In [2]:
```
#upload tiny imagenet folder into jupyter project

#import zipfile as zf
#files = zf.ZipFile("tiny-imagenet-200.zip", 'r')
#files.extractall()
#files.close
```

In [3]:
```
%matplotlib inline
import matplotlib.pyplot as plt
import numpy as np
import torch
import torchvision
import torchvision.datasets as datasets
import torch.utils.data as data
from torchvision.utils import make_grid
import torchvision.transforms as transforms
import torch.nn as nn
import torch.nn.functional as F
import torch.optim as optim

import os

import vgg
import resnet
import googlenet
import alexnet
```

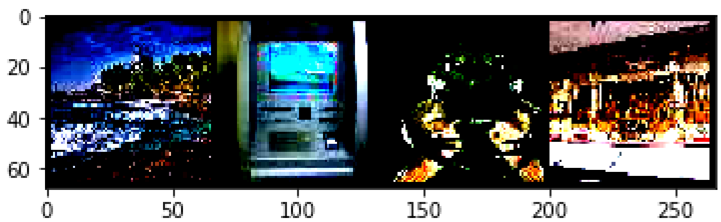
In [4]:
```
# If there are GPUs, choose the first one for computing. Otherwise use CPU.
device = torch.device("cuda:0" if torch.cuda.is_available() else "cpu")
print(device)
# If 'cuda:0' is printed, it means GPU is available.
```

```
cuda:0
```

In [5]:
```
data_transforms = {
    'train': transforms.Compose([
        transforms.RandomRotation(20),
        transforms.RandomHorizontalFlip(0.5),
        transforms.ToTensor(),
        transforms.Normalize([0.4802, 0.4481, 0.3975], [0.2302, 0.2265, 0.2262])
    ]),
    'val': transforms.Compose([
        transforms.ToTensor(),
        transforms.Normalize([0.4802, 0.4481, 0.3975], [0.2302, 0.2265, 0.2262])
    ]),
    'test': transforms.Compose([
        transforms.ToTensor(),
        transforms.Normalize([0.4802, 0.4481, 0.3975], [0.2302, 0.2265, 0.2262])
    ])
}

data_dir = 'tiny-imagenet-200/'

num_workers = {
    'train' : 100,
    'val'   : 0,
    'test'  : 0
}

image_datasets = {x: datasets.ImageFolder(os.path.join(data_dir, x),
                                          data_transforms[x])
                  for x in ['train', 'val','test']}
dataloaders = {x: torch.utils.data.DataLoader(image_datasets[x], batch_size=100,
                                             shuffle=True,num_workers=num_worker
              for x in ['train', 'val', 'test']}
dataset_sizes = {x: len(image_datasets[x]) for x in ['train', 'val', 'test']}
```

In [8]:
```
def imshow(img):
    img = img.numpy().transpose((1, 2, 0))
    img = np.clip(img, 0, 1)
    plt.imshow(img)

images, labels = next(iter(dataloaders['val']))
print(labels)
grid = make_grid(images[:4], nrow=4)
imshow(grid)
```

```
tensor([0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
        0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
        0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
        0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
        0, 0, 0, 0])
```



In [22]:
```
net = vgg.vgg16()   # Create the network instance.
net.to(device)
```

Out[22]:
```
VGG(
  (features): Sequential(
    (0): Conv2d(3, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (1): ReLU(inplace=True)
    (2): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (3): ReLU(inplace=True)
    (4): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=Fa
lse)
    (5): Conv2d(64, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (6): ReLU(inplace=True)
    (7): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (8): ReLU(inplace=True)
    (9): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=Fa
lse)
    (10): Conv2d(128, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (11): ReLU(inplace=True)
    (12): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (13): ReLU(inplace=True)
    (14): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (15): ReLU(inplace=True)
    (16): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=F
alse)
    (17): Conv2d(256, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (18): ReLU(inplace=True)
    (19): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (20): ReLU(inplace=True)
    (21): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (22): ReLU(inplace=True)
    (23): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=F
alse)
    (24): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (25): ReLU(inplace=True)
    (26): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (27): ReLU(inplace=True)
    (28): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (29): ReLU(inplace=True)
    (30): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=F
alse)
  )
  (avgpool): AdaptiveAvgPool2d(output_size=(7, 7))
  (classifier): Sequential(
    (0): Linear(in_features=25088, out_features=4096, bias=True)
    (1): ReLU(inplace=True)
    (2): Dropout(p=0.5, inplace=False)
    (3): Linear(in_features=4096, out_features=4096, bias=True)
    (4): ReLU(inplace=True)
    (5): Dropout(p=0.5, inplace=False)
    (6): Linear(in_features=4096, out_features=1000, bias=True)
  )
)
```

