In [ ]:
```
#code : No code
```

Discussion: 이번 단원에서는 CNN, Convolutional Neural Network에 대해 소개하고, 기존까지 배웠던 MLP방식과 이를 비교하여 이미지 데이터 처리에서 CNN이 왜 더 적합한지, 그리고 CNN이 이미지를 어떻게 처리하는지에 대한 설명을 진행하였다. 지금까지 배웠던 MLP방식은 Table Data를 다루기 적합했지만, 고해상도의 이미지와 같이 고차원적인 Data에 대해서는 사용하기 적합하지 않다는 단점이 있었다. 아무리 hidden layer의 차원을 줄이려 해도, 엄청난 양의 parameter를 학습해야 하므로 현재의 컴퓨팅 자원으로는 매우 비효율적인 방법이다. 다행히, 이미지가 가지고 있는 구조적 특성 덕에, 다른 방법을 생각해 낼 수 있었는데 그것이 바로 합성곱을 이용한 CNN방식이다.

CNN은 이미지가 가진 구조적 특성 - translation invariance / locality를 활용하여 효율적으로 학습할 수 있도록 한다. (사실 나의 배경 지식 및 현재 상태의 이해력의 한계로 완전히 이해하지 못하지만, 내가 이해한 수준에서 설명하면) Translation invariance는, 이미지 내의 물체가 다른 위치에 위치하더라도, 동일한 물체는 동일하게 인식되는 특성이다. 예를 들어, 다양한 사진 속 "비행기"를 찾는 상황에서, 비행기는 이미지 속 상단에 위치할 수도, 하단에 위치할 수도, 좌/우에 위치할 수도 있지만, 비행기의 특징 자체는 동일하다는 것이다. 이는 CNN이 동일한 patch의 이미지를 다른 위치에서 발견해도 동일하게 반응할 수 있도록 한다. Locality는, 이미지 내에서 정보를 처리할 때 국소적인 영역에 집중하는 것을 뜻한다. 처음 단계에서 이미지의 작은 부분만을 분석하고(local filter 영역), 이후 단계에서 정보를 종합하여 전체 이미지를 분석하는 것이다.

그렇다면 CNN은 어떻게 이러한 것을 가능하게 할까? -> 바로 수업시간에도 배웠던 convolution, 합성곱 연산을 통해 파라미터 수를 줄이며 학습 능력을 유지하도록 도와주었다. 합성곱이란 한 함수를 뒤집고, 다른 함수와 겹쳐서 측정하는 것이다. 수학적으로는, (f*g)(x) = integral{f(z)g(x-z)dz}으로 될 수 있고, 뭐 이제 실제 컴퓨터에서는 이산적이므로 시그마로 바꾸거나 2차원 텐서를 위해 변수가 여러개인 함수를 한다던가 그런 식이다. 또한 CNN은 channel이란 것을 활용하는데, 이는 단순한 흑백 픽셀 값을 넘어 RGB 값도 인식에 반영될 수 있게 하는 것이다.

이번 단원에는 Code가 없으므로 exercise 문제들 중 하나를 진행해 보려고 한다.

6. 합성곱 f*g가 symmetric, 즉 f*g = g*f임을 증명. 수식 기호가 없어 일단 이렇게 진행한다. (f*g)(t) = inte{ f(t) g(t - r) dr} u = t - r로 치환하면, du = -dr (f*g)(t) = inte{ f(t-u) g(u) -du } = inte { g(u) f(t-u) du } = g*f(t) 간단한 치환적분을 통해 증명하였다.

In [1]:
```python
import torch
from torch import nn
from d2l import torch as d2l
```

In [8]:
```python
def corr2d(X,K):
    h,w=K.shape
    Y = torch.zeros((X.shape[0]-h+1, X.shape[1]-w+1))
    for i in range(Y.shape[0]):
        for j in range(Y.shape[1]):
            Y[i,j] = (X[i:i+h,j:j+w]*K).sum()
    return Y
```

In [9]:
```python
X = torch.tensor([[0.0,1.0,2.0],[3.0,4.0,5.0],[6.0,7.0,8.0]])
K = torch.tensor([[0.0,1.0],[2.0,3.0]])
corr2d(X,K)
```

Out[9]:
```
tensor([[19., 25.],
        [37., 43.]])
```

In [10]:
```python
class Conv2D(nn.Module):
    def __init__(self, kernel_size):
        super().__init__()
        self.weight = nn.Parameter(torch.rand(kernel_size))
        self.bias = nn.Parameter(torch.zeros(1))

    def forward(self, x):
        return corr2d(x, self.weight) + self.bias
```

In [11]:
```python
X = torch.ones((6, 8))
X[:, 2:6] = 0
X
```

Out[11]:
```
tensor([[1., 1., 0., 0., 0., 0., 1., 1.],
        [1., 1., 0., 0., 0., 0., 1., 1.],
        [1., 1., 0., 0., 0., 0., 1., 1.],
        [1., 1., 0., 0., 0., 0., 1., 1.],
        [1., 1., 0., 0., 0., 0., 1., 1.],
        [1., 1., 0., 0., 0., 0., 1., 1.]])
```

In [12]:
```python
K = torch.tensor([[1.0, -1.0]])
```

In [14]:
```python
Y = corr2d(X, K)
Y
```

Out[14]:
```
tensor([[ 0.,  1.,  0.,  0.,  0., -1.,  0.],
        [ 0.,  1.,  0.,  0.,  0., -1.,  0.],
        [ 0.,  1.,  0.,  0.,  0., -1.,  0.],
        [ 0.,  1.,  0.,  0.,  0., -1.,  0.],
        [ 0.,  1.,  0.,  0.,  0., -1.,  0.],
        [ 0.,  1.,  0.,  0.,  0., -1.,  0.]])
```

In [15]:
```python
corr2d(X.t(), K)
```

Out[15]:
```
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.]])
```

In [16]:
```python
conv2d = nn.LazyConv2d(1, kernel_size=(1, 2), bias=False)

X = X.reshape((1, 1, 6, 8))
Y = Y.reshape((1, 1, 6, 7))
lr = 3e-2

for i in range(10):
    Y_hat = conv2d(X)
    l = (Y_hat - Y) ** 2
    conv2d.zero_grad()
    l.sum().backward()
    conv2d.weight.data[:] -= lr * conv2d.weight.grad
    if (i + 1) % 2 == 0:
        print(f'epoch {i + 1}, loss {l.sum():.3f}')
```

```
epoch 2, loss 9.297
epoch 4, loss 3.076
epoch 6, loss 1.137
epoch 8, loss 0.445
epoch 10, loss 0.179
```

```
/Library/Frameworks/Python.framework/Versions/3.11/lib/python3.11/site-packages/torch/nn/modules/lazy.py:181: UserWarning: Lazy modules are a new feature under heavy development so changes to the API or functionality can happen at any moment.
  warnings.warn('Lazy modules are a new feature under heavy development '
```

In [17]:
```
conv2d.weight.data.reshape((1, 2))
```

Out[17]:
```
tensor([[ 0.9476, -1.0343]])
```

Discussion: 이번 단원에서는 CNN에서 사용하는 연산인 Cross-correlation Operation, 교차 상관 연산을 배웠다. 입력 텐서와 커널 텐서를 통해 출력 텐서를 생성하는 연산이다. 이때, 커널이라 함은 이미지의 특정 부분에 대한 연산을 수행하여 해당 부분에서 중요한 정보를 추출하는 것을 뜻한다. 합성곱에서는 커널을 뒤집어서 연산을 수행하지만, 이러한 커널은 학습 과정에서 데이터를 통해 자동으로 학습되기 때문에 CNN에서는 이 cross-correlation과 합성곱 연산의 결과가 크게 달라지지 않고 그렇기 때문에 이 분야에서는 거의 비슷한 말로 쓰인다고 한다. 이런 방식으로 이미지 데이터의 특징을 활용해서 다른 방법의 학습을 한다는 것이 인상적이었다.

Excercise: Construct an image X with diagonal edges. <- X를 단순화해서 대각행렬로 생각해보자.

What happens if you apply the kernel K in this section to it?

어떤 K를 말하는지 살짝 헷갈리지만, 수정한 K = [1 -1]을 기준으로 하면, 이 커널은 가로로 인접한 두 픽셀의 차이를 계산하기 때문에, 어느 정도 이미지 경계를 파악 가능하지만 가로+세로를 포함한 대각선의 특성상 완전한 탐지는 힘들 것 같다.

