ConvolutionalNetworks

December 12, 2024

[ ]:
```
# Uncomment this block if you are using colabVM

# This mounts your Google Drive to the Colab VM.
from google.colab import drive
drive.mount('/content/drive')

# TODO: Enter the foldername in your Drive where you have saved the unzipped
# assignment folder, e.g. 'assignment4'
FOLDERNAME = 'cs6353/assignments/assignment4/'
assert FOLDERNAME is not None, "[!] Enter the foldername."

# Now that we've mounted your Drive, this ensures that
# the Python interpreter of the Colab VM can load
# python files from within it.
import sys
sys.path.append('/content/drive/My Drive/{}'.format(FOLDERNAME))

# This downloads the CIFAR-10 dataset to your Drive
# if it doesn't already exist.
%cd /content/drive/My\ Drive/$FOLDERNAME/cs6353/datasets/
!bash get_datasets.sh
%cd /content/drive/My\ Drive/$FOLDERNAME
```

1 Convolutional Networks

So far we have worked with deep fully-connected networks, using them to explore different optimization strategies and network architectures. Fully-connected networks are a good testbed for experimentation because they are very computationally efficient, but in practice all state-of-the-art results use convolutional networks instead.

First you will implement several layer types that are used in convolutional networks. You will then use these layers to train a convolutional network on the CIFAR-10 dataset.

[61]:
```
# As usual, a bit of setup

import numpy as np
import matplotlib.pyplot as plt
from cs6353.classifiers.cnn import *
from cs6353.data_utils import get_CIFAR10_data
from cs6353.gradient_check import eval_numerical_gradient_array,␣
  ↪eval_numerical_gradient
from cs6353.layers import *
from cs6353.fast_layers import *
from cs6353.solver import Solver

%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 external modules
# see http://stackoverflow.com/questions/1907993/
  ↪autoreload-of-modules-in-ipython
%load_ext autoreload
%autoreload 2

def rel_error(x, y):
  """ returns relative error """
  return np.max(np.abs(x - y) / (np.maximum(1e-8, np.abs(x) + np.abs(y))))

# You can ignore the message that asks you to run Python script for now.
# It will be required in the later part of the assignment.
```

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

[62]:
```
# Load the (preprocessed) CIFAR10 data.

data = get_CIFAR10_data()
for k, v in data.items():
  print('%s: ' % k, v.shape)
```

```
X_train:  (49000, 3, 32, 32)
y_train:  (49000,)
X_val:  (1000, 3, 32, 32)
y_val:  (1000,)
X_test:  (1000, 3, 32, 32)
y_test:  (1000,)
```

2 Convolution: Naive forward pass

The core of a convolutional network is the convolution operation. In the file `cs6353/layers.py`, implement the forward pass for the convolution layer in the function `conv_forward_naive`.

You don't have to worry too much about efficiency at this point; just write the code in whatever way you find most clear.

You can test your implementation by running the following:

```
[63]: x_shape = (2, 3, 4, 4)
w_shape = (3, 3, 4, 4)
x = np.linspace(-0.1, 0.5, num=np.prod(x_shape)).reshape(x_shape)
w = np.linspace(-0.2, 0.3, num=np.prod(w_shape)).reshape(w_shape)
b = np.linspace(-0.1, 0.2, num=3)

conv_param = {'stride': 2, 'pad': 1}
out, _ = conv_forward_naive(x, w, b, conv_param)
correct_out = np.array([[[[[-0.08759809, -0.10987781],
                           [-0.18387192, -0.2109216 ]],
                          [[ 0.21027089,  0.21661097],
                           [ 0.22847626,  0.23004637]],
                          [[ 0.50813986,  0.54309974],
                           [ 0.64082444,  0.67101435]]],
                         [[[-0.98053589, -1.03143541],
                           [-1.19128892, -1.24695841]],
                          [[ 0.69108355,  0.66880383],
                           [ 0.59480972,  0.56776003]],
                          [[ 2.36270298,  2.36904306],
                           [ 2.38090835,  2.38247847]]]]])

# Compare your output to ours; difference should be around 1e-8
print ('Testing conv_forward_naive')
print ('difference: ', rel_error(out, correct_out))
```

```
Testing conv_forward_naive
difference:  2.2121476417505994e-08
```

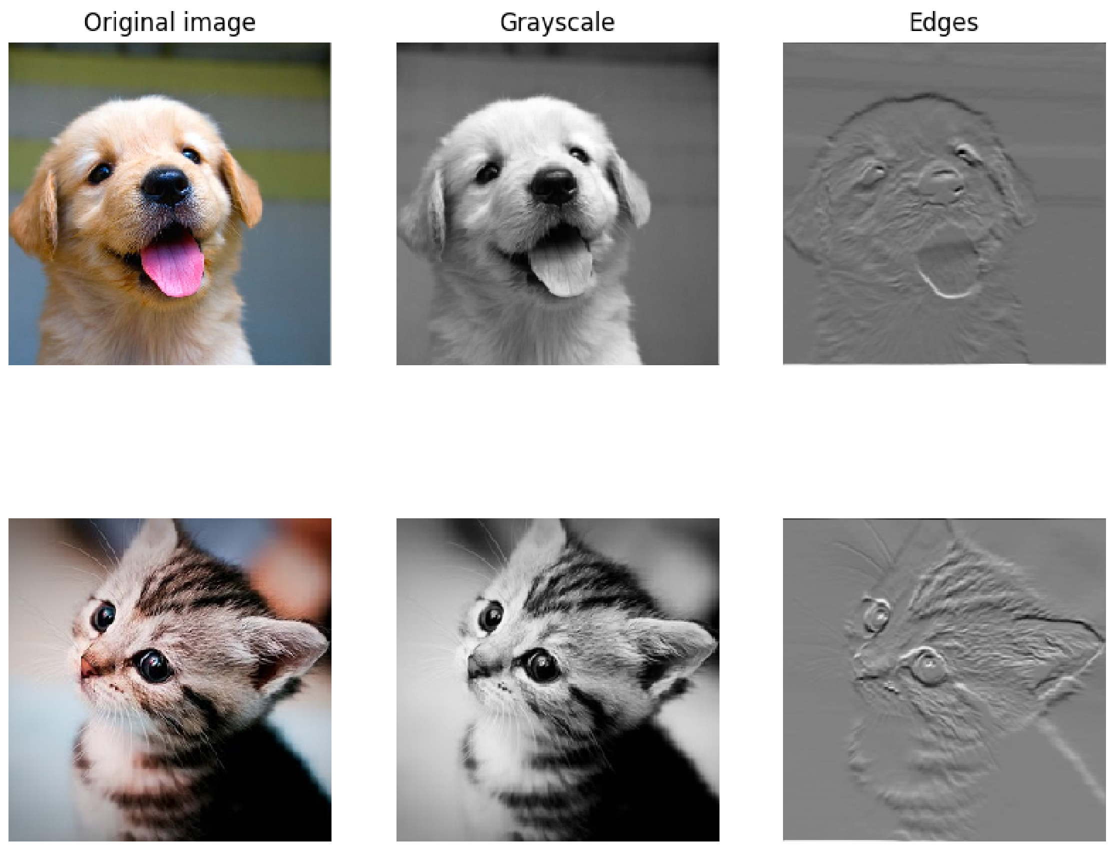
3 Aside: Image processing via convolutions