In [11]:
```
# We use cross-entropy as loss function.
loss_func = nn.CrossEntropyLoss()
# We use stochastic gradient descent (SGD) as optimizer.
opt = optim.SGD(net.parameters(), lr=0.001, momentum=0.9)
```

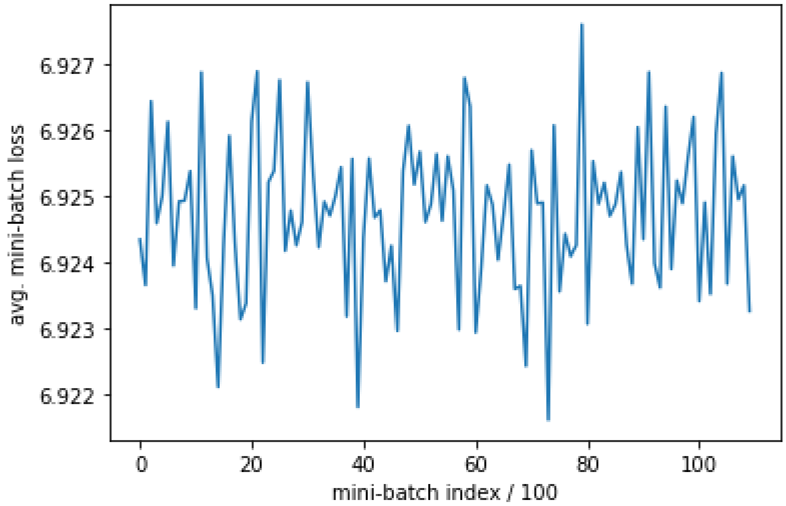
In [23]:
```python
avg_losses = []   # Avg. losses.
epochs = 11       # Total epochs.
print_freq = 100  # Print frequency.

for epoch in range(epochs):  # Loop over the dataset multiple times.
    running_loss = 0.0       # Initialize running loss.
    for i, data in enumerate(dataloaders['train'], 0):

        net.train()

        # Get the inputs
        inputs, labels = data

        # Move the inputs to the specified device.
        inputs, labels = inputs.to(device), labels.to(device)

        # Zero the parameter gradients.
        opt.zero_grad()

        # Forward step.
        outputs = net(inputs)
        loss = loss_func(outputs, labels)

        # Backward step.
        loss.backward()

        # Optimization step (update the parameters).
        opt.step()

        # Print statistics.
        running_loss += loss.item()
        if i % print_freq == print_freq - 1: # Print every several mini-batches.
            avg_loss = running_loss / print_freq
            print('[epoch: {}, i: {:5d}] avg mini-batch loss: {:.3f}'.format(
                epoch, i, avg_loss))
            avg_losses.append(avg_loss)
            running_loss = 0.0

print('Finished Training.')
```

```
[epoch: 0, i:    99] avg mini-batch loss: 6.924
[epoch: 0, i:   199] avg mini-batch loss: 6.924
[epoch: 0, i:   299] avg mini-batch loss: 6.926
[epoch: 0, i:   399] avg mini-batch loss: 6.925
[epoch: 0, i:   499] avg mini-batch loss: 6.925
[epoch: 0, i:   599] avg mini-batch loss: 6.926
[epoch: 0, i:   699] avg mini-batch loss: 6.924
[epoch: 0, i:   799] avg mini-batch loss: 6.925
[epoch: 0, i:   899] avg mini-batch loss: 6.925
[epoch: 0, i:   999] avg mini-batch loss: 6.925
[epoch: 1, i:    99] avg mini-batch loss: 6.923
[epoch: 1, i:   199] avg mini-batch loss: 6.927
[epoch: 1, i:   299] avg mini-batch loss: 6.924
[epoch: 1, i:   399] avg mini-batch loss: 6.923
[epoch: 1, i:   499] avg mini-batch loss: 6.922
[epoch: 1, i:   599] avg mini-batch loss: 6.924
```

```
[epoch: 1, i:   699] avg mini-batch loss: 6.926
[epoch: 1, i:   799] avg mini-batch loss: 6.924
[epoch: 1, i:   899] avg mini-batch loss: 6.923
[epoch: 1, i:   999] avg mini-batch loss: 6.923
[epoch: 2, i:    99] avg mini-batch loss: 6.926
[epoch: 2, i:   199] avg mini-batch loss: 6.927