What happens if you transpose X?

X를 행렬로 표현하면, 대각 행렬이고, 대각 행렬에서 X^T = X 이므로 동일할 것이다.

What happens if you transpose K?

K를 전치하면, 세로의 변화를 잘 감지할 것 같다. K^T를 해보면..

In [1]:
```python
import torch
from torch import nn
```

In [2]:
```python
def comp_conv2d(conv2d, X):
    X = X.reshape((1, 1) + X.shape)
    Y = conv2d(X)
    return Y.reshape(Y.shape[2:])

conv2d = nn.LazyConv2d(1, kernel_size=3, padding=1)
X = torch.rand(size=(8, 8))
comp_conv2d(conv2d, X).shape
```

```
/Library/Frameworks/Python.framework/Versions/3.11/lib/python3.11/site-packages/torch/nn/modules/lazy.py:181: UserWarning: Lazy modules are a new feature under heavy development so changes to the API or functionality can happen at any moment.
  warnings.warn('Lazy modules are a new feature under heavy development '
```

Out[2]:
```
torch.Size([8, 8])
```

In [3]:
```python
conv2d = nn.LazyConv2d(1, kernel_size=(5, 3), padding=(2, 1))
comp_conv2d(conv2d, X).shape
```

Out[3]:
```
torch.Size([8, 8])
```

In [4]:
```python
conv2d = nn.LazyConv2d(1, kernel_size=3, padding=1, stride=2)
comp_conv2d(conv2d, X).shape
```

Out[4]:
```
torch.Size([4, 4])
```

In [5]:
```python
conv2d = nn.LazyConv2d(1, kernel_size=(3, 5), padding=(0, 1), stride=(3, 4))
comp_conv2d(conv2d, X).shape
```

Out[5]:
```
torch.Size([2, 2])
```

Discussion: 이 단원에서는 Padding과 Stride의 개념에 대해 설명하고 있다. Padding은 이미지의 경계 픽셀이 잘 사용되지 않는 문제를 해결하기 위해 사용되는 것으로, CNN에서 연속적인 합성곱 연산을 거칠 때 출력 이미지가 점점 작아지면서 경계 근처의 중요한 정보가 손실되는 것을 막는다. 이미지의 가장자리에 0으로 된 픽셀을 추가하는 것이다. Stride는 커널이 이미지 위를 이동하는 간격을 의미한다. 수업 시간에 이것으로 간단한 수식을 계산했던 것으로 어렴풋이 기억에 남는데, 출력 크기를 줄이거나 계산 속도를 향상시키고자 할때 유용하다. 출력 크기 = [입력 - 커널 + (2x패딩)/스트리이드] + 1로, 이때 분수에서의 소숫점은 버리는 것으로 알고 있다.

Exercise: Given the final code example in this section with kernel size (3,5), padding (0,1), and stride (3,4), calculate the output shape to check if it is consistent with the experimental result.

입력은 8x8인 것을 아므로, 한번 공식을 대입해보면.

출력 높이 = (8-3/3)(반내림) + 1 = 1+1 = 2, 출력 너비 (8-5+2/4) + 1 = 1+ 1 = 2, 따라서 2x2로 일치한다. 사실 뭐 컴퓨터가 틀릴 일은 없으니 당연한 결과이긴 하다.

In [1]:
```
import torch
from d2l import torch as d2l
```

In [2]:
```
def corr2d_multi_in(X, K):
    return sum(d2l.corr2d(x, k) for x, k in zip(X, K))
```

In [3]:
```
X = torch.tensor([[[0.0, 1.0, 2.0], [3.0, 4.0, 5.0], [6.0, 7.0, 8.0]],
               [[1.0, 2.0, 3.0], [4.0, 5.0, 6.0], [7.0, 8.0, 9.0]]])
K = torch.tensor([[[0.0, 1.0], [2.0, 3.0]], [[1.0, 2.0], [3.0, 4.0]]])

corr2d_multi_in(X, K)
```

Out[3]:
```
tensor([[ 56.,  72.],
        [104., 120.]])
```

In [4]:
```
def corr2d_multi_in_out(X, K):
    return torch.stack([corr2d_multi_in(X, k) for k in K], 0)
```

In [5]:
```
K = torch.stack((K, K + 1, K + 2), 0)
K.shape
```

Out[5]:
```
torch.Size([3, 2, 2, 2])
```

In [6]:
```
corr2d_multi_in_out(X, K)
```

Out[6]:
```
tensor([[[ 56.,  72.],
         [104., 120.]],

        [[ 76., 100.],
         [148., 172.]],

        [[ 96., 128.],
         [192., 224.]]])
```

In [7]:
```
def corr2d_multi_in_out_1x1(X, K):
    c_i, h, w = X.shape
    c_o = K.shape[0]
    X = X.reshape((c_i, h * w))
    K = K.reshape((c_o, c_i))
    Y = torch.matmul(K, X)
    return Y.reshape((c_o, h, w))
```

In [8]:
```
X = torch.normal(0, 1, (3, 3, 3))
K = torch.normal(0, 1, (2, 3, 1, 1))
Y1 = corr2d_multi_in_out_1x1(X, K)
Y2 = corr2d_multi_in_out(X, K)
assert float(torch.abs(Y1 - Y2).sum()) < 1e-6
```

Discussion: 이번 단원에서는, CNN 중 다중 입력 채널 및 다중 출력 채널을 처리하는 방식에 대해서 배웠다. 이번 단원 전까지는 단일 입력 채널에 대해서 다루었지만, 실제로 이미지에서는 여러 입력 채널이 있기 때문에 이를 처리하여야 한다. 합성곱 연산을 할 때에도 마찬가지로, 입력 데이터의 채널의 수 n에 따라 각 채널마다 별도의 n차원 커널이 필요하다는 것이 중요한 내용이라고 생각했다. 이러한 입력 채널과 마찬가지로, 출력 채널 역시 다중 출력 채널이 필요한데, 이미지에는 한 가지 특징이 있는 것이 아니라 많은 특징을 가지고 있기 때문에 다중 출력 채널을 사용하여 여러 가지 특징을 동시에 학습할 수 있도록 해준다. (예를 들어, edge를 감지하는 채널과 shape를 감지하는 채널) 이렇게 출력 채널을 여러 개로 만들기 위해서는, 각 출력 채널마다 별도의 합성곱 커널이 필요하기 때문에 커널의 수가 많아지는 특징이 있다. 추가적으로, 1x1 convolution이라는 개념도 나왔는데, 이는 각 픽셀에서 channel간의 상호작용을 학습하는 데 사용하여. 각 픽셀의 다양한 특징을 통합하는 기능을 한다고 한다. 각 픽셀의 값을 linear combination하여 새로운 값을 계산하는 방식이라고 한다. 다른 기초 전공들 (자구, 알고)에서 많이 다루었던 개념이기도 한 연산의 cost는 O(input size x output size)라고 한다.

Exercise Assume that we have two convolution kernels of size k1 and k2, respectively (with no nonlinearity in between).

Prove that the result of the operation can be expressed by a single convolution. 합성곱 연산은 결합법칙이 성립하기 때문에, 차례로 입력하는 상황인 (X * k1) * k2와 , 한번에 적용하는 상황인 X * (k1 * k2)가 같을 수 밖에 없다. What is the dimensionality of the equivalent single convolution? (k1 + k2 - 1)^2 (k1의 차원이 k1^2이고, k2의 차원이 k2^2라고 가정) Is the converse true, i.e., can you always decompose a convolution into two smaller ones? 이것은 잘 몰라서 인터넷을 검색해 봤는데, 특정한 경우를 제외하고는 불가능하다고 한다. 왜냐하면 큰 커널의 학습 과정에서 얻어진 복잡한 필터링은 더 작은 커널로 나누었을 때 동일한 필터링 결과를 보장할 수 없기 때문이라고 한다.