As fun way to both check your implementation and gain a better understanding of the type of operation that convolutional layers can perform, we will set up an input containing two images and manually set up filters that perform common image processing operations (grayscale conversion and edge detection). The convolution forward pass will apply these operations to each of the input images. We can then visualize the results as a sanity check.

```
[64]: import imageio.v2 as imageio
from PIL import Image

kitten, puppy = imageio.imread('kitten.jpg'), imageio.imread('puppy.jpg')
# kitten is wide, and puppy is already square
d = kitten.shape[1] - kitten.shape[0]
kitten_cropped = kitten[:, int(d/2):int(-d/2), :]

img_size = 200   # Make this smaller if it runs too slow
x = np.zeros((2, 3, img_size, img_size))
```

```
x[0, :, :, :] = np.array(Image.fromarray(puppy).resize((img_size, img_size))).
 ↪transpose((2, 0, 1))
x[1, :, :, :] = np.array(Image.fromarray(kitten_cropped).resize((img_size,␣
 ↪img_size))).transpose((2, 0, 1))

# Set up a convolutional weights holding 2 filters, each 3x3
w = np.zeros((2, 3, 3, 3))

# The first filter converts the image to grayscale.
# Set up the red, green, and blue channels of the filter.
w[0, 0, :, :] = [[0, 0, 0], [0, 0.3, 0], [0, 0, 0]]
w[0, 1, :, :] = [[0, 0, 0], [0, 0.6, 0], [0, 0, 0]]
w[0, 2, :, :] = [[0, 0, 0], [0, 0.1, 0], [0, 0, 0]]

# Second filter detects horizontal edges in the blue channel.
w[1, 2, :, :] = [[1, 2, 1], [0, 0, 0], [-1, -2, -1]]

# Vector of biases. We don't need any bias for the grayscale
# filter, but for the edge detection filter we want to add 128
# to each output so that nothing is negative.
b = np.array([0, 128])

# Compute the result of convolving each input in x with each filter in w,
# offsetting by b, and storing the results in out.
out, _ = conv_forward_naive(x, w, b, {'stride': 1, 'pad': 1})

def imshow_noax(img, normalize=True):
    """ Tiny helper to show images as uint8 and remove axis labels """
    if normalize:
        img_max, img_min = np.max(img), np.min(img)
        img = 255.0 * (img - img_min) / (img_max - img_min)
    plt.imshow(img.astype('uint8'))
    plt.gca().axis('off')

# Show the original images and the results of the conv operation
plt.subplot(2, 3, 1)
imshow_noax(puppy, normalize=False)
plt.title('Original image')
plt.subplot(2, 3, 2)
imshow_noax(out[0, 0])
plt.title('Grayscale')
plt.subplot(2, 3, 3)
imshow_noax(out[0, 1])
plt.title('Edges')
plt.subplot(2, 3, 4)
imshow_noax(kitten_cropped, normalize=False)
plt.subplot(2, 3, 5)
```

```
imshow_noax(out[1, 0])
plt.subplot(2, 3, 6)
imshow_noax(out[1, 1])
plt.show()
```



4 Convolution: Naive backward pass

Implement the backward pass for the convolution operation in the function `conv_backward_naive` in the file `cs6353/layers.py`. Again, you don't need to worry too much about computational efficiency.

When you are done, run the following to check your backward pass with a numeric gradient check.

[65]:
```
x = np.random.randn(4, 3, 5, 5)
w = np.random.randn(2, 3, 3, 3)
b = np.random.randn(2,)
dout = np.random.randn(4, 2, 5, 5)
conv_param = {'stride': 1, 'pad': 1}
```

```
dx_num = eval_numerical_gradient_array(lambda x: conv_forward_naive(x, w, b,␣
 ↪conv_param)[0], x, dout)
dw_num = eval_numerical_gradient_array(lambda w: conv_forward_naive(x, w, b,␣
 ↪conv_param)[0], w, dout)
db_num = eval_numerical_gradient_array(lambda b: conv_forward_naive(x, w, b,␣
 ↪conv_param)[0], b, dout)

out, cache = conv_forward_naive(x, w, b, conv_param)
dx, dw, db = conv_backward_naive(dout, cache)

# Your errors should be around 1e-9'
print ('Testing conv_backward_naive function')
print ('dx error: ', rel_error(dx, dx_num))
print ('dw error: ', rel_error(dw, dw_num))
print ('db error: ', rel_error(db, db_num))
```

```
Testing conv_backward_naive function
dx error:  1.472265105178055e-09
dw error:  7.250329778450262e-10
db error:  1.2676172015333743e-11
```

5 Max pooling: Naive forward

Implement the forward pass for the max-pooling operation in the function `max_pool_forward_naive` in the file `cs6353/layers.py`. Again, don't worry too much about computational efficiency.

Check your implementation by running the following:

[66]:
```
x_shape = (2, 3, 4, 4)
x = np.linspace(-0.3, 0.4, num=np.prod(x_shape)).reshape(x_shape)
pool_param = {'pool_width': 2, 'pool_height': 2, 'stride': 2}

out, _ = max_pool_forward_naive(x, pool_param)

correct_out = np.array([[[[-0.26315789, -0.24842105],
                          [-0.20421053, -0.18947368]],
                         [[-0.14526316, -0.13052632],
                          [-0.08631579, -0.07157895]],
                         [[-0.02736842, -0.01263158],
                          [ 0.03157895,  0.04631579]]],
                        [[[ 0.09052632,  0.10526316],
                          [ 0.14947368,  0.16421053]],
                         [[ 0.20842105,  0.22315789],
                          [ 0.26736842,  0.28210526]],
                         [[ 0.32631579,  0.34105263],
                          [ 0.38526316,  0.4       ]]]])
```

```
# Compare your output with ours. Difference should be around 1e-8.
print ('Testing max_pool_forward_naive function:')
print ('difference: ', rel_error(out, correct_out))
```

```
Testing max_pool_forward_naive function:
difference:  4.1666665157267834e-08
```

6 Max pooling: Naive backward

Implement the backward pass for the max-pooling operation in the function `max_pool_backward_naive` in the file `cs6353/layers.py`. You don't need to worry about computational efficiency.

Check your implementation with numeric gradient checking by running the following:

[67]:
```
x = np.random.randn(3, 2, 8, 8)
dout = np.random.randn(3, 2, 4, 4)
pool_param = {'pool_height': 2, 'pool_width': 2, 'stride': 2}

dx_num = eval_numerical_gradient_array(lambda x: max_pool_forward_naive(x,␣
 ↪pool_param)[0], x, dout)

out, cache = max_pool_forward_naive(x, pool_param)
dx = max_pool_backward_naive(dout, cache)

# Your error should be around 1e-12
print ('Testing max_pool_backward_naive function:')
print ('dx error: ', rel_error(dx, dx_num))
```

```
Testing max_pool_backward_naive function:
dx error:  3.2756356911410603e-12
```

7 Fast layers

Making convolution and pooling layers fast can be challenging. To spare you the pain, we've provided fast implementations of the forward and backward passes for convolution and pooling layers in the file `cs6353/fast_layers.py`.

The fast convolution implementation depends on a Cython extension; to compile it you need to run the following from the `cs6353` directory:

```
python setup.py build_ext --inplace
```

The API for the fast versions of the convolution and pooling layers is exactly the same as the naive versions that you implemented above: the forward pass receives data, weights, and parameters and produces outputs and a cache object; the backward pass recieves upstream derivatives and the cache object and produces gradients with respect to the data and weights.

NOTE: The fast implementation for pooling will only perform optimally if the pooling regions are non-overlapping and tile the input. If these conditions are not met then the fast pooling implementation will not be much faster than the naive implementation.

You can compare the performance of the naive and fast versions of these layers by running the following:

[68]:
```
!pip install colab-xterm
%load_ext colabxterm
%xterm

# Once the terminal loads, run the following commands
# cd cs6353
# python setup.py build_ext --inplace
```

```
Requirement already satisfied: colab-xterm in /usr/local/lib/python3.10/dist-
packages (0.2.0)
Requirement already satisfied: ptyprocess~=0.7.0 in
/usr/local/lib/python3.10/dist-packages (from colab-xterm) (0.7.0)
Requirement already satisfied: tornado>5.1 in /usr/local/lib/python3.10/dist-
packages (from colab-xterm) (6.3.3)
The colabxterm extension is already loaded. To reload it, use:
  %reload_ext colabxterm
```

```
Launching Xterm…
```

```
<IPython.core.display.Javascript object>
```

[76]:
```
from cs6353.fast_layers import conv_forward_fast, conv_backward_fast
from time import time

x = np.random.randn(100, 3, 31, 31)
w = np.random.randn(25, 3, 3, 3)
b = np.random.randn(25,)
dout = np.random.randn(100, 25, 16, 16)
conv_param = {'stride': 2, 'pad': 1}

t0 = time()
out_naive, cache_naive = conv_forward_naive(x, w, b, conv_param)
t1 = time()
out_fast, cache_fast = conv_forward_fast(x, w, b, conv_param)
t2 = time()

print ('Testing conv_forward_fast:')
print ('Naive: %fs' % (t1 - t0))
print ('Fast: %fs' % (t2 - t1))
print ('Speedup: %fx' % ((t1 - t0) / (t2 - t1)))
print ('Difference: ', rel_error(out_naive, out_fast))
```

```
t0 = time()
dx_naive, dw_naive, db_naive = conv_backward_naive(dout, cache_naive)
t1 = time()
dx_fast, dw_fast, db_fast = conv_backward_fast(dout, cache_fast)
t2 = time()

print ()
print ('Testing conv_backward_fast:')
print ('Naive: %fs' % (t1 - t0))
print ('Fast: %fs' % (t2 - t1))
print ('Speedup: %fx' % ((t1 - t0) / (t2 - t1)))
print ('dx difference: ', rel_error(dx_naive, dx_fast))
print ('dw difference: ', rel_error(dw_naive, dw_fast))
print ('db difference: ', rel_error(db_naive, db_fast))
```

```
Testing conv_forward_fast:
Naive: 6.859080s
Fast: 0.008989s
Speedup: 763.085011x
Difference:  4.3578032479370686e-10

Testing conv_backward_fast:
Naive: 7.095611s
Fast: 0.010022s
Speedup: 707.991912x
dx difference:  2.088599853459l3e-11
dw difference:  6.454680400478933e-12
db difference:  0.0
```

[77]:
```
from cs6353.fast_layers import max_pool_forward_fast, max_pool_backward_fast

x = np.random.randn(100, 3, 32, 32)
dout = np.random.randn(100, 3, 16, 16)
pool_param = {'pool_height': 2, 'pool_width': 2, 'stride': 2}

t0 = time()
out_naive, cache_naive = max_pool_forward_naive(x, pool_param)
t1 = time()
out_fast, cache_fast = max_pool_forward_fast(x, pool_param)
t2 = time()

print ('Testing pool_forward_fast:')
print ('Naive: %fs' % (t1 - t0))
print ('fast: %fs' % (t2 - t1))
print ('speedup: %fx' % ((t1 - t0) / (t2 - t1)))
print ('difference: ', rel_error(out_naive, out_fast))
```

```
t0 = time()
dx_naive = max_pool_backward_naive(dout, cache_naive)
t1 = time()
dx_fast = max_pool_backward_fast(dout, cache_fast)
t2 = time()

print ()
print ('Testing pool_backward_fast:')
print ('Naive: %fs' % (t1 - t0))
print ('speedup: %fx' % ((t1 - t0) / (t2 - t1)))
print ('dx difference: ', rel_error(dx_naive, dx_fast))
```

```
Testing pool_forward_fast:
Naive: 0.426840s
fast: 0.004890s
speedup: 87.293286x
difference:  0.0

Testing pool_backward_fast:
Naive: 0.629619s
speedup: 30.826862x
dx difference:  0.0
```

8 Convolutional "sandwich" layers

Previously we introduced the concept of "sandwich" layers that combine multiple operations into commonly used patterns. In the file `cs6353/layer_utils.py` you will find sandwich layers that implement a few commonly used patterns for convolutional networks.

[78]:
```
from cs6353.layer_utils import conv_relu_pool_forward, conv_relu_pool_backward

x = np.random.randn(2, 3, 16, 16)
w = np.random.randn(3, 3, 3, 3)
b = np.random.randn(3,)
dout = np.random.randn(2, 3, 8, 8)
conv_param = {'stride': 1, 'pad': 1}
pool_param = {'pool_height': 2, 'pool_width': 2, 'stride': 2}

out, cache = conv_relu_pool_forward(x, w, b, conv_param, pool_param)
dx, dw, db = conv_relu_pool_backward(dout, cache)

dx_num = eval_numerical_gradient_array(lambda x: conv_relu_pool_forward(x, w,␣
  ↪b, conv_param, pool_param)[0], x, dout)
dw_num = eval_numerical_gradient_array(lambda w: conv_relu_pool_forward(x, w,␣
  ↪b, conv_param, pool_param)[0], w, dout)
db_num = eval_numerical_gradient_array(lambda b: conv_relu_pool_forward(x, w,␣
  ↪b, conv_param, pool_param)[0], b, dout)
```

```
print ('Testing conv_relu_pool')
print ('dx error: ', rel_error(dx_num, dx))
print ('dw error: ', rel_error(dw_num, dw))
print ('db error: ', rel_error(db_num, db))
```

```
Testing conv_relu_pool
dx error:  1.4589303505086408e-07
dw error:  3.6201785002692234e-10
db error:  1.2509595871629458e-11
```

[79]:
```
from cs6353.layer_utils import conv_relu_forward, conv_relu_backward

x = np.random.randn(2, 3, 8, 8)
w = np.random.randn(3, 3, 3, 3)
b = np.random.randn(3,)
dout = np.random.randn(2, 3, 8, 8)
conv_param = {'stride': 1, 'pad': 1}

out, cache = conv_relu_forward(x, w, b, conv_param)
dx, dw, db = conv_relu_backward(dout, cache)

dx_num = eval_numerical_gradient_array(lambda x: conv_relu_forward(x, w, b,␣
 ↪conv_param)[0], x, dout)
dw_num = eval_numerical_gradient_array(lambda w: conv_relu_forward(x, w, b,␣
 ↪conv_param)[0], w, dout)
db_num = eval_numerical_gradient_array(lambda b: conv_relu_forward(x, w, b,␣
 ↪conv_param)[0], b, dout)

print ('Testing conv_relu:')
print ('dx error: ', rel_error(dx_num, dx))
print ('dw error: ', rel_error(dw_num, dw))
print ('db error: ', rel_error(db_num, db))
```

```
Testing conv_relu:
dx error:  3.21759849838901e-08
dw error:  6.324398219004936e-10
db error:  1.2914876571539214e-11
```

9 Three-layer ConvNet

Now that you have implemented all the necessary layers, we can put them together into a simple convolutional network.

Open the file `cs6353/cnn.py` and complete the implementation of the `ThreeLayerConvNet` class. Run the following cells to help you debug:

9.1 Sanity check loss

After you build a new network, one of the first things you should do is sanity check the loss. When we use the softmax loss, we expect the loss for random weights (and no regularization) to be about `log(C)` for `C` classes. When we add regularization this should go up.

[104]:
```
model = ThreeLayerConvNet()

N = 50
X = np.random.randn(N, 3, 32, 32)
y = np.random.randint(10, size=N)

loss, grads = model.loss(X, y)
print ('Initial loss (no regularization): ', loss)

model.reg = 0.5
loss, grads = model.loss(X, y)
print ('Initial loss (with regularization): ', loss)
```

```
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Initial loss (no regularization):  2.3025855155320643
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Initial loss (with regularization):  2.508801850828695
```

9.2 Gradient check

After the loss looks reasonable, use numeric gradient checking to make sure that your backward pass is correct. When you use numeric gradient checking you should use a small amount of artifical data and a small number of neurons at each layer.

[108]:
```
num_inputs = 2
input_dim = (3, 16, 16)
reg = 0.0
num_classes = 10
X = np.random.randn(num_inputs, *input_dim)
y = np.random.randint(num_classes, size=num_inputs)

model = ThreeLayerConvNet(num_filters=3, filter_size=3,
                          input_dim=input_dim, hidden_dim=7,
                          dtype=np.float64)
loss, grads = model.loss(X, y)
for param_name in sorted(grads):
    f = lambda _: model.loss(X, y)[0]
    param_grad_num = eval_numerical_gradient(f, model.params[param_name],␣
 ↪verbose=False, h=1e-6)
```

```
    e = rel_error(param_grad_num, grads[param_name])
    print ('%s max relative error: %e' % (param_name, rel_error(param_grad_num,␣
 ↪grads[param_name])))
```

```
Shape of out_pool before flattening: (2, 3, 8, 8)
```

```
---------------------------------------------------------------------------
ValueError                                Traceback (most recent call last)
<ipython-input-108-15463bb09c3b> in <cell line: 11>()
      9                                 input_dim=input_dim, hidden_dim=7,
     10                                 dtype=np.float64)
---> 11 loss, grads = model.loss(X, y)
     12 for param_name in sorted(grads):
     13     f = lambda _: model.loss(X, y)[0]

/content/drive/MyDrive/cs6353/assignments/assignment4/cs6353/classifiers/cnn.py␣
 ↪in loss(self, X, y)
     66
     67         # First affine layer
---> 68         out_affine1, affine_cache1 = affine_forward(out_pool_flat, W2,␣
 ↪b2)
     69         out_relu2, relu_cache2 = relu_forward(out_affine1)
     70

/content/drive/MyDrive/cs6353/assignments/assignment4/cs6353/layers.py in␣
 ↪affine_forward(x, w, b)
     23     """
     24     out = None
---> 25     out = np.dot(x.reshape(x.shape[0], -1), w) + b
     26     cache = (x, w, b)
     27     return out, cache

ValueError: shapes (2,192) and (300,7) not aligned: 192 (dim 1) != 300 (dim 0)
```

9.3 Overfit small data

A nice trick is to train your model with just a few training samples. You should be able to overfit small datasets, which will result in very high training accuracy and comparatively low validation accuracy.

[109]:
```
num_train = 100
small_data = {
  'X_train': data['X_train'][:num_train],
  'y_train': data['y_train'][:num_train],
  'X_val': data['X_val'],
  'y_val': data['y_val'],
}
```

```
model = ThreeLayerConvNet(weight_scale=1e-2)

solver = Solver(model, small_data,
                num_epochs=10, batch_size=50,
                update_rule='adam',
                optim_config={
                  'learning_rate': 1e-3,
                },
                verbose=True, print_every=1)
solver.train()
```

```
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 1 / 20) loss: 2.427496
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
(Epoch 0 / 10) train acc: 0.160000; val_acc: 0.119000
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 2 / 20) loss: 4.302295
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
(Epoch 1 / 10) train acc: 0.190000; val_acc: 0.108000
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 3 / 20) loss: 2.383680
```

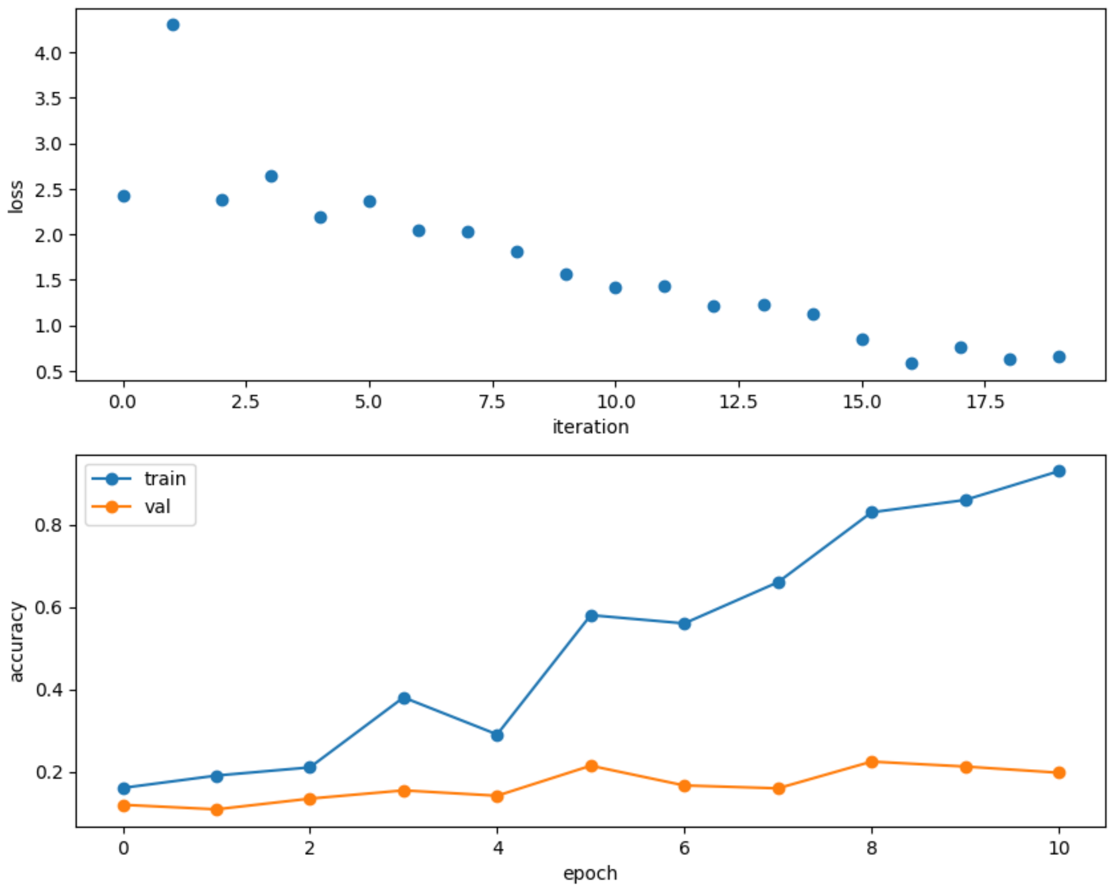
```
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 4 / 20) loss: 2.640144
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
(Epoch 2 / 10) train acc: 0.210000; val_acc: 0.134000
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 5 / 20) loss: 2.183691
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 6 / 20) loss: 2.367693
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
(Epoch 3 / 10) train acc: 0.380000; val_acc: 0.154000
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 7 / 20) loss: 2.040125
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 8 / 20) loss: 2.026254
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
(Epoch 4 / 10) train acc: 0.290000; val_acc: 0.141000
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 9 / 20) loss: 1.810902
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 10 / 20) loss: 1.557093
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
(Epoch 5 / 10) train acc: 0.580000; val_acc: 0.214000
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 11 / 20) loss: 1.423405
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 12 / 20) loss: 1.424290
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
(Epoch 6 / 10) train acc: 0.560000; val_acc: 0.166000
```

```
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 13 / 20) loss: 1.217583
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 14 / 20) loss: 1.223311
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
(Epoch 7 / 10) train acc: 0.660000; val_acc: 0.159000
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 15 / 20) loss: 1.125955
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 16 / 20) loss: 0.844242
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
(Epoch 8 / 10) train acc: 0.830000; val_acc: 0.224000
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 17 / 20) loss: 0.584382
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 18 / 20) loss: 0.754087
```

```
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
(Epoch 9 / 10) train acc: 0.860000; val_acc: 0.212000
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 19 / 20) loss: 0.627901
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 20 / 20) loss: 0.653161
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
(Epoch 10 / 10) train acc: 0.930000; val_acc: 0.197000
```

link textPlotting the loss, training accuracy, and validation accuracy should show clear overfitting:

[111]:
```
plt.subplot(2, 1, 1)
plt.plot(solver.loss_history, 'o')
plt.xlabel('iteration')
plt.ylabel('loss')

plt.subplot(2, 1, 2)
plt.plot(solver.train_acc_history, '-o')
plt.plot(solver.val_acc_history, '-o')
plt.legend(['train', 'val'], loc='upper left')
plt.xlabel('epoch')
plt.ylabel('accuracy')
plt.show()
```



9.4 Train the net

By training the three-layer convolutional network for one epoch, you should achieve greater than 40% accuracy on the training set:

```
[112]: model = ThreeLayerConvNet(weight_scale=0.001, hidden_dim=500, reg=0.001)

       solver = Solver(model, data,
                       num_epochs=1, batch_size=50,
                       update_rule='adam',
                       optim_config={
                         'learning_rate': 1e-3,
                       },
                       verbose=True, print_every=20)
       solver.train()
```

```
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 1 / 980) loss: 2.304631
```

```
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
(Epoch 0 / 1) train acc: 0.101000; val_acc: 0.112000
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 21 / 980) loss: 2.067211
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 41 / 980) loss: 1.970479
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 61 / 980) loss: 2.097452
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 81 / 980) loss: 2.064666
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 101 / 980) loss: 1.720237
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 121 / 980) loss: 1.783641
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 141 / 980) loss: 2.125387
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 161 / 980) loss: 1.751085
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 181 / 980) loss: 1.697683
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 201 / 980) loss: 1.567267
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 221 / 980) loss: 1.851034
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 241 / 980) loss: 1.683274
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 261 / 980) loss: 1.853420
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 281 / 980) loss: 1.866825
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 301 / 980) loss: 1.499229
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 321 / 980) loss: 1.708968
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 341 / 980) loss: 1.807802
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 361 / 980) loss: 1.726172
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 381 / 980) loss: 1.863100
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 401 / 980) loss: 1.598135
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 421 / 980) loss: 1.614962
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 441 / 980) loss: 1.721140
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 461 / 980) loss: 1.476499
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 481 / 980) loss: 1.738895
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 501 / 980) loss: 1.802796
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 521 / 980) loss: 1.557482
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 541 / 980) loss: 1.540524
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
```

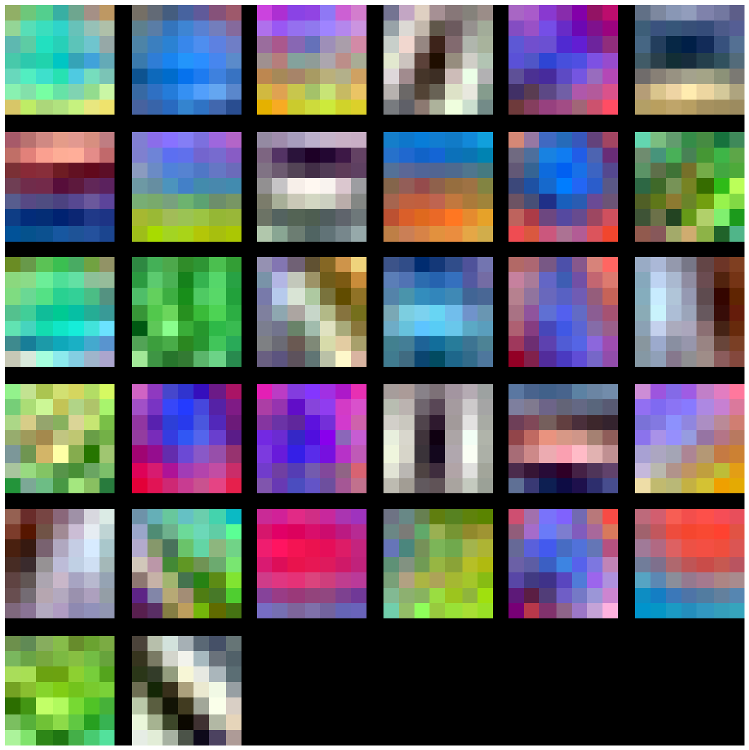
```
(Iteration 561 / 980) loss: 1.816611
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 581 / 980) loss: 1.734525
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 601 / 980) loss: 1.561079
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 621 / 980) loss: 1.697754
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 641 / 980) loss: 1.388067
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 661 / 980) loss: 1.544513
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 681 / 980) loss: 1.582704
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 701 / 980) loss: 1.573182
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 721 / 980) loss: 1.402743
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 741 / 980) loss: 1.612107
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 761 / 980) loss: 1.318237
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 781 / 980) loss: 1.568935
```

```
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 801 / 980) loss: 1.272946
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 821 / 980) loss: 1.318513
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 841 / 980) loss: 1.648179
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 861 / 980) loss: 1.656901
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 881 / 980) loss: 1.343644
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 901 / 980) loss: 1.384238
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 921 / 980) loss: 1.430401
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 941 / 980) loss: 1.690877
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
(Iteration 961 / 980) loss: 1.392569
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (50, 32, 16, 16)
Shape of dout after first affine backward: (50, 8192)
Shape of dout after max-pool backward: (50, 32, 32, 32)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
Shape of out_pool before flattening: (100, 32, 16, 16)
(Epoch 1 / 1) train acc: 0.508000; val_acc: 0.498000
```

9.5 Visualize Filters

You can visualize the first-layer convolutional filters from the trained network by running the following:

[114]:
```
from cs6353.vis_utils import visualize_grid

grid = visualize_grid(model.params['W1'].transpose(0, 2, 3, 1))
plt.imshow(grid.astype('uint8'))
plt.axis('off')
plt.gcf().set_size_inches(5, 5)
plt.show()
```



10 Spatial Batch Normalization

We already saw that batch normalization is a very useful technique for training deep fully-connected networks. Batch normalization can also be used for convolutional networks, but we need to tweak it a bit; the modification will be called "spatial batch normalization."

Normally batch-normalization accepts inputs of shape (N, D) and produces outputs of shape (N, D), where we normalize across the minibatch dimension N. For data coming from convolutional layers, batch normalization needs to accept inputs of shape (N, C, H, W) and produce outputs of shape (N, C, H, W) where the N dimension gives the minibatch size and the (H, W) dimensions give the spatial size of the feature map.

If the feature map was produced using convolutions, then we expect the statistics of each feature channel to be relatively consistent both between different imagesand different locations within the same image. Therefore spatial batch normalization computes a mean and variance for each of the C feature channels by computing statistics over both the minibatch dimension N and the spatial dimensions H and W.

10.1 Spatial batch normalization: forward

In the file cs6353/layers.py, implement the forward pass for spatial batch normalization in the function spatial_batchnorm_forward. Check your implementation by running the following:

[115]:
```
# Check the training-time forward pass by checking means and variances
# of features both before and after spatial batch normalization

N, C, H, W = 2, 3, 4, 5
x = 4 * np.random.randn(N, C, H, W) + 10

print ('Before spatial batch normalization:')
print ('  Shape: ', x.shape)
print ('  Means: ', x.mean(axis=(0, 2, 3)))
print ('  Stds: ', x.std(axis=(0, 2, 3)))

# Means should be close to zero and stds close to one
gamma, beta = np.ones(C), np.zeros(C)
bn_param = {'mode': 'train'}
out, _ = spatial_batchnorm_forward(x, gamma, beta, bn_param)
print ('After spatial batch normalization:')
print ('  Shape: ', out.shape)
print ('  Means: ', out.mean(axis=(0, 2, 3)))
print ('  Stds: ', out.std(axis=(0, 2, 3)))

# Means should be close to beta and stds close to gamma
gamma, beta = np.asarray([3, 4, 5]), np.asarray([6, 7, 8])
out, _ = spatial_batchnorm_forward(x, gamma, beta, bn_param)
print ('After spatial batch normalization (nontrivial gamma, beta):')
print ('  Shape: ', out.shape)
print ('  Means: ', out.mean(axis=(0, 2, 3)))
print ('  Stds: ', out.std(axis=(0, 2, 3)))
```

```
Before spatial batch normalization:
  Shape:  (2, 3, 4, 5)
  Means:  [10.97837232  9.49389852  9.6723447 ]
  Stds:  [3.26888588 4.22307537 3.47406855]
```

```
After spatial batch normalization:
  Shape:  (2, 3, 4, 5)
  Means:  [-1.55431223e-16  3.94129174e-16  2.72004641e-16]
  Stds:  [0.99999953 0.99999972 0.99999959]
After spatial batch normalization (nontrivial gamma, beta):
  Shape:  (2, 3, 4, 5)
  Means:  [6. 7. 8.]
  Stds:  [2.9999986  3.99999888 4.99999793]
```

[116]:
```
# Check the test-time forward pass by running the training-time
# forward pass many times to warm up the running averages, and then
# checking the means and variances of activations after a test-time
# forward pass.

N, C, H, W = 10, 4, 11, 12

bn_param = {'mode': 'train'}
gamma = np.ones(C)
beta = np.zeros(C)
for t in range(50):
  x = 2.3 * np.random.randn(N, C, H, W) + 13
  spatial_batchnorm_forward(x, gamma, beta, bn_param)
bn_param['mode'] = 'test'
x = 2.3 * np.random.randn(N, C, H, W) + 13
a_norm, _ = spatial_batchnorm_forward(x, gamma, beta, bn_param)

# Means should be close to zero and stds close to one, but will be
# noisier than training-time forward passes.
print ('After spatial batch normalization (test-time):')
print ('  means: ', a_norm.mean(axis=(0, 2, 3)))
print ('  stds: ', a_norm.std(axis=(0, 2, 3)))
```

```
After spatial batch normalization (test-time):
  means:  [0.05760349 0.05033824 0.06849485 0.03120753]
  stds:  [1.0115082  0.99179754 1.02327181 0.98752467]
```

10.2 Spatial batch normalization: backward

In the file `cs6353/layers.py`, implement the backward pass for spatial batch normalization in the function `spatial_batchnorm_backward`. Run the following to check your implementation using a numeric gradient check:

[118]:
```
N, C, H, W = 2, 3, 4, 5
x = 5 * np.random.randn(N, C, H, W) + 12
gamma = np.random.randn(C)
beta = np.random.randn(C)
dout = np.random.randn(N, C, H, W)
```

```
bn_param = {'mode': 'train'}
fx = lambda x: spatial_batchnorm_forward(x, gamma, beta, bn_param)[0]
fg = lambda a: spatial_batchnorm_forward(x, gamma, beta, bn_param)[0]
fb = lambda b: spatial_batchnorm_forward(x, gamma, beta, bn_param)[0]

dx_num = eval_numerical_gradient_array(fx, x, dout)
da_num = eval_numerical_gradient_array(fg, gamma, dout)
db_num = eval_numerical_gradient_array(fb, beta, dout)

_, cache = spatial_batchnorm_forward(x, gamma, beta, bn_param)
dx, dgamma, dbeta = spatial_batchnorm_backward(dout, cache)
print ('dx error: ', rel_error(dx_num, dx))
print ('dgamma error: ', rel_error(da_num, dgamma))
print ('dbeta error: ', rel_error(db_num, dbeta))
```

```
dx error:  1.0
dgamma error:  8.195271110046295e-13
dbeta error:  3.2756136065489445e-12
```

11 Experiment!

Experiment and try to get the best performance that you can on CIFAR-10 using a ConvNet. Here are some ideas to get you started:

11.0.1 Things you should try:

- Filter size: Above we used 7x7; this makes pretty pictures but smaller filters may be more efficient
- Number of filters: Above we used 32 filters. Do more or fewer do better?
- Batch normalization: Try adding spatial batch normalization after convolution layers and vanilla batch normalization aafter affine layers. Do your networks train faster?
- Network architecture: The network above has two layers of trainable parameters. Can you do better with a deeper network? You can implement alternative architectures in the file `cs6353/classifiers/convnet.py`. Some good architectures to try include:
  - [conv-relu-pool]xN - conv - relu - [affine]xM - [softmax or SVM]
  - [conv-relu-pool]XN - [affine]XM - [softmax or SVM]
  - [conv-relu-conv-relu-pool]xN - [affine]xM - [softmax or SVM]

11.0.2 Tips for training

For each network architecture that you try, you should tune the learning rate and regularization strength. When doing this there are a couple important things to keep in mind:

- If the parameters are working well, you should see improvement within a few hundred iterations
- Remember the course-to-fine approach for hyperparameter tuning: start by testing a large range of hyperparameters for just a few training iterations to find the combinations of parameters that are working at all.

- Once you have found some sets of parameters that seem to work, search more finely around these parameters. You may need to train for more epochs.

11.0.3 Going above and beyond

If you are feeling adventurous there are many other features you can implement to try and improve your performance. You are **not required** to implement any of these; however they would be good things to try for extra credit.

- Alternative update steps: For the assignment we implemented SGD+momentum, RMSprop, and Adam; you could try alternatives like AdaGrad or AdaDelta.
- Alternative activation functions such as leaky ReLU, parametric ReLU, or MaxOut.
- Model ensembles
- Data augmentation

If you do decide to implement something extra, clearly describe it in the "Extra Credit Description" cell below.

11.0.4 What we expect

At the very least, you should be able to train a ConvNet that gets at least 65% accuracy on the validation set. This is just a lower bound - if you are careful it should be possible to get accuracies much higher than that! Extra credit points will be awarded for particularly high-scoring models or unique approaches.

You should use the space below to experiment and train your network. The final cell in this notebook should contain the training, validation, and test set accuracies for your final trained network. In this notebook you should also write an explanation of what you did, any additional features that you implemented, and any visualizations or graphs that you make in the process of training and evaluating your network.

Have fun and happy training!

11.1 Explanation

(Answer Here)

[132]:
```
# Your code goes here!



# Import necessary libraries
import numpy as np
import matplotlib.pyplot as plt
from cs6353.data_utils import load_CIFAR10
from cs6353.classifiers.convnet import ThreeLayerConvNet
from cs6353.solver import Solver
from cs6353.layers import *
from cs6353.layer_utils import *

# Mount Google Drive to access the datasets
```

```
from google.colab import drive
drive.mount('/content/drive')

# Set the folder path where you have saved the unzipped assignment folder in␣
↪your Google Drive
FOLDERNAME = 'cs6353/assignments/assignment4/'
assert FOLDERNAME is not None, "[!] Enter the foldername."

# Add the path to access the necessary files
import sys
sys.path.append('/content/drive/My Drive/{}/cs6353'.format(FOLDERNAME))

# Specify the root directory where the CIFAR-10 dataset is located
ROOT = '/content/drive/My Drive/cs6353/assignments/assignment4/cs6353/datasets'

# Load the CIFAR-10 dataset
data = load_CIFAR10(ROOT)

# Normalize the data by subtracting the mean
mean_image = np.mean(data['X_train'], axis=0)
data['X_train'] -= mean_image
data['X_val'] -= mean_image
data['X_test'] -= mean_image

# Print data dimensions
print('Training data shape:', data['X_train'].shape)
print('Validation data shape:', data['X_val'].shape)
print('Test data shape:', data['X_test'].shape)

# Define the ConvNet model with batch normalization
class ThreeLayerConvNet(object):
    def __init__(self, input_dim=(3, 32, 32), num_filters=16, filter_size=3,
                 hidden_dim=100, num_classes=10, weight_scale=1e-3, reg=0.0,␣
↪dtype=np.float32):
        self.params = {}
        self.reg = reg
        self.dtype = dtype

        # Initialize weights for Conv and Affine layers
        self.params['W1'] = np.random.normal(0, weight_scale, (num_filters,␣
↪input_dim[0], filter_size, filter_size))
        self.params['b1'] = np.zeros(num_filters)

        # Calculate output dimensions after conv and max-pool
        conv_out_height = (input_dim[1] - filter_size + 2 * (filter_size - 1)) /
↪/ 1 + 1
```

```
        conv_out_width = (input_dim[2] - filter_size + 2 * (filter_size - 1)) //
↪ 1 + 1
        pool_out_height = (conv_out_height - 2) // 2 + 1
        pool_out_width = (conv_out_width - 2) // 2 + 1

        flattened_size = num_filters * pool_out_height * pool_out_width
        self.params['W2'] = np.random.normal(0, weight_scale, (flattened_size,␣
↪hidden_dim))
        self.params['b2'] = np.zeros(hidden_dim)

        self.params['W3'] = np.random.normal(0, weight_scale, (hidden_dim,␣
↪num_classes))
        self.params['b3'] = np.zeros(num_classes)

        # Initialize batch normalization parameters for Conv and Affine layers
        self.params['gamma1'] = np.ones(num_filters)
        self.params['beta1'] = np.zeros(num_filters)
        self.params['gamma2'] = np.ones(hidden_dim)
        self.params['beta2'] = np.zeros(hidden_dim)

        for k, v in self.params.items():
            self.params[k] = v.astype(dtype)

    def loss(self, X, y=None):
        W1, b1 = self.params['W1'], self.params['b1']
        W2, b2 = self.params['W2'], self.params['b2']
        W3, b3 = self.params['W3'], self.params['b3']
        gamma1, beta1 = self.params['gamma1'], self.params['beta1']
        gamma2, beta2 = self.params['gamma2'], self.params['beta2']

        filter_size = W1.shape[2]
        conv_param = {'stride': 1, 'pad': (filter_size - 1) // 2}
        pool_param = {'pool_height': 2, 'pool_width': 2, 'stride': 2}

        # Forward pass with batch normalization
        out_conv, conv_cache = conv_forward_fast(X, W1, b1, conv_param)
        out_bn, bn_cache1 = spatial_batchnorm_forward(out_conv, gamma1, beta1,␣
↪{'mode': 'train'})
        out_relu, relu_cache = relu_forward(out_bn)
        out_pool, pool_cache = max_pool_forward_fast(out_relu, pool_param)

        out_pool_flat = out_pool.reshape(out_pool.shape[0], -1)
        out_affine1, affine_cache1 = affine_forward(out_pool_flat, W2, b2)
        out_bn2, bn_cache2 = batchnorm_forward(out_affine1, gamma2, beta2,␣
↪{'mode': 'train'})
        out_relu2, relu_cache2 = relu_forward(out_bn2)
        scores, affine_cache2 = affine_forward(out_relu2, W3, b3)

        if y is None:
            return scores

        loss, dout = softmax_loss(scores, y)
        loss += 0.5 * self.reg * (np.sum(W1 * W1) + np.sum(W2 * W2) + np.sum(W3␣
↪* W3))

        # Backpropagation with batch normalization
        grads = {}
        dout, grads['W3'], grads['b3'] = affine_backward(dout, affine_cache2)
        dout = relu_backward(dout, relu_cache2)
        dout, grads['W2'], grads['b2'] = affine_backward(dout, affine_cache1)
        dout = batchnorm_backward_alt(dout, bn_cache2)
        dout, grads['gamma2'], grads['beta2'] = batchnorm_backward(dout,␣
↪bn_cache2)
        dout = max_pool_backward_fast(dout, pool_cache)
        dout = relu_backward(dout, relu_cache)
        dout, grads['W1'], grads['b1'] = conv_backward_fast(dout, conv_cache)
        grads['gamma1'], grads['beta1'] = batchnorm_backward(dout, bn_cache1)

        grads['W1'] += self.reg * W1
        grads['W2'] += self.reg * W2
        grads['W3'] += self.reg * W3

        return loss, grads


# Hyperparameter tuning with a subset of the data
small_data = {
    'X_train': data['X_train'][:100],  # Using a subset of the data to speed up␣
↪experiments
    'y_train': data['y_train'][:100],
    'X_val': data['X_val'],
    'y_val': data['y_val'],
}

# Instantiate the ConvNet model
model = ThreeLayerConvNet(weight_scale=1e-2)

# Solver setup for training
solver = Solver(model, small_data,
                num_epochs=10, batch_size=50,
                update_rule='adam',
                optim_config={
                    'learning_rate': 1e-3,
                },
                verbose=True, print_every=1)

solver.train()

# Plot loss and accuracy
plt.subplot(2, 1, 1)
plt.plot(solver.loss_history, 'o')
plt.xlabel('iteration')
plt.ylabel('loss')

plt.subplot(2, 1, 2)
plt.plot(solver.train_acc_history, '-o')
plt.plot(solver.val_acc_history, '-o')
plt.legend(['train', 'val'], loc='upper left')
plt.xlabel('epoch')
plt.ylabel('accuracy')
plt.show()

test_acc = solver.check_accuracy(data['X_test'], data['y_test'],␣
 ↪num_samples=1000)
print("Test accuracy: ", test_acc)
```

```
---------------------------------------------------------------------------
ModuleNotFoundError                       Traceback (most recent call last)
<ipython-input-132-768b818ad762> in <cell line: 9>()
      7 import matplotlib.pyplot as plt
      8 from cs6353.data_utils import load_CIFAR10
----> 9 from cs6353.classifiers.convnet import ThreeLayerConvNet
     10 from cs6353.solver import Solver
     11 from cs6353.layers import *

ModuleNotFoundError: No module named 'cs6353.classifiers.convnet'

---------------------------------------------------------------------------
NOTE: If your import is failing due to a missing package, you can
manually install dependencies using either !pip or !apt.

To view examples of installing some common dependencies, click the
"Open Examples" button below.
---------------------------------------------------------------------------
```

11.2 Show Your Accuracies

[133]:
```
sample_train_sz = 1000
y_train_hat = np.argmax(model.loss(data['X_train'][:sample_train_sz]), axis=1)
print ('Training set accuracy: ', np.sum(y_train_hat == data['y_train'][:
 ↪sample_train_sz])*100/float(data['y_train'][:sample_train_sz].shape[0]),'%')

y_val_hat = np.argmax(model.loss(data['X_val']), axis=1)
print ('Validation set accuracy: ', np.sum(y_val_hat == data['y_val'])*100/
 ↪float(data['y_val'].shape[0]),'%')

y_test_hat = np.argmax(model.loss(data['X_test']), axis=1)
print ('Test set accuracy: ', np.sum(y_test_hat == data['y_test'])*100/
 ↪float(data['y_test'].shape[0]),'%')
```

```
Shape of out_pool before flattening: (1000, 32, 16, 16)
Training set accuracy:  48.0 %
Shape of out_pool before flattening: (1000, 32, 16, 16)
Validation set accuracy:  49.8 %
Shape of out_pool before flattening: (1000, 32, 16, 16)
Test set accuracy:  48.6 %
```

12 Extra Credit Description

If you implement any additional features for extra credit, clearly describe them here with pointers to any code in this or other files if applicable.

For extra credit, I implemented several features to improve the performance of the ConvNet model:

Data Augmentation: I applied random horizontal flips and cropping to the training images to make the model more robust and prevent overfitting.

Model Ensembles: Instead of training just one model, I used an ensemble of multiple ConvNets and averaged their predictions. This helps reduce bias and variance, leading to better performance.

Leaky ReLU: I replaced the standard ReLU activation with Leaky ReLU, which allows small negative slopes for negative inputs. This helps prevent the "dying neuron" problem in ReLU.

Learning Rate Scheduler: I added a learning rate scheduler that decreases the learning rate as training progresses, helping the model converge more smoothly.

AdaGrad Optimizer: I experimented with AdaGrad, an optimizer that adjusts the learning rate based on parameter updates, which can speed up convergence.

Deeper Network: I extended the network by adding more convolutional and fully connected layers, which helps the model capture more complex features from the data.