[epoch: 2, i:   299] avg mini-batch loss: 6.922
[epoch: 2, i:   399] avg mini-batch loss: 6.925
[epoch: 2, i:   499] avg mini-batch loss: 6.925
[epoch: 2, i:   599] avg mini-batch loss: 6.927
[epoch: 2, i:   699] avg mini-batch loss: 6.924
[epoch: 2, i:   799] avg mini-batch loss: 6.925
[epoch: 2, i:   899] avg mini-batch loss: 6.924
[epoch: 2, i:   999] avg mini-batch loss: 6.925
[epoch: 3, i:    99] avg mini-batch loss: 6.927
[epoch: 3, i:   199] avg mini-batch loss: 6.925
[epoch: 3, i:   299] avg mini-batch loss: 6.924
[epoch: 3, i:   399] avg mini-batch loss: 6.925
[epoch: 3, i:   499] avg mini-batch loss: 6.925
[epoch: 3, i:   599] avg mini-batch loss: 6.925
[epoch: 3, i:   699] avg mini-batch loss: 6.925
[epoch: 3, i:   799] avg mini-batch loss: 6.923
[epoch: 3, i:   899] avg mini-batch loss: 6.926
[epoch: 3, i:   999] avg mini-batch loss: 6.922
[epoch: 4, i:    99] avg mini-batch loss: 6.924
[epoch: 4, i:   199] avg mini-batch loss: 6.926
[epoch: 4, i:   299] avg mini-batch loss: 6.925
[epoch: 4, i:   399] avg mini-batch loss: 6.925
[epoch: 4, i:   499] avg mini-batch loss: 6.924
[epoch: 4, i:   599] avg mini-batch loss: 6.924
[epoch: 4, i:   699] avg mini-batch loss: 6.923
[epoch: 4, i:   799] avg mini-batch loss: 6.925
[epoch: 4, i:   899] avg mini-batch loss: 6.926
[epoch: 4, i:   999] avg mini-batch loss: 6.925
[epoch: 5, i:    99] avg mini-batch loss: 6.926
[epoch: 5, i:   199] avg mini-batch loss: 6.925
[epoch: 5, i:   299] avg mini-batch loss: 6.925
[epoch: 5, i:   399] avg mini-batch loss: 6.926
[epoch: 5, i:   499] avg mini-batch loss: 6.925
[epoch: 5, i:   599] avg mini-batch loss: 6.926
[epoch: 5, i:   699] avg mini-batch loss: 6.925
[epoch: 5, i:   799] avg mini-batch loss: 6.923
[epoch: 5, i:   899] avg mini-batch loss: 6.927
[epoch: 5, i:   999] avg mini-batch loss: 6.926
[epoch: 6, i:    99] avg mini-batch loss: 6.923
[epoch: 6, i:   199] avg mini-batch loss: 6.924
[epoch: 6, i:   299] avg mini-batch loss: 6.925
[epoch: 6, i:   399] avg mini-batch loss: 6.925
[epoch: 6, i:   499] avg mini-batch loss: 6.924
[epoch: 6, i:   599] avg mini-batch loss: 6.925
[epoch: 6, i:   699] avg mini-batch loss: 6.925
[epoch: 6, i:   799] avg mini-batch loss: 6.924
[epoch: 6, i:   899] avg mini-batch loss: 6.924
[epoch: 6, i:   999] avg mini-batch loss: 6.922
[epoch: 7, i:    99] avg mini-batch loss: 6.926
[epoch: 7, i:   199] avg mini-batch loss: 6.925
[epoch: 7, i:   299] avg mini-batch loss: 6.925
```

```
[epoch: 7, i:   399] avg mini-batch loss: 6.922
