VGG_network

```
root@autodl-container-b197439d52-cc536be3:~/autodl-tmp# python VGG_network.py
Sequential output shape:         torch.Size([1, 16, 112, 112])
Sequential output shape:         torch.Size([1, 32, 56, 56])
Sequential output shape:         torch.Size([1, 64, 28, 28])
Sequential output shape:         torch.Size([1, 128, 14, 14])
Sequential output shape:         torch.Size([1, 128, 7, 7])
Flatten output shape:    torch.Size([1, 6272])
Linear output shape:     torch.Size([1, 4096])
ReLU output shape:       torch.Size([1, 4096])
Dropout output shape:    torch.Size([1, 4096])
Linear output shape:     torch.Size([1, 4096])
ReLU output shape:       torch.Size([1, 4096])
Dropout output shape:    torch.Size([1, 4096])
Linear output shape:     torch.Size([1, 10])
Downloading http://fashion-mnist.s3-website.eu-central-1.amazonaws.com/train-
images-idx3-ubyte.gz
Downloading http://fashion-mnist.s3-website.eu-central-1.amazonaws.com/train-
images-idx3-ubyte.gz to ./data/FashionMNIST/raw/train-images-idx3-ubyte.gz
100%|
| 26.4M/26.4M [02:36<00:00, 169kB/s]
Extracting ./data/FashionMNIST/raw/train-images-idx3-ubyte.gz to
./data/FashionMNIST/raw

Downloading http://fashion-mnist.s3-website.eu-central-1.amazonaws.com/train-
labels-idx1-ubyte.gz
Downloading http://fashion-mnist.s3-website.eu-central-1.amazonaws.com/train-
labels-idx1-ubyte.gz to ./data/FashionMNIST/raw/train-labels-idx1-ubyte.gz
100%|
| 29.5k/29.5k [00:00<00:00, 173kB/s]
Extracting ./data/FashionMNIST/raw/train-labels-idx1-ubyte.gz to
./data/FashionMNIST/raw

Downloading http://fashion-mnist.s3-website.eu-central-1.amazonaws.com/t10k-
images-idx3-ubyte.gz
Downloading http://fashion-mnist.s3-website.eu-central-1.amazonaws.com/t10k-
images-idx3-ubyte.gz to ./data/FashionMNIST/raw/t10k-images-idx3-ubyte.gz
100%|
| 4.42M/4.42M [00:04<00:00, 965kB/s]
Extracting ./data/FashionMNIST/raw/t10k-images-idx3-ubyte.gz to
./data/FashionMNIST/raw

Downloading http://fashion-mnist.s3-website.eu-central-1.amazonaws.com/t10k-
labels-idx1-ubyte.gz
Downloading http://fashion-mnist.s3-website.eu-central-1.amazonaws.com/t10k-
labels-idx1-ubyte.gz to ./data/FashionMNIST/raw/t10k-labels-idx1-ubyte.gz
100%|
| 5.15k/5.15k [00:00<00:00, 3.35MB/s]
```

```
Extracting ./data/FashionMNIST/raw/t10k-labels-idx1-ubyte.gz to
./data/FashionMNIST/raw

Training on cuda
Epoch 1: train loss 2.3028, train acc 0.099, test acc 0.100, time 28.5 sec
Epoch 2: train loss 2.3028, train acc 0.099, test acc 0.100, time 26.7 sec
Epoch 3: train loss 2.3027, train acc 0.098, test acc 0.100, time 29.8 sec
Epoch 4: train loss 2.3027, train acc 0.100, test acc 0.100, time 28.0 sec
Epoch 5: train loss 2.3026, train acc 0.100, test acc 0.100, time 28.5 sec
Epoch 6: train loss 2.2769, train acc 0.175, test acc 0.399, time 33.7 sec
Epoch 7: train loss 0.8582, train acc 0.672, test acc 0.813, time 28.2 sec
Epoch 8: train loss 0.4424, train acc 0.834, test acc 0.849, time 26.6 sec
Epoch 9: train loss 0.3487, train acc 0.873, test acc 0.879, time 28.6 sec
Epoch 10: train loss 0.3022, train acc 0.889, test acc 0.889, time 28.2 sec
root@autodl-container-b197439d52-cc536be3:~/autodl-tmp#
```

(降低学习率 + 加入 momentum）：

- 原来学习率 0.05 容易震荡，模型学不动
- 现在 `lr=0.01` + `momentum=0.9`

```
root@autodl-container-b197439d52-cc536be3:~/autodl-tmp# python VGG_network.py
Sequential output shape:         torch.Size([1, 16, 112, 112])
Sequential output shape:         torch.Size([1, 32, 56, 56])
Sequential output shape:         torch.Size([1, 64, 28, 28])
Sequential output shape:         torch.Size([1, 128, 14, 14])
Sequential output shape:         torch.Size([1, 128, 7, 7])
Flatten output shape:    torch.Size([1, 6272])
Linear output shape:     torch.Size([1, 4096])
ReLU output shape:       torch.Size([1, 4096])
Dropout output shape:    torch.Size([1, 4096])
Linear output shape:     torch.Size([1, 4096])
ReLU output shape:       torch.Size([1, 4096])
Dropout output shape:    torch.Size([1, 4096])
Linear output shape:     torch.Size([1, 10])
Training on cuda
Epoch 1: train loss 2.3029, train acc 0.099, test acc 0.100, time 28.0 sec
Epoch 2: train loss 2.3028, train acc 0.100, test acc 0.100, time 28.3 sec
Epoch 3: train loss 1.4855, train acc 0.449, test acc 0.799, time 28.7 sec
Epoch 4: train loss 0.4449, train acc 0.836, test acc 0.866, time 28.7 sec
Epoch 5: train loss 0.3274, train acc 0.880, test acc 0.888, time 29.0 sec
Epoch 6: train loss 0.2765, train acc 0.897, test acc 0.902, time 28.0 sec
Epoch 7: train loss 0.2409, train acc 0.911, test acc 0.907, time 28.2 sec
Epoch 8: train loss 0.2178, train acc 0.920, test acc 0.912, time 26.9 sec
Epoch 9: train loss 0.1954, train acc 0.928, test acc 0.919, time 28.2 sec
Epoch 10: train loss 0.1777, train acc 0.933, test acc 0.915, time 27.9 sec
```