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 (http://vision.stanford.edu/teaching/cs231n/assignments.html)](http://vision.stanford.edu/teaching/cs231n/assignments.html) 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.

In [1]:
```
import random
import numpy as np
from cs231n.data_utils import load_CIFAR10
import matplotlib.pyplot as plt

from __future__ import print_function

%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-i
python
%load_ext autoreload
%autoreload 2
```

Load data

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

In [2]:
```
from cs231n.features import color_histogram_hsv, hog_feature

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

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

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 the bonus section.

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.

In [3]:
```
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_c
olor_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 featu
re
# 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
```

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.

In [4]:
```
# Use the validation set to tune the learning rate and regularization stren
gth

from cs231n.classifiers.linear_classifier import LinearSVM

learning_rates = [1e-9, 1e-8, 1e-7]
regularization_strengths = [5e4, 5e5, 5e6]

results = {}
best_val = -1
best_svm = None

pass
###########################################################################
#####
# TODO:
#
# Use the validation set to set the learning rate and regularization streng
th. #
# 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.
#
###########################################################################
#####
for alpha_ in learning_rates:
    for beta_ in regularization_strengths:
        print('----------------------------')
        print('Current alpha: %f, beta: %f' % (alpha_, beta_))
        svm = LinearSVM()
        loss_hist = svm.train(X_train_feats, y_train, learning_rate=alpha_,
reg=beta_,
                      num_iters=1500, verbose=False)
        y_val_pred = svm.predict(X_val_feats)
        val_pred_acc = np.mean(y_val_pred == y_val)
        y_train_pred = svm.predict(X_train_feats)
        train_pred_acc = np.mean(y_train_pred == y_train)
        results[(alpha_,beta_)]=(train_pred_acc, val_pred_acc)
        if val_pred_acc > best_val:
            best_val=val_pred_acc
            best_svm=svm
###########################################################################
#####
#                              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' % bes
t_val)
```

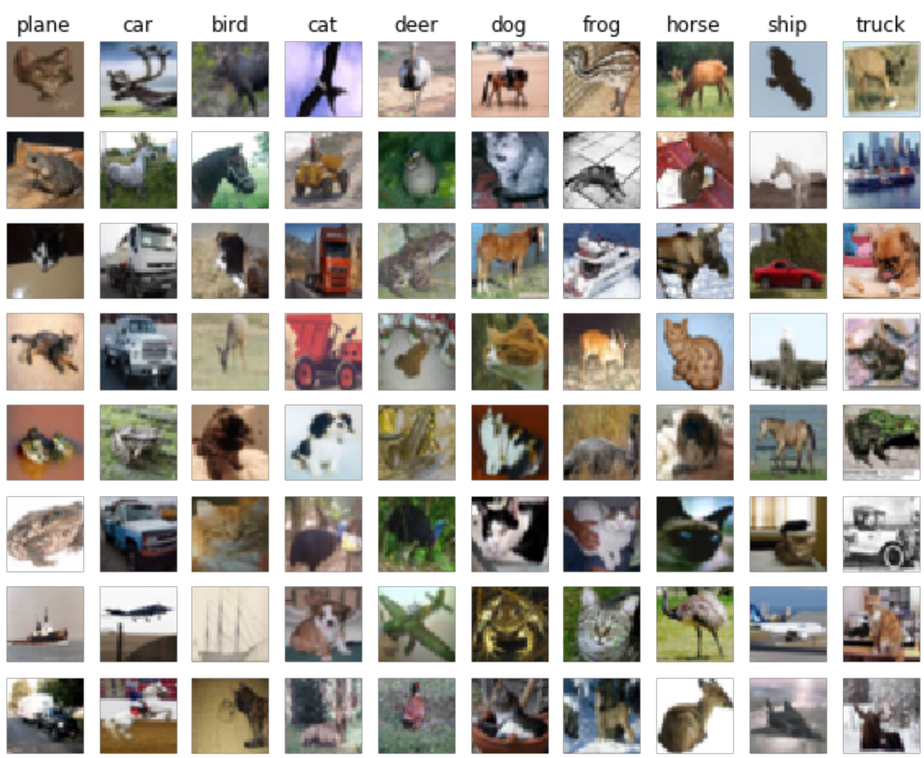
```
----------------------------
Current alpha: 0.000000, beta: 50000.000000
----------------------------
Current alpha: 0.000000, beta: 500000.000000
----------------------------
Current alpha: 0.000000, beta: 5000000.000000
----------------------------
Current alpha: 0.000000, beta: 50000.000000
----------------------------
Current alpha: 0.000000, beta: 500000.000000
----------------------------
Current alpha: 0.000000, beta: 5000000.000000
----------------------------
Current alpha: 0.000000, beta: 50000.000000
----------------------------
Current alpha: 0.000000, beta: 500000.000000
----------------------------
Current alpha: 0.000000, beta: 5000000.000000
lr 1.000000e-09 reg 5.000000e+04 train accuracy: 0.109224 val accuracy: 0.1
11000
lr 1.000000e-09 reg 5.000000e+05 train accuracy: 0.102204 val accuracy: 0.0
99000
lr 1.000000e-09 reg 5.000000e+06 train accuracy: 0.412082 val accuracy: 0.4
11000
lr 1.000000e-08 reg 5.000000e+04 train accuracy: 0.104571 val accuracy: 0.0
94000
lr 1.000000e-08 reg 5.000000e+05 train accuracy: 0.411735 val accuracy: 0.4
06000
lr 1.000000e-08 reg 5.000000e+06 train accuracy: 0.411184 val accuracy: 0.4
02000
lr 1.000000e-07 reg 5.000000e+04 train accuracy: 0.413469 val accuracy: 0.4
15000
lr 1.000000e-07 reg 5.000000e+05 train accuracy: 0.405020 val accuracy: 0.4
19000
lr 1.000000e-07 reg 5.000000e+06 train accuracy: 0.339347 val accuracy: 0.3
68000
best validation accuracy achieved during cross-validation: 0.419000
```

In [5]:

