        if i == 0:
            plt.title(cls_name)
plt.show()
```

1.3.1 Inline question 1:

Describe the misclassification results that you see. Do they make sense?

1.4 Neural Network on image features

Earlier in this assigment we saw that training a two-layer neural network on raw pixels achieved better classification performance than linear classifiers on raw pixels. In this notebook we have seen that linear classifiers on image features outperform linear classifiers on raw pixels.

For completeness, we should also try training a neural network on image features. This approach should outperform all previous approaches: you should easily be able to achieve over 55% classification accuracy on the test set; our best model achieves about 60% classification accuracy.

[27]:
```
# Preprocessing: Remove the bias dimension
# Make sure to run this cell only ONCE
print(X_train_feats.shape)
X_train_feats = X_train_feats[:, :-1]
X_val_feats = X_val_feats[:, :-1]
X_test_feats = X_test_feats[:, :-1]
```

```
print(X_train_feats.shape)
```

```
(49000, 155)
(49000, 154)
```

[28]:
```
from cs231n.classifiers.neural_net import TwoLayerNet

input_dim = X_train_feats.shape[1]
hidden_dim = 500
num_classes = 10

#net = TwoLayerNet(input_dim, hidden_dim, num_classes) # Not needed
best_net = None

################################################################################
# TODO: Train a two-layer neural network on image features. You may want to    #
# cross-validate various parameters as in previous sections. Store your best   #
# model in the best_net variable.                                              #
################################################################################

batch_size = 200
num_epochs = 7
num_iters  = batch_size * num_epochs
best_val   = -1
best_hyps  = {}
results    = {}
progress   = 1

# Hyperparameters generation function
def hyperparameters(gen_fns, fn_args, num):
    """Hyperparameter generator.

    Args:
        gen_fns: Generator functions
        fn_args: Arguments of gen_fns
        num: Number of iterations

    Returns:
        Tuple of hyperparameters
    """
    for _ in range(num):
        yield tuple(fn(*params) for (fn, params) in zip(gen_fns, fn_args))

lr_dec = 0.95       # This one is set
hsize = hidden_dim  # This one is set
```

```python
# Fine search (random instead of grid search, random values in focused ranges)
print("Fine search")
print("-----------")
lr_range = (-0.1, -1.0)    # Best seems to be around 10^0
reg_range = (-3.5, -1.75)  # Best seems to be around 10^-3 towards 10^-1.5
# random generation functions
lr_fn = reg_fn = lambda x,y: 10**np.random.uniform(x,y)
# number of combinations
num = 20
progress = 1
results = {}

for lr, reg in hyperparameters((lr_fn, reg_fn), (lr_range, reg_range), num):
    # Info
    print("Progress: {}/{} (lr: {:.3e}, lr_dec: {:.2f}, reg: {:.3e}, hsize: {})"
          .format(progress, num, lr, lr_dec, reg, int(hsize)))
    # Train
    net = TwoLayerNet(input_dim, hsize, num_classes)
    train_hist = net.train(X_train_feats, y_train, X_val_feats, y_val,
                           learning_rate=lr, learning_rate_decay=lr_dec,␣
↪reg=reg, num_iters=num_iters,
                           batch_size=batch_size, verbose=False)
    # Predict
    acc_val   = (net.predict(X_val_feats) == y_val).mean()
    acc_train = (net.predict(X_train_feats) == y_train).mean()
    print("prediction accuracy - train: {:.3f} | val: {:.3f}".format(acc_train,␣
↪acc_val))
    # Choose best network
    if acc_val > best_val:
        best_val = acc_val
        best_hyps = {'lr': lr, 'lr_dec': lr_dec, 'reg': reg, 'hsize': hsize}
        best_net = net
    results[(lr, lr_dec, reg, hsize)] = (acc_train, acc_val)
    progress += 1
print("Done training.\n")

# Best validation results
print("Best validation accuracy: {} ".format(best_val))
print("Hyperparameters: {}\n".format(best_hyps))

################################################################################
#                              END OF YOUR CODE                                #
################################################################################
```