In [1]:
```python
import torch
from torch import nn
from d2l import torch as d2l
```

In [2]:
```python
def pool2d(X, pool_size, mode='max'):
    p_h, p_w = pool_size
    Y = torch.zeros((X.shape[0] - p_h + 1, X.shape[1] - p_w + 1))
    for i in range(Y.shape[0]):
        for j in range(Y.shape[1]):
            if mode == 'max':
                Y[i, j] = X[i: i + p_h, j: j + p_w].max()
            elif mode == 'avg':
                Y[i, j] = X[i: i + p_h, j: j + p_w].mean()
    return Y
```

In [3]:
```python
X = torch.tensor([[0.0, 1.0, 2.0], [3.0, 4.0, 5.0], [6.0, 7.0, 8.0]])
pool2d(X, (2, 2))
```

Out[3]:
```
tensor([[4., 5.],
        [7., 8.]])
```

In [4]:
```python
pool2d(X, (2, 2), 'avg')
```

Out[4]:
```
tensor([[2., 3.],
        [5., 6.]])
```

In [5]:
```python
X = torch.arange(16, dtype=torch.float32).reshape((1, 1, 4, 4))
X
```

Out[5]:
```
tensor([[[[ 0.,  1.,  2.,  3.],
          [ 4.,  5.,  6.,  7.],
          [ 8.,  9., 10., 11.],
          [12., 13., 14., 15.]]]])
```

In [6]:
```python
pool2d = nn.MaxPool2d(3)
pool2d(X)
```

Out[6]:
```
tensor([[[[10.]]]])
```

In [7]:
```python
pool2d = nn.MaxPool2d(3, padding=1, stride=2)
pool2d(X)
```

Out[7]:
```
tensor([[[[ 5.,  7.],
          [13., 15.]]]])
```

In [8]:
```python
pool2d = nn.MaxPool2d((2, 3), stride=(2, 3), padding=(0, 1))
pool2d(X)
```

Out[8]:
```
tensor([[[[ 5.,  7.],
          [13., 15.]]]])
```

In [9]:
```python
X = torch.cat((X, X + 1), 1)
X
```

Out[9]:
```
tensor([[[[ 0.,  1.,  2.,  3.],
          [ 4.,  5.,  6.,  7.],
          [ 8.,  9., 10., 11.],
          [12., 13., 14., 15.]],

         [[ 1.,  2.,  3.,  4.],
          [ 5.,  6.,  7.,  8.],
          [ 9., 10., 11., 12.],
          [13., 14., 15., 16.]]]])
```

In [10]:
```python
pool2d = nn.MaxPool2d(3, padding=1, stride=2)
pool2d(X)
```

Out[10]:
```
tensor([[[[ 5.,  7.],
          [13., 15.]],

         [[ 6.,  8.],
          [14., 16.]]]])
```

Discussion: 이번 단원에서는 Pooling층이 어떻게 작동하는 것이고, 왜 필요한 것인지에 대해 설명하였다. 풀링은 공간 해상도와 모델의 위치에 대한 민감도를 줄이는 데 사용되며, 출력 차원을 줄여 계산의 효율을 높이는 역할을 한다. Max pooling과 Avg pooling이 있는데, Max는 말 그대로 그 pooling window 중 max값을, avg는 avg값을 출력하는 느낌이다.

나머지 코드는 이러한 pooling 층에서도 앞 단원의 padding과 stride를 설정할 수 있다는 내용에 관한 것이고, 또한 풀링은 다중 채널을 처리할 때 각 채널별로 독립적으로 작동하므로 합성곱 층과는 달리 채널 간에 값을 합치지 않고 각 채널에 대해 동일한 풀링 연산을 적용한다고 한다.

Exercise: Prove that max-pooling cannot be implemented through a convolution alone.

합성곱은 linear한 커널을 입력에 적용하는 것이다. 입력 값들의 가중 합을 계산하는 방식으로 이루어진다. 그러나, linear한 연산은 아무리 합성을 하고 하고 하고 해봐도 linear한 특성을 벗어날 수 없다. 그러나 max pooling은 max()함수를 사용하는 nonlinear operation이고, 따라서 아무리 합성곱만을 가지고 이러한 저러한 조작을 해도 절대 max pooling을 구현할 수 없다.

In [1]:
```python
import torch
from torch import nn
from d2l import torch as d2l
```

In [2]:
```python
def init_cnn(module):
    """Initialize weights for CNNs."""
    if type(module) == nn.Linear or type(module) == nn.Conv2d:
        nn.init.xavier_uniform_(module.weight)

class LeNet(d2l.Classifier):
    """The LeNet-5 model."""
    def __init__(self, lr=0.1, num_classes=10):
        super().__init__()
        self.save_hyperparameters()
        self.net = nn.Sequential(
            nn.LazyConv2d(6, kernel_size=5, padding=2), nn.Sigmoid(),
            nn.AvgPool2d(kernel_size=2, stride=2),
            nn.LazyConv2d(16, kernel_size=5), nn.Sigmoid(),
            nn.AvgPool2d(kernel_size=2, stride=2),
            nn.Flatten(),
            nn.LazyLinear(120), nn.Sigmoid(),
            nn.LazyLinear(84), nn.Sigmoid(),
            nn.LazyLinear(num_classes))
```

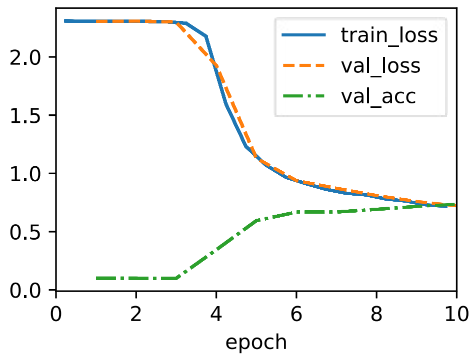
In [3]:
```python
@d2l.add_to_class(d2l.Classifier)
def layer_summary(self, X_shape):
    X = torch.randn(*X_shape)
    for layer in self.net:
        X = layer(X)
        print(layer.__class__.__name__, 'output shape:\t', X.shape)

model = LeNet()
model.layer_summary((1, 1, 28, 28))
```

```
Conv2d output shape:     torch.Size([1, 6, 28, 28])
Sigmoid output shape:    torch.Size([1, 6, 28, 28])
AvgPool2d output shape:  torch.Size([1, 6, 14, 14])
Conv2d output shape:     torch.Size([1, 16, 10, 10])
Sigmoid output shape:    torch.Size([1, 16, 10, 10])
AvgPool2d output shape:  torch.Size([1, 16, 5, 5])
Flatten output shape:    torch.Size([1, 400])
Linear output shape:     torch.Size([1, 120])
Sigmoid output shape:    torch.Size([1, 120])
Linear output shape:     torch.Size([1, 84])
Sigmoid output shape:    torch.Size([1, 84])
Linear output shape:     torch.Size([1, 10])
```

```
/Library/Frameworks/Python.framework/Versions/3.11/lib/python3.11/site-packages/torch/nn/modules/lazy.py:181: UserWarning: Lazy modules are a new feature under heavy development so changes to the API or functionality can happen at any moment.
  warnings.warn('Lazy modules are a new feature under heavy development '
```

In [4]:
```python
trainer = d2l.Trainer(max_epochs=10, num_gpus=1)
data = d2l.FashionMNIST(batch_size=128)
model = LeNet(lr=0.1)
model.apply_init([next(iter(data.get_dataloader(True)))[0]], init_cnn)
trainer.fit(model, data)
```



Discussion: 이번 단원에서는 LeNet이라는 초기의 합성곱 신경망 구조에 대해 설명하고, 실제 구현을 해보는 단원이다. 이번 7-1~7-5의 복습 및 정리 단원 느낌이고, 수업 시간에 LeNet과 관련되어 어렴풋이 들은 기억만 남아 있는데, 이 부분에 대해 실제로 구현해보고 설명을 보며 조금 더 잘 이해할 수 있게 된 것 같다. LeNet은 크게, 두개의 합성층 곱과 pooling 층으로 이루어진 Convolutional Encoder과, 3개의 완전 연결된 층으로 구성되어있다고 한다. 첫 번째 합성곱 층이 6개의 채널을 생성하고, 시그모이드 사용하고, 두 번째 합성곱이 16개 생성하고, 시그모이드 사용하고, lazy linear로 120 -> 84 -> 19으로 해서 10개의 class로 분류한다고 생각하면 된다. Lenet 학습도 진행하였는데, 위의 코드에서는 10epoch동안 학습하였고 손실 함수로는 교차 엔트로피를 사용하고, mini-batch SGD로 최적화하였다.