```
# 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)
```

```
0.411
```

In [6]:
```
# An important way to gain intuition about how an algorithm works is to
# visualize the mistakes that it makes. In this visualization, we show exam
ples
# 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) + cl
s + 1)
        plt.imshow(X_test[idx].astype('uint8'))
        plt.axis('off')
        if i == 0:
            plt.title(cls_name)
plt.show()
```



Inline question 1:

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

Neural Network on image features

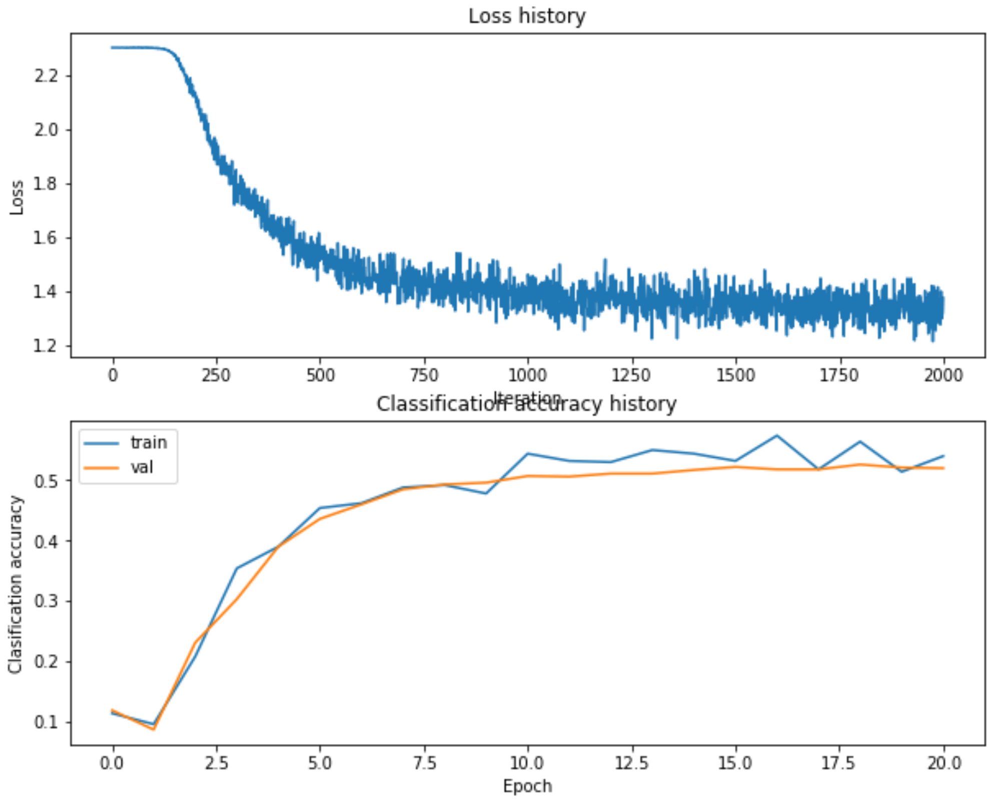
Earlier in this assigment we saw that training a two-layer neural network on raw pixels achieved better classification performance than linear classifiers on raw pixels. In this notebook we have seen that linear classifiers on image features outperform linear classifiers on raw pixels.

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.

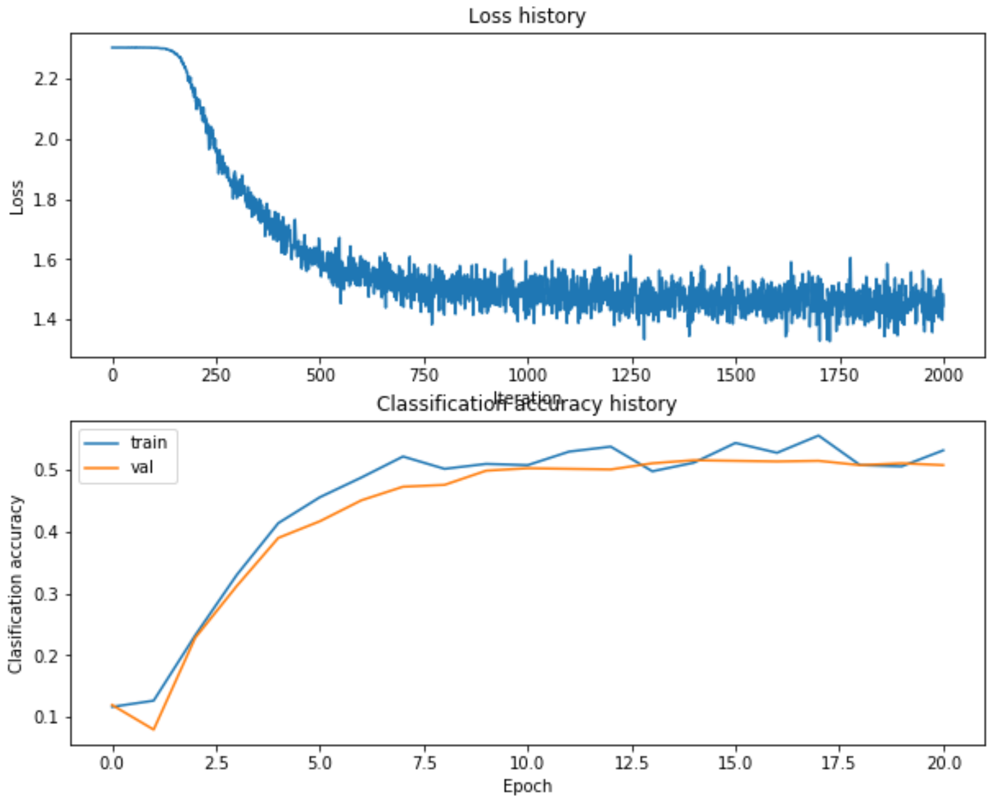
In [ ]:
```
print(X_train_feats.shape)
```

In [48]:
```
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)
best_net = None
best_val_acc = 0