```
Fine search
-----------
Progress: 1/20 (lr: 3.275e-01, lr_dec: 0.95, reg: 3.359e-04, hsize: 500)
```

```
prediction accuracy - train: 0.638 | val: 0.582
Progress: 2/20 (lr: 1.682e-01, lr_dec: 0.95, reg: 4.344e-03, hsize: 500)
prediction accuracy - train: 0.558 | val: 0.537
Progress: 3/20 (lr: 2.702e-01, lr_dec: 0.95, reg: 5.233e-04, hsize: 500)
prediction accuracy - train: 0.615 | val: 0.589
Progress: 4/20 (lr: 4.419e-01, lr_dec: 0.95, reg: 1.299e-02, hsize: 500)
prediction accuracy - train: 0.544 | val: 0.534
Progress: 5/20 (lr: 2.470e-01, lr_dec: 0.95, reg: 6.509e-03, hsize: 500)
prediction accuracy - train: 0.567 | val: 0.548
Progress: 6/20 (lr: 1.412e-01, lr_dec: 0.95, reg: 6.564e-03, hsize: 500)
prediction accuracy - train: 0.543 | val: 0.532
Progress: 7/20 (lr: 1.518e-01, lr_dec: 0.95, reg: 2.786e-03, hsize: 500)
prediction accuracy - train: 0.555 | val: 0.536
Progress: 8/20 (lr: 3.564e-01, lr_dec: 0.95, reg: 2.825e-03, hsize: 500)
prediction accuracy - train: 0.619 | val: 0.582
Progress: 9/20 (lr: 3.005e-01, lr_dec: 0.95, reg: 8.598e-03, hsize: 500)
prediction accuracy - train: 0.561 | val: 0.552
Progress: 10/20 (lr: 7.894e-01, lr_dec: 0.95, reg: 1.026e-03, hsize: 500)
prediction accuracy - train: 0.676 | val: 0.573
Progress: 11/20 (lr: 2.263e-01, lr_dec: 0.95, reg: 6.179e-03, hsize: 500)
prediction accuracy - train: 0.564 | val: 0.548
Progress: 12/20 (lr: 1.282e-01, lr_dec: 0.95, reg: 3.978e-04, hsize: 500)
prediction accuracy - train: 0.552 | val: 0.523
Progress: 13/20 (lr: 2.772e-01, lr_dec: 0.95, reg: 9.078e-04, hsize: 500)
prediction accuracy - train: 0.611 | val: 0.565
Progress: 14/20 (lr: 1.616e-01, lr_dec: 0.95, reg: 1.291e-03, hsize: 500)
prediction accuracy - train: 0.566 | val: 0.551
Progress: 15/20 (lr: 4.175e-01, lr_dec: 0.95, reg: 3.401e-04, hsize: 500)
prediction accuracy - train: 0.652 | val: 0.562
Progress: 16/20 (lr: 3.335e-01, lr_dec: 0.95, reg: 3.297e-03, hsize: 500)
prediction accuracy - train: 0.607 | val: 0.561
Progress: 17/20 (lr: 2.648e-01, lr_dec: 0.95, reg: 1.359e-03, hsize: 500)
prediction accuracy - train: 0.608 | val: 0.574
Progress: 18/20 (lr: 3.430e-01, lr_dec: 0.95, reg: 3.341e-03, hsize: 500)
prediction accuracy - train: 0.607 | val: 0.571
Progress: 19/20 (lr: 1.954e-01, lr_dec: 0.95, reg: 4.858e-03, hsize: 500)
prediction accuracy - train: 0.564 | val: 0.538
Progress: 20/20 (lr: 6.464e-01, lr_dec: 0.95, reg: 1.748e-03, hsize: 500)
prediction accuracy - train: 0.659 | val: 0.588
Done training.

Best validation accuracy: 0.589
Hyperparameters: {'lr': 0.2702112765183649, 'lr_dec': 0.95, 'reg':
0.0005233031088174872, 'hsize': 500}
```

[29]:
```
# Run your best neural net classifier on the test set. You should be able
# to get more than 55% accuracy.

test_acc = (best_net.predict(X_test_feats) == y_test).mean()
print(test_acc)
```

```
0.569
```