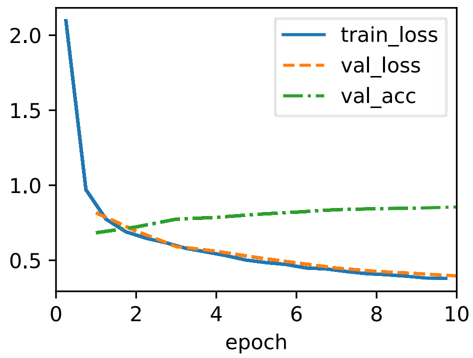
Exercise - Let's modernize LeNet. Implement and test the following changes:

Replace average pooling with max-pooling.

Replace the softmax layer with ReLU.

In [9]:
```
class modernizedLeNet(d2l.Classifier):
    """The LeNet-5 model."""
    def __init__(self, lr=0.1, num_classes=10):
        super().__init__()
        self.save_hyperparameters()
        self.net = nn.Sequential(
            nn.LazyConv2d(6, kernel_size=5, padding=2), nn.Sigmoid(),
            nn.MaxPool2d(kernel_size=2, stride=2),
            nn.LazyConv2d(16, kernel_size=5), nn.Sigmoid(),
            nn.MaxPool2d(kernel_size=2, stride=2),
            nn.Flatten(),
            nn.LazyLinear(120), nn.ReLU(),
            nn.LazyLinear(84), nn.ReLU(),
            nn.LazyLinear(num_classes))
```

In [10]:
```
trainer = d2l.Trainer(max_epochs=10, num_gpus=1)
data = d2l.FashionMNIST(batch_size=128)
model = modernizedLeNet(lr=0.1)
model.apply_init([next(iter(data.get_dataloader(True)))[0]], init_cnn)
trainer.fit(model, data)
```



해본 결과, train, val loss가 훨씬 더 적은 epoch부터 줄어든 결과가 나왔다.