[epoch: 7, i:   499] avg mini-batch loss: 6.926
[epoch: 7, i:   599] avg mini-batch loss: 6.924
[epoch: 7, i:   699] avg mini-batch loss: 6.924
[epoch: 7, i:   799] avg mini-batch loss: 6.924
[epoch: 7, i:   899] avg mini-batch loss: 6.924
[epoch: 7, i:   999] avg mini-batch loss: 6.928
[epoch: 8, i:    99] avg mini-batch loss: 6.923
[epoch: 8, i:   199] avg mini-batch loss: 6.926
[epoch: 8, i:   299] avg mini-batch loss: 6.925
[epoch: 8, i:   399] avg mini-batch loss: 6.925
[epoch: 8, i:   499] avg mini-batch loss: 6.925
[epoch: 8, i:   599] avg mini-batch loss: 6.925
[epoch: 8, i:   699] avg mini-batch loss: 6.925
[epoch: 8, i:   799] avg mini-batch loss: 6.924
[epoch: 8, i:   899] avg mini-batch loss: 6.924
[epoch: 8, i:   999] avg mini-batch loss: 6.926
[epoch: 9, i:    99] avg mini-batch loss: 6.924
[epoch: 9, i:   199] avg mini-batch loss: 6.927
[epoch: 9, i:   299] avg mini-batch loss: 6.924
[epoch: 9, i:   399] avg mini-batch loss: 6.924
[epoch: 9, i:   499] avg mini-batch loss: 6.926
[epoch: 9, i:   599] avg mini-batch loss: 6.924
[epoch: 9, i:   699] avg mini-batch loss: 6.925
[epoch: 9, i:   799] avg mini-batch loss: 6.925
[epoch: 9, i:   899] avg mini-batch loss: 6.926
[epoch: 9, i:   999] avg mini-batch loss: 6.926
[epoch: 10, i:    99] avg mini-batch loss: 6.923
[epoch: 10, i:   199] avg mini-batch loss: 6.925
[epoch: 10, i:   299] avg mini-batch loss: 6.924
[epoch: 10, i:   399] avg mini-batch loss: 6.926
[epoch: 10, i:   499] avg mini-batch loss: 6.927
[epoch: 10, i:   599] avg mini-batch loss: 6.924
[epoch: 10, i:   699] avg mini-batch loss: 6.926
[epoch: 10, i:   799] avg mini-batch loss: 6.925
[epoch: 10, i:   899] avg mini-batch loss: 6.925
[epoch: 10, i:   999] avg mini-batch loss: 6.923
Finished Training.
```

In [24]:
```
plt.plot(avg_losses)
plt.xlabel('mini-batch index / {}'.format(print_freq))
plt.ylabel('avg. mini-batch loss')
plt.show()
```




In [25]:
```
# Get test accuracy.
correct = 0
total = 0
with torch.no_grad():
    for i, data in enumerate(dataloaders['test']):
        net.eval()
        images, labels = data
        images, labels = images.to(device), labels.to(device)
        outputs = net(images)
        _, predicted = torch.max(outputs.data, 1)
        total += labels.size(0)
        correct += (predicted == labels).sum().item()

print('Accuracy of the network on the 10000 test images: %d %%' % (
    100 * correct / total))
```

```
Accuracy of the network on the 10000 test images: 23 %
```

In [ ]: