```
This code was originally written for CS 231n at Stanford University
(cs231n.stanford.edu). It has been modified in various areas for use in the
ECE 239AS class at UCLA. This includes the descriptions of what code to
implement as well as some slight potential changes in variable names to be
consistent with class nomenclature. We thank Justin Johnson & Serena Yeung for
permission to use this code. To see the original version, please visit
cs231n.stanford.edu.
\# {affine - [batch norm] - relu - [dropout]} \times (L - 1) - affine - softmax
def affine batch relu forward(x, w, b, gamma, beta, bn param):
  Convenience layer that performs an affine transform followed by batchnorm followed by a ReLU
  Inputs:
  - x: Input to the affine layer
  - w, b: Weights for the affine layer
 Returns a tuple of:
  - out: Output from the ReLU
  - cache: Object to give to the backward pass
  x1, fc cache = affine forward(x, w, b)
  x2, bn cache = batchnorm forward(x1, gamma, beta, bn param)
  out, relu cache = relu forward(x2)
  cache = (fc cache, bn cache, relu cache)
  return out, cache
def affine batch relu backward(dout, cache):
  Backward pass for the affine-batchnorm-relu convenience layer
  fc_cache, bn_cache, relu_cache = cache
  da = relu backward(dout, relu_cache)
  dx bn, dgamma, dbeta = batchnorm backward(da, bn cache)
  # print(len(db norm))
  dx, dw, db = affine backward(dx bn, fc cache)
  return dx, dw, db, dgamma, dbeta
def affine relu forward(x, w, b):
  Convenience layer that performs an affine transform followed by a ReLU
  Inputs:
  - x: Input to the affine layer
  - w, b: Weights for the affine layer
 Returns a tuple of:
  - out: Output from the ReLU
  - cache: Object to give to the backward pass
  a, fc_cache = affine_forward(x, w, b)
  out, relu cache = relu forward(a)
  cache = (fc cache, relu cache)
  return out, cache
def affine relu backward(dout, cache):
  Backward pass for the affine-relu convenience layer
```

from .layers import *

fc_cache, relu_cache = cache
da = relu_backward(dout, relu_cache)
dx, dw, db = affine_backward(da, fc_cache)
return dx, dw, db