In [1]:
```
!pip install d2l
```

```
Collecting d2l
  Downloading d2l-1.0.3-py3-none-any.whl.metadata (556 bytes)
Collecting jupyter==1.0.0 (from d2l)
  Downloading jupyter-1.0.0-py2.py3-none-any.whl.metadata (995 bytes)
Collecting numpy==1.23.5 (from d2l)
  Downloading numpy-1.23.5-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (2.3 kB)
Collecting matplotlib==3.7.2 (from d2l)
  Downloading matplotlib-3.7.2-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (5.6 kB)
Collecting matplotlib-inline==0.1.6 (from d2l)
  Downloading matplotlib_inline-0.1.6-py3-none-any.whl.metadata (2.8 kB)
Collecting requests==2.31.0 (from d2l)
  Downloading requests-2.31.0-py3-none-any.whl.metadata (4.6 kB)
Collecting pandas==2.0.3 (from d2l)
  Downloading pandas-2.0.3-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (18 kB)
Collecting scipy==1.10.1 (from d2l)
  Downloading scipy-1.10.1-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (58 kB)
     ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 58.9/58.9 kB 2.0 MB/s eta 0:00:00
Requirement already satisfied: notebook in /usr/local/lib/python3.10/dist-packages (from jupyter==1.0.0->d2l) (6
.5.5)
Collecting qtconsole (from jupyter==1.0.0->d2l)
  Downloading qtconsole-5.6.0-py3-none-any.whl.metadata (5.0 kB)
Requirement already satisfied: jupyter-console in /usr/local/lib/python3.10/dist-packages (from jupyter==1.0.0->
d2l) (6.1.0)
Requirement already satisfied: nbconvert in /usr/local/lib/python3.10/dist-packages (from jupyter==1.0.0->d2l) (
6.5.4)
Requirement already satisfied: ipykernel in /usr/local/lib/python3.10/dist-packages (from jupyter==1.0.0->d2l) (
5.5.6)
Requirement already satisfied: ipywidgets in /usr/local/lib/python3.10/dist-packages (from jupyter==1.0.0->d2l)
(7.7.1)
Requirement already satisfied: contourpy>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib==3.7
.2->d2l) (1.3.0)
Requirement already satisfied: cycler>=0.10 in /usr/local/lib/python3.10/dist-packages (from matplotlib==3.7.2->
d2l) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib==3.
7.2->d2l) (4.54.1)
Requirement already satisfied: kiwisolver>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib==3.
7.2->d2l) (1.4.7)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib==3.7.
2->d2l) (24.1)
Requirement already satisfied: pillow>=6.2.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib==3.7.2-
>d2l) (10.4.0)
Collecting pyparsing<3.1,>=2.3.1 (from matplotlib==3.7.2->d2l)
  Downloading pyparsing-3.0.9-py3-none-any.whl.metadata (4.2 kB)
Requirement already satisfied: python-dateutil>=2.7 in /usr/local/lib/python3.10/dist-packages (from matplotlib=
=3.7.2->d2l) (2.8.2)
Requirement already satisfied: traitlets in /usr/local/lib/python3.10/dist-packages (from matplotlib-inline==0.1
.6->d2l) (5.7.1)
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.10/dist-packages (from pandas==2.0.3->d2l)
(2024.2)
Requirement already satisfied: tzdata>=2022.1 in /usr/local/lib/python3.10/dist-packages (from pandas==2.0.3->d2
l) (2024.2)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.10/dist-packages (from request
s==2.31.0->d2l) (3.3.2)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.10/dist-packages (from requests==2.31.0->d
2l) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.10/dist-packages (from requests==2.3
1.0->d2l) (2.2.3)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.10/dist-packages (from requests==2.3
1.0->d2l) (2024.8.30)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.10/dist-packages (from python-dateutil>=2.7->m
atplotlib==3.7.2->d2l) (1.16.0)
Requirement already satisfied: ipython-genutils in /usr/local/lib/python3.10/dist-packages (from ipykernel->jupy
ter==1.0.0->d2l) (0.2.0)
Requirement already satisfied: ipython>=5.0.0 in /usr/local/lib/python3.10/dist-packages (from ipykernel->jupyte
r==1.0.0->d2l) (7.34.0)
Requirement already satisfied: jupyter-client in /usr/local/lib/python3.10/dist-packages (from ipykernel->jupyte
r==1.0.0->d2l) (6.1.12)
Requirement already satisfied: tornado>=4.2 in /usr/local/lib/python3.10/dist-packages (from ipykernel->jupyter=
=1.0.0->d2l) (6.3.3)
Requirement already satisfied: widgetsnbextension~=3.6.0 in /usr/local/lib/python3.10/dist-packages (from ipywid
gets->jupyter==1.0.0->d2l) (3.6.9)
Requirement already satisfied: jupyterlab-widgets>=1.0.0 in /usr/local/lib/python3.10/dist-packages (from ipywid
gets->jupyter==1.0.0->d2l) (3.0.13)
Requirement already satisfied: prompt-toolkit!=3.0.0,!=3.0.1,<3.1.0,>=2.0.0 in /usr/local/lib/python3.10/dist-pa
ckages (from jupyter-console->jupyter==1.0.0->d2l) (3.0.48)
Requirement already satisfied: pygments in /usr/local/lib/python3.10/dist-packages (from jupyter-console->jupyte
r==1.0.0->d2l) (2.18.0)
Requirement already satisfied: lxml in /usr/local/lib/python3.10/dist-packages (from nbconvert->jupyter==1.0.0->
d2l) (4.9.4)
Requirement already satisfied: beautifulsoup4 in /usr/local/lib/python3.10/dist-packages (from nbconvert->jupyte
```

```
r==1.0.0->d2l) (4.12.3)
Requirement already satisfied: bleach in /usr/local/lib/python3.10/dist-packages (from nbconvert->jupyter==1.0.0
->d2l) (6.1.0)
Requirement already satisfied: defusedxml in /usr/local/lib/python3.10/dist-packages (from nbconvert->jupyter==1
.0.0->d2l) (0.7.1)
Requirement already satisfied: entrypoints>=0.2.2 in /usr/local/lib/python3.10/dist-packages (from nbconvert->ju
pyter==1.0.0->d2l) (0.4)
Requirement already satisfied: jinja2>=3.0 in /usr/local/lib/python3.10/dist-packages (from nbconvert->jupyter==
1.0.0->d2l) (3.1.4)
Requirement already satisfied: jupyter-core>=4.7 in /usr/local/lib/python3.10/dist-packages (from nbconvert->jup
yter==1.0.0->d2l) (5.7.2)
Requirement already satisfied: jupyterlab-pygments in /usr/local/lib/python3.10/dist-packages (from nbconvert->j
upyter==1.0.0->d2l) (0.3.0)
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.10/dist-packages (from nbconvert->jupyt
er==1.0.0->d2l) (2.1.5)
Requirement already satisfied: mistune<2,>=0.8.1 in /usr/local/lib/python3.10/dist-packages (from nbconvert->jup
yter==1.0.0->d2l) (0.8.4)
Requirement already satisfied: nbclient>=0.5.0 in /usr/local/lib/python3.10/dist-packages (from nbconvert->jupyt
er==1.0.0->d2l) (0.10.0)
Requirement already satisfied: nbformat>=5.1 in /usr/local/lib/python3.10/dist-packages (from nbconvert->jupyter
==1.0.0->d2l) (5.10.4)
Requirement already satisfied: pandocfilters>=1.4.1 in /usr/local/lib/python3.10/dist-packages (from nbconvert->
jupyter==1.0.0->d2l) (1.5.1)
Requirement already satisfied: tinycss2 in /usr/local/lib/python3.10/dist-packages (from nbconvert->jupyter==1.0
.0->d2l) (1.3.0)
Requirement already satisfied: pyzmq<25,>=17 in /usr/local/lib/python3.10/dist-packages (from notebook->jupyter=
=1.0.0->d2l) (24.0.1)
Requirement already satisfied: argon2-cffi in /usr/local/lib/python3.10/dist-packages (from notebook->jupyter==1
.0.0->d2l) (23.1.0)
Requirement already satisfied: nest-asyncio>=1.5 in /usr/local/lib/python3.10/dist-packages (from notebook->jupy
ter==1.0.0->d2l) (1.6.0)
Requirement already satisfied: Send2Trash>=1.8.0 in /usr/local/lib/python3.10/dist-packages (from notebook->jupy
ter==1.0.0->d2l) (1.8.3)
Requirement already satisfied: terminado>=0.8.3 in /usr/local/lib/python3.10/dist-packages (from notebook->jupyt
er==1.0.0->d2l) (0.18.1)
Requirement already satisfied: prometheus-client in /usr/local/lib/python3.10/dist-packages (from notebook->jupy
ter==1.0.0->d2l) (0.21.0)
Requirement already satisfied: nbclassic>=0.4.7 in /usr/local/lib/python3.10/dist-packages (from notebook->jupyt
er==1.0.0->d2l) (1.1.0)
Collecting qtpy>=2.4.0 (from qtconsole->jupyter==1.0.0->d2l)
  Downloading QtPy-2.4.1-py3-none-any.whl.metadata (12 kB)
Requirement already satisfied: setuptools>=18.5 in /usr/local/lib/python3.10/dist-packages (from ipython>=5.0.0-
>ipykernel->jupyter==1.0.0->d2l) (71.0.4)
Collecting jedi>=0.16 (from ipython>=5.0.0->ipykernel->jupyter==1.0.0->d2l)
  Using cached jedi-0.19.1-py2.py3-none-any.whl.metadata (22 kB)
Requirement already satisfied: decorator in /usr/local/lib/python3.10/dist-packages (from ipython>=5.0.0->ipyker
nel->jupyter==1.0.0->d2l) (4.4.2)
Requirement already satisfied: pickleshare in /usr/local/lib/python3.10/dist-packages (from ipython>=5.0.0->ipyk
ernel->jupyter==1.0.0->d2l) (0.7.5)
Requirement already satisfied: backcall in /usr/local/lib/python3.10/dist-packages (from ipython>=5.0.0->ipykern
el->jupyter==1.0.0->d2l) (0.2.0)
Requirement already satisfied: pexpect>4.3 in /usr/local/lib/python3.10/dist-packages (from ipython>=5.0.0->ipyk
ernel->jupyter==1.0.0->d2l) (4.9.0)
Requirement already satisfied: platformdirs>=2.5 in /usr/local/lib/python3.10/dist-packages (from jupyter-core>=
4.7->nbconvert->jupyter==1.0.0->d2l) (4.3.6)
Requirement already satisfied: notebook-shim>=0.2.3 in /usr/local/lib/python3.10/dist-packages (from nbclassic>=
0.4.7->notebook->jupyter==1.0.0->d2l) (0.2.4)
Requirement already satisfied: fastjsonschema>=2.15 in /usr/local/lib/python3.10/dist-packages (from nbformat>=5
.1->nbconvert->jupyter==1.0.0->d2l) (2.20.0)
Requirement already satisfied: jsonschema>=2.6 in /usr/local/lib/python3.10/dist-packages (from nbformat>=5.1->n
bconvert->jupyter==1.0.0->d2l) (4.23.0)
Requirement already satisfied: wcwidth in /usr/local/lib/python3.10/dist-packages (from prompt-toolkit!=3.0.0,!=
3.0.1,<3.1.0,>=2.0.0->jupyter-console->jupyter==1.0.0->d2l) (0.2.13)
Requirement already satisfied: ptyprocess in /usr/local/lib/python3.10/dist-packages (from terminado>=0.8.3->not
ebook->jupyter==1.0.0->d2l) (0.7.0)
Requirement already satisfied: argon2-cffi-bindings in /usr/local/lib/python3.10/dist-packages (from argon2-cffi
->notebook->jupyter==1.0.0->d2l) (21.2.0)
Requirement already satisfied: soupsieve>1.2 in /usr/local/lib/python3.10/dist-packages (from beautifulsoup4->nb
convert->jupyter==1.0.0->d2l) (2.6)
Requirement already satisfied: webencodings in /usr/local/lib/python3.10/dist-packages (from bleach->nbconvert->
jupyter==1.0.0->d2l) (0.5.1)
Requirement already satisfied: parso<0.9.0,>=0.8.3 in /usr/local/lib/python3.10/dist-packages (from jedi>=0.16->
ipython>=5.0.0->ipykernel->jupyter==1.0.0->d2l) (0.8.4)
Requirement already satisfied: attrs>=22.2.0 in /usr/local/lib/python3.10/dist-packages (from jsonschema>=2.6->n
bformat>=5.1->nbconvert->jupyter==1.0.0->d2l) (24.2.0)
Requirement already satisfied: jsonschema-specifications>=2023.03.6 in /usr/local/lib/python3.10/dist-packages (
from jsonschema>=2.6->nbformat>=5.1->nbconvert->jupyter==1.0.0->d2l) (2023.12.1)
Requirement already satisfied: referencing>=0.28.4 in /usr/local/lib/python3.10/dist-packages (from jsonschema>=
2.6->nbformat>=5.1->nbconvert->jupyter==1.0.0->d2l) (0.35.1)
Requirement already satisfied: rpds-py>=0.7.1 in /usr/local/lib/python3.10/dist-packages (from jsonschema>=2.6->
nbformat>=5.1->nbconvert->jupyter==1.0.0->d2l) (0.20.0)
Requirement already satisfied: jupyter-server<3,>=1.8 in /usr/local/lib/python3.10/dist-packages (from notebook-
shim>=0.2.3->nbclassic>=0.4.7->notebook->jupyter==1.0.0->d2l) (1.24.0)
```

```
Requirement already satisfied: cffi>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from argon2-cffi-bindings
->argon2-cffi->notebook->jupyter==1.0.0->d2l) (1.17.1)
Requirement already satisfied: pycparser in /usr/local/lib/python3.10/dist-packages (from cffi>=1.0.1->argon2-cf
fi-bindings->argon2-cffi->notebook->jupyter==1.0.0->d2l) (2.22)
Requirement already satisfied: anyio<4,>=3.1.0 in /usr/local/lib/python3.10/dist-packages (from jupyter-server<3
,>=1.8->notebook-shim>=0.2.3->nbclassic>=0.4.7->notebook->jupyter==1.0.0->d2l) (3.7.1)
Requirement already satisfied: websocket-client in /usr/local/lib/python3.10/dist-packages (from jupyter-server<
3,>=1.8->notebook-shim>=0.2.3->nbclassic>=0.4.7->notebook->jupyter==1.0.0->d2l) (1.8.0)
Requirement already satisfied: sniffio>=1.1 in /usr/local/lib/python3.10/dist-packages (from anyio<4,>=3.1.0->ju
pyter-server<3,>=1.8->notebook-shim>=0.2.3->nbclassic>=0.4.7->notebook->jupyter==1.0.0->d2l) (1.3.1)
Requirement already satisfied: exceptiongroup in /usr/local/lib/python3.10/dist-packages (from anyio<4,>=3.1.0->
jupyter-server<3,>=1.8->notebook-shim>=0.2.3->nbclassic>=0.4.7->notebook->jupyter==1.0.0->d2l) (1.2.2)
Downloading d2l-1.0.3-py3-none-any.whl (111 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 111.7/111.7 kB 3.5 MB/s eta 0:00:00
Downloading jupyter-1.0.0-py2.py3-none-any.whl (2.7 kB)
Downloading matplotlib-3.7.2-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (11.6 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 11.6/11.6 MB 33.7 MB/s eta 0:00:00
Downloading matplotlib_inline-0.1.6-py3-none-any.whl (9.4 kB)
Downloading numpy-1.23.5-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (17.1 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 17.1/17.1 MB 33.3 MB/s eta 0:00:00
Downloading pandas-2.0.3-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (12.3 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 12.3/12.3 MB 37.1 MB/s eta 0:00:00
Downloading requests-2.31.0-py3-none-any.whl (62 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 62.6/62.6 kB 2.7 MB/s eta 0:00:00
Downloading scipy-1.10.1-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (34.4 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 34.4/34.4 MB 14.8 MB/s eta 0:00:00
Downloading pyparsing-3.0.9-py3-none-any.whl (98 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 98.3/98.3 kB 5.4 MB/s eta 0:00:00
Downloading qtconsole-5.6.0-py3-none-any.whl (124 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 124.7/124.7 kB 8.9 MB/s eta 0:00:00
Downloading QtPy-2.4.1-py3-none-any.whl (93 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 93.5/93.5 kB 7.8 MB/s eta 0:00:00
Using cached jedi-0.19.1-py2.py3-none-any.whl (1.6 MB)
Installing collected packages: requests, qtpy, pyparsing, numpy, matplotlib-inline, jedi, scipy, pandas, matplot
lib, qtconsole, jupyter, d2l
  Attempting uninstall: requests
    Found existing installation: requests 2.32.3
    Uninstalling requests-2.32.3:
      Successfully uninstalled requests-2.32.3
  Attempting uninstall: pyparsing
    Found existing installation: pyparsing 3.1.4
    Uninstalling pyparsing-3.1.4:
      Successfully uninstalled pyparsing-3.1.4
  Attempting uninstall: numpy
    Found existing installation: numpy 1.26.4
    Uninstalling numpy-1.26.4:
      Successfully uninstalled numpy-1.26.4
  Attempting uninstall: matplotlib-inline
    Found existing installation: matplotlib-inline 0.1.7
    Uninstalling matplotlib-inline-0.1.7:
      Successfully uninstalled matplotlib-inline-0.1.7
  Attempting uninstall: scipy
    Found existing installation: scipy 1.13.1
    Uninstalling scipy-1.13.1:
      Successfully uninstalled scipy-1.13.1
  Attempting uninstall: pandas
    Found existing installation: pandas 2.2.2
    Uninstalling pandas-2.2.2:
      Successfully uninstalled pandas-2.2.2
  Attempting uninstall: matplotlib
    Found existing installation: matplotlib 3.7.1
    Uninstalling matplotlib-3.7.1:
      Successfully uninstalled matplotlib-3.7.1
ERROR: pip's dependency resolver does not currently take into account all the packages that are installed. This
behaviour is the source of the following dependency conflicts.
albucore 0.0.16 requires numpy>=1.24, but you have numpy 1.23.5 which is incompatible.
albumentations 1.4.15 requires numpy>=1.24.4, but you have numpy 1.23.5 which is incompatible.
bigframes 1.19.0 requires numpy>=1.24.0, but you have numpy 1.23.5 which is incompatible.
chex 0.1.87 requires numpy>=1.24.1, but you have numpy 1.23.5 which is incompatible.
google-colab 1.0.0 requires pandas==2.2.2, but you have pandas 2.0.3 which is incompatible.
google-colab 1.0.0 requires requests==2.32.3, but you have requests 2.31.0 which is incompatible.
jax 0.4.33 requires numpy>=1.24, but you have numpy 1.23.5 which is incompatible.
jaxlib 0.4.33 requires numpy>=1.24, but you have numpy 1.23.5 which is incompatible.
mizani 0.11.4 requires pandas>=2.1.0, but you have pandas 2.0.3 which is incompatible.
plotnine 0.13.6 requires pandas<3.0.0,>=2.1.0, but you have pandas 2.0.3 which is incompatible.
xarray 2024.9.0 requires numpy>=1.24, but you have numpy 1.23.5 which is incompatible.
xarray 2024.9.0 requires pandas>=2.1, but you have pandas 2.0.3 which is incompatible.
Successfully installed d2l-1.0.3 jedi-0.19.1 jupyter-1.0.0 matplotlib-3.7.2 matplotlib-inline-0.1.6 numpy-1.23.5
pandas-2.0.3 pyparsing-3.0.9 qtconsole-5.6.0 qtpy-2.4.1 requests-2.31.0 scipy-1.10.1
```