###########################################################################
#####
# TODO: Train a two-layer neural network on image features. You may want to
#
# cross-validate various parameters as in previous sections. Store your bes
t   #
# model in the best_net variable.
#
###########################################################################
#####
hidden_sizes = [800]
learning_rates = [0.1, 0.5]
regs = [1e-3, 5e-3]

for hidden_size in hidden_sizes:
    for learning_rate in learning_rates:
        for reg in regs:
            print('Hidden size: ', hidden_size, ' Learning rate: ', learnin
g_rate, ' Reg: ', reg)
            net = TwoLayerNet(input_dim, hidden_size, num_classes)
            # Train the network
            stats = net.train(X_train_feats, y_train, X_val_feats, y_val,
                        num_iters=2000, batch_size=500,
                        learning_rate=learning_rate, learning_rate_decay=0.
95,
                        reg=reg, verbose=False)
            val_acc = (net.predict(X_val_feats) == y_val).mean()
            print('Validation accuracy: ', val_acc)
            print('\n')
            if val_acc > best_val_acc:
                best_val_acc = val_acc
                best_net = net
            # Plot the loss function and train / validation accuracies
            plt.subplot(2, 1, 1)
            plt.plot(stats['loss_history'])
            plt.title('Loss history')
            plt.xlabel('Iteration')
            plt.ylabel('Loss')