In [1]:

```python
import torch
from torch import nn
from d2l import torch as d2l
```

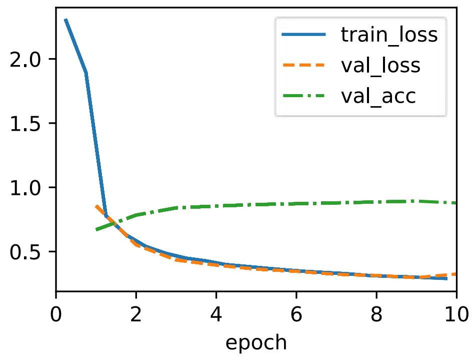
In [2]:
```python
def vgg_block(num_convs, out_channels):
    layers = []
    for _ in range(num_convs):
        layers.append(nn.LazyConv2d(out_channels, kernel_size=3, padding=1))
        layers.append(nn.ReLU())
    layers.append(nn.MaxPool2d(kernel_size=2,stride=2))
    return nn.Sequential(*layers)
```

In [3]:
```python
class VGG(d2l.Classifier):
    def __init__(self, arch, lr=0.1, num_classes=10):
        super().__init__()
        self.save_hyperparameters()
        conv_blks = []
        for (num_convs, out_channels) in arch:
            conv_blks.append(vgg_block(num_convs, out_channels))
        self.net = nn.Sequential(
            *conv_blks, nn.Flatten(),
            nn.LazyLinear(4096), nn.ReLU(), nn.Dropout(0.5),
            nn.LazyLinear(4096), nn.ReLU(), nn.Dropout(0.5),
            nn.LazyLinear(num_classes))
        self.net.apply(d2l.init_cnn)
```

In [4]:
```python
VGG(arch=((1, 64), (1, 128), (2, 256), (2, 512), (2, 512))).layer_summary(
    (1, 1, 224, 224))
```

```
Sequential output shape:         torch.Size([1, 64, 112, 112])
Sequential output shape:         torch.Size([1, 128, 56, 56])
Sequential output shape:         torch.Size([1, 256, 28, 28])
Sequential output shape:         torch.Size([1, 512, 14, 14])
Sequential output shape:         torch.Size([1, 512, 7, 7])
Flatten output shape:    torch.Size([1, 25088])
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])
```

In [5]:
```python
model = VGG(arch=((1, 16), (1, 32), (2, 64), (2, 128), (2, 128)), lr=0.01)
trainer = d2l.Trainer(max_epochs=10, num_gpus=1)
data = d2l.FashionMNIST(batch_size=128, resize=(224, 224))
model.apply_init([next(iter(data.get_dataloader(True)))[0]], d2l.init_cnn)
trainer.fit(model, data)
```



Discussion: 일단 지금까지의 단원들에 비해 시간이 압도적으로 많이 걸렸다. 이번 단원은 VGG Network에 대해서 설명하는데, 이는 CNN의 여러 구조중 하나로, block구조를 사용한다는 특징을 가지고 있다. 각각의 block은 합성곱층, ReLU, Max Pooling의 순서로 이루어진다고 한다. 이러한 block들이 구성된 network는 크게 2가지 부분으로 나뉘는데, 합성곱 및 풀링 층으로 구성된 부분에서는 이미지에서 특징을 추출하고 입력 공간의 크기를 점차 줄여나가는 역할을 하고, 완전 연결 층으로 구성된 부분에서는 마지막에 추출된 특징을 사용하여 이미지를 분리하는 역할을 하였다. 이 예시에서는 5개의 block으로 구성되어 있으며, 위의 코드의 VGG구조에서 각 block의 채널 수나 계산 비용이 매우 크다는 것을 알 수 있다. 실제 훈련 및 실험에서는 채널 수를 줄여 더 간단한 버전을 훈련하였다. (그러나, 아무래도 데스크탑도 아니고 노트북으로 진행하고 있다 보니 시간이 여전히 너무 오래 걸렸다. 그래서 처음으로 colab으로 변경해서 진행하였다.)