            plt.subplot(2, 1, 2)
            plt.plot(stats['train_acc_history'], label='train')
            plt.plot(stats['val_acc_history'], label='val')
            plt.title('Classification accuracy history')
            plt.xlabel('Epoch')
            plt.ylabel('Clasification accuracy')
            plt.legend()
            plt.show()
###########################################################################
#####
#                              END OF YOUR CODE
#
###########################################################################
#####
```

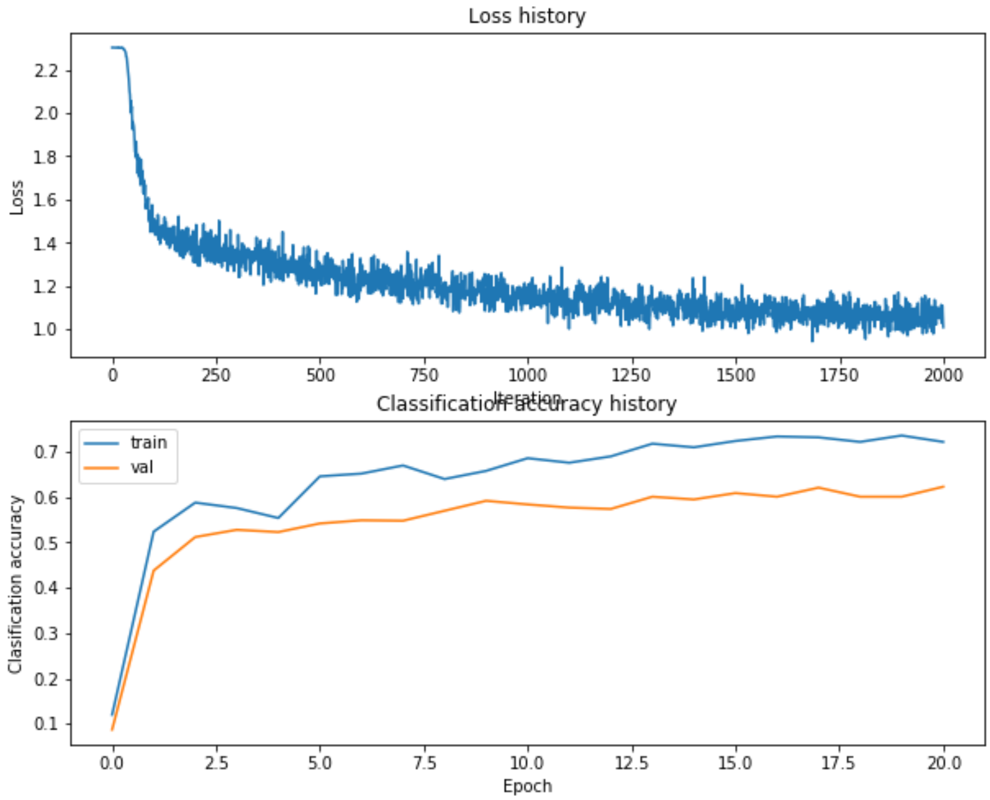
```
Hidden size:  800  Learning rate:  0.1  Reg:  0.001
Validation accuracy:  0.519
```



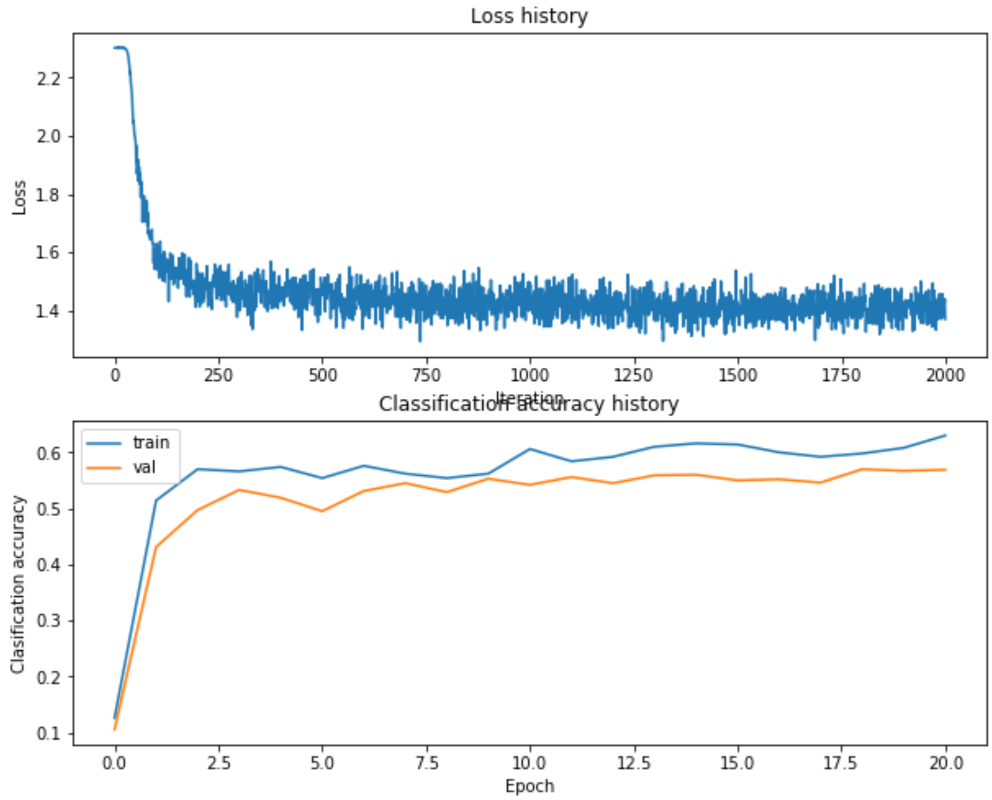

```
Hidden size:  800  Learning rate:  0.1  Reg:  0.005
Validation accuracy:  0.517
```




```
Hidden size:  800  Learning rate:  0.5  Reg:  0.001
Validation accuracy:  0.602
```




```
Hidden size:  800  Learning rate:  0.5  Reg:  0.005
Validation accuracy:  0.565
```



In [49]:
```
# Run your 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)
```

```
0.595
```

Bonus: Design your own features!

You have seen that simple image features can improve classification performance. So far we have tried HOG and color histograms, but other types of features may be able to achieve even better classification performance.

For bonus points, design and implement a new type of feature and use it for image classification on CIFAR-10. Explain how your feature works and why you expect it to be useful for image classification. Implement it in this notebook, cross-validate any hyperparameters, and compare its performance to the HOG + Color histogram baseline.

Bonus: Do something extra!

Use the material and code we have presented in this assignment to do something interesting. Was there another question we should have asked? Did any cool ideas pop into your head as you were working on the assignment? This is your chance to show off!