In [3]:
```
!pip install d2l
```

```
Collecting d2l
  Downloading d2l-1.0.3-py3-none-any.whl.metadata (556 bytes)
Collecting jupyter==1.0.0 (from d2l)
  Downloading jupyter-1.0.0-py2.py3-none-any.whl.metadata (995 bytes)
Collecting numpy==1.23.5 (from d2l)
  Downloading numpy-1.23.5-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (2.3 kB)
Collecting matplotlib==3.7.2 (from d2l)
  Downloading matplotlib-3.7.2-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (5.6 kB)
Collecting matplotlib-inline==0.1.6 (from d2l)
  Downloading matplotlib_inline-0.1.6-py3-none-any.whl.metadata (2.8 kB)
Collecting requests==2.31.0 (from d2l)
  Downloading requests-2.31.0-py3-none-any.whl.metadata (4.6 kB)
Collecting pandas==2.0.3 (from d2l)
  Downloading pandas-2.0.3-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (18 kB)
Collecting scipy==1.10.1 (from d2l)
  Downloading scipy-1.10.1-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (58 kB)
     ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 58.9/58.9 kB 5.2 MB/s eta 0:00:00
Requirement already satisfied: notebook in /usr/local/lib/python3.10/dist-packages (from jupyter==1.0.0->d2l) (6
.5.5)
Collecting qtconsole (from jupyter==1.0.0->d2l)
  Downloading qtconsole-5.6.0-py3-none-any.whl.metadata (5.0 kB)
Requirement already satisfied: jupyter-console in /usr/local/lib/python3.10/dist-packages (from jupyter==1.0.0->
d2l) (6.1.0)
Requirement already satisfied: nbconvert in /usr/local/lib/python3.10/dist-packages (from jupyter==1.0.0->d2l) (
6.5.4)
Requirement already satisfied: ipykernel in /usr/local/lib/python3.10/dist-packages (from jupyter==1.0.0->d2l) (
5.5.6)
Requirement already satisfied: ipywidgets in /usr/local/lib/python3.10/dist-packages (from jupyter==1.0.0->d2l)
(7.7.1)
Requirement already satisfied: contourpy>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib==3.7
.2->d2l) (1.3.0)
Requirement already satisfied: cycler>=0.10 in /usr/local/lib/python3.10/dist-packages (from matplotlib==3.7.2->
d2l) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib==3.
7.2->d2l) (4.54.1)
Requirement already satisfied: kiwisolver>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib==3.
7.2->d2l) (1.4.7)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib==3.7.
2->d2l) (24.1)
Requirement already satisfied: pillow>=6.2.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib==3.7.2-
>d2l) (10.4.0)
Collecting pyparsing<3.1,>=2.3.1 (from matplotlib==3.7.2->d2l)
  Downloading pyparsing-3.0.9-py3-none-any.whl.metadata (4.2 kB)
Requirement already satisfied: python-dateutil>=2.7 in /usr/local/lib/python3.10/dist-packages (from matplotlib=
=3.7.2->d2l) (2.8.2)
Requirement already satisfied: traitlets in /usr/local/lib/python3.10/dist-packages (from matplotlib-inline==0.1
.6->d2l) (5.7.1)
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.10/dist-packages (from pandas==2.0.3->d2l)
(2024.2)
Requirement already satisfied: tzdata>=2022.1 in /usr/local/lib/python3.10/dist-packages (from pandas==2.0.3->d2
l) (2024.2)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.10/dist-packages (from request
s==2.31.0->d2l) (3.3.2)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.10/dist-packages (from requests==2.31.0->d
2l) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.10/dist-packages (from requests==2.3
1.0->d2l) (2.2.3)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.10/dist-packages (from requests==2.3
1.0->d2l) (2024.8.30)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.10/dist-packages (from python-dateutil>=2.7->m
atplotlib==3.7.2->d2l) (1.16.0)
Requirement already satisfied: ipython-genutils in /usr/local/lib/python3.10/dist-packages (from ipykernel->jupy
ter==1.0.0->d2l) (0.2.0)
Requirement already satisfied: ipython>=5.0.0 in /usr/local/lib/python3.10/dist-packages (from ipykernel->jupyte
r==1.0.0->d2l) (7.34.0)
Requirement already satisfied: jupyter-client in /usr/local/lib/python3.10/dist-packages (from ipykernel->jupyte
r==1.0.0->d2l) (6.1.12)
Requirement already satisfied: tornado>=4.2 in /usr/local/lib/python3.10/dist-packages (from ipykernel->jupyter=
=1.0.0->d2l) (6.3.3)
Requirement already satisfied: widgetsnbextension~=3.6.0 in /usr/local/lib/python3.10/dist-packages (from ipywid
gets->jupyter==1.0.0->d2l) (3.6.9)
Requirement already satisfied: jupyterlab-widgets>=1.0.0 in /usr/local/lib/python3.10/dist-packages (from ipywid
gets->jupyter==1.0.0->d2l) (3.0.13)
Requirement already satisfied: prompt-toolkit!=3.0.0,!=3.0.1,<3.1.0,>=2.0.0 in /usr/local/lib/python3.10/dist-pa
ckages (from jupyter-console->jupyter==1.0.0->d2l) (3.0.48)
Requirement already satisfied: pygments in /usr/local/lib/python3.10/dist-packages (from jupyter-console->jupyte
r==1.0.0->d2l) (2.18.0)
Requirement already satisfied: lxml in /usr/local/lib/python3.10/dist-packages (from nbconvert->jupyter==1.0.0->
d2l) (4.9.4)
Requirement already satisfied: beautifulsoup4 in /usr/local/lib/python3.10/dist-packages (from nbconvert->jupyte
```

```
r==1.0.0->d2l) (4.12.3)
Requirement already satisfied: bleach in /usr/local/lib/python3.10/dist-packages (from nbconvert->jupyter==1.0.0
->d2l) (6.1.0)
Requirement already satisfied: defusedxml in /usr/local/lib/python3.10/dist-packages (from nbconvert->jupyter==1
.0.0->d2l) (0.7.1)
Requirement already satisfied: entrypoints>=0.2.2 in /usr/local/lib/python3.10/dist-packages (from nbconvert->ju
pyter==1.0.0->d2l) (0.4)
Requirement already satisfied: jinja2>=3.0 in /usr/local/lib/python3.10/dist-packages (from nbconvert->jupyter==
1.0.0->d2l) (3.1.4)
Requirement already satisfied: jupyter-core>=4.7 in /usr/local/lib/python3.10/dist-packages (from nbconvert->jup
yter==1.0.0->d2l) (5.7.2)
Requirement already satisfied: jupyterlab-pygments in /usr/local/lib/python3.10/dist-packages (from nbconvert->j
upyter==1.0.0->d2l) (0.3.0)
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.10/dist-packages (from nbconvert->jupyt
er==1.0.0->d2l) (2.1.5)
Requirement already satisfied: mistune<2,>=0.8.1 in /usr/local/lib/python3.10/dist-packages (from nbconvert->jup
yter==1.0.0->d2l) (0.8.4)
Requirement already satisfied: nbclient>=0.5.0 in /usr/local/lib/python3.10/dist-packages (from nbconvert->jupyt
er==1.0.0->d2l) (0.10.0)
Requirement already satisfied: nbformat>=5.1 in /usr/local/lib/python3.10/dist-packages (from nbconvert->jupyter
==1.0.0->d2l) (5.10.4)
Requirement already satisfied: pandocfilters>=1.4.1 in /usr/local/lib/python3.10/dist-packages (from nbconvert->
jupyter==1.0.0->d2l) (1.5.1)
Requirement already satisfied: tinycss2 in /usr/local/lib/python3.10/dist-packages (from nbconvert->jupyter==1.0
.0->d2l) (1.3.0)
Requirement already satisfied: pyzmq<25,>=17 in /usr/local/lib/python3.10/dist-packages (from notebook->jupyter=
=1.0.0->d2l) (24.0.1)
Requirement already satisfied: argon2-cffi in /usr/local/lib/python3.10/dist-packages (from notebook->jupyter==1
.0.0->d2l) (23.1.0)
Requirement already satisfied: nest-asyncio>=1.5 in /usr/local/lib/python3.10/dist-packages (from notebook->jupy
ter==1.0.0->d2l) (1.6.0)
Requirement already satisfied: Send2Trash>=1.8.0 in /usr/local/lib/python3.10/dist-packages (from notebook->jupy
ter==1.0.0->d2l) (1.8.3)
Requirement already satisfied: terminado>=0.8.3 in /usr/local/lib/python3.10/dist-packages (from notebook->jupyt
er==1.0.0->d2l) (0.18.1)
Requirement already satisfied: prometheus-client in /usr/local/lib/python3.10/dist-packages (from notebook->jupy
ter==1.0.0->d2l) (0.21.0)
Requirement already satisfied: nbclassic>=0.4.7 in /usr/local/lib/python3.10/dist-packages (from notebook->jupyt
er==1.0.0->d2l) (1.1.0)
Collecting qtpy>=2.4.0 (from qtconsole->jupyter==1.0.0->d2l)
  Downloading QtPy-2.4.1-py3-none-any.whl.metadata (12 kB)
Requirement already satisfied: setuptools>=18.5 in /usr/local/lib/python3.10/dist-packages (from ipython>=5.0.0-
>ipykernel->jupyter==1.0.0->d2l) (71.0.4)
Collecting jedi>=0.16 (from ipython>=5.0.0->ipykernel->jupyter==1.0.0->d2l)
  Using cached jedi-0.19.1-py2.py3-none-any.whl.metadata (22 kB)
Requirement already satisfied: decorator in /usr/local/lib/python3.10/dist-packages (from ipython>=5.0.0->ipyker
nel->jupyter==1.0.0->d2l) (4.4.2)
Requirement already satisfied: pickleshare in /usr/local/lib/python3.10/dist-packages (from ipython>=5.0.0->ipyk
ernel->jupyter==1.0.0->d2l) (0.7.5)
Requirement already satisfied: backcall in /usr/local/lib/python3.10/dist-packages (from ipython>=5.0.0->ipykern
el->jupyter==1.0.0->d2l) (0.2.0)
Requirement already satisfied: pexpect>4.3 in /usr/local/lib/python3.10/dist-packages (from ipython>=5.0.0->ipyk
ernel->jupyter==1.0.0->d2l) (4.9.0)
Requirement already satisfied: platformdirs>=2.5 in /usr/local/lib/python3.10/dist-packages (from jupyter-core>=
4.7->nbconvert->jupyter==1.0.0->d2l) (4.3.6)
Requirement already satisfied: notebook-shim>=0.2.3 in /usr/local/lib/python3.10/dist-packages (from nbclassic>=
0.4.7->notebook->jupyter==1.0.0->d2l) (0.2.4)
Requirement already satisfied: fastjsonschema>=2.15 in /usr/local/lib/python3.10/dist-packages (from nbformat>=5
.1->nbconvert->jupyter==1.0.0->d2l) (2.20.0)
Requirement already satisfied: jsonschema>=2.6 in /usr/local/lib/python3.10/dist-packages (from nbformat>=5.1->n
bconvert->jupyter==1.0.0->d2l) (4.23.0)
Requirement already satisfied: wcwidth in /usr/local/lib/python3.10/dist-packages (from prompt-toolkit!=3.0.0,!=
3.0.1,<3.1.0,>=2.0.0->jupyter-console->jupyter==1.0.0->d2l) (0.2.13)
Requirement already satisfied: ptyprocess in /usr/local/lib/python3.10/dist-packages (from terminado>=0.8.3->not
ebook->jupyter==1.0.0->d2l) (0.7.0)
Requirement already satisfied: argon2-cffi-bindings in /usr/local/lib/python3.10/dist-packages (from argon2-cffi
->notebook->jupyter==1.0.0->d2l) (21.2.0)
Requirement already satisfied: soupsieve>1.2 in /usr/local/lib/python3.10/dist-packages (from beautifulsoup4->nb
convert->jupyter==1.0.0->d2l) (2.6)
Requirement already satisfied: webencodings in /usr/local/lib/python3.10/dist-packages (from bleach->nbconvert->
jupyter==1.0.0->d2l) (0.5.1)
Requirement already satisfied: parso<0.9.0,>=0.8.3 in /usr/local/lib/python3.10/dist-packages (from jedi>=0.16->
ipython>=5.0.0->ipykernel->jupyter==1.0.0->d2l) (0.8.4)
Requirement already satisfied: attrs>=22.2.0 in /usr/local/lib/python3.10/dist-packages (from jsonschema>=2.6->n
bformat>=5.1->nbconvert->jupyter==1.0.0->d2l) (24.2.0)
Requirement already satisfied: jsonschema-specifications>=2023.03.6 in /usr/local/lib/python3.10/dist-packages (
from jsonschema>=2.6->nbformat>=5.1->nbconvert->jupyter==1.0.0->d2l) (2023.12.1)
Requirement already satisfied: referencing>=0.28.4 in /usr/local/lib/python3.10/dist-packages (from jsonschema>=
2.6->nbformat>=5.1->nbconvert->jupyter==1.0.0->d2l) (0.35.1)
Requirement already satisfied: rpds-py>=0.7.1 in /usr/local/lib/python3.10/dist-packages (from jsonschema>=2.6->
nbformat>=5.1->nbconvert->jupyter==1.0.0->d2l) (0.20.0)
Requirement already satisfied: jupyter-server<3,>=1.8 in /usr/local/lib/python3.10/dist-packages (from notebook-
shim>=0.2.3->nbclassic>=0.4.7->notebook->jupyter==1.0.0->d2l) (1.24.0)
```

```
Requirement already satisfied: cffi>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from argon2-cffi-bindings
->argon2-cffi->notebook->jupyter==1.0.0->d2l) (1.17.1)
Requirement already satisfied: pycparser in /usr/local/lib/python3.10/dist-packages (from cffi>=1.0.1->argon2-cf
fi-bindings->argon2-cffi->notebook->jupyter==1.0.0->d2l) (2.22)
Requirement already satisfied: anyio<4,>=3.1.0 in /usr/local/lib/python3.10/dist-packages (from jupyter-server<3
,>=1.8->notebook-shim>=0.2.3->nbclassic>=0.4.7->notebook->jupyter==1.0.0->d2l) (3.7.1)
Requirement already satisfied: websocket-client in /usr/local/lib/python3.10/dist-packages (from jupyter-server<
3,>=1.8->notebook-shim>=0.2.3->nbclassic>=0.4.7->notebook->jupyter==1.0.0->d2l) (1.8.0)
Requirement already satisfied: sniffio>=1.1 in /usr/local/lib/python3.10/dist-packages (from anyio<4,>=3.1.0->ju
pyter-server<3,>=1.8->notebook-shim>=0.2.3->nbclassic>=0.4.7->notebook->jupyter==1.0.0->d2l) (1.3.1)
Requirement already satisfied: exceptiongroup in /usr/local/lib/python3.10/dist-packages (from anyio<4,>=3.1.0->
jupyter-server<3,>=1.8->notebook-shim>=0.2.3->nbclassic>=0.4.7->notebook->jupyter==1.0.0->d2l) (1.2.2)
Downloading d2l-1.0.3-py3-none-any.whl (111 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 111.7/111.7 kB 11.0 MB/s eta 0:00:00
Downloading jupyter-1.0.0-py2.py3-none-any.whl (2.7 kB)
Downloading matplotlib-3.7.2-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (11.6 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 11.6/11.6 MB 72.6 MB/s eta 0:00:00
Downloading matplotlib_inline-0.1.6-py3-none-any.whl (9.4 kB)
Downloading numpy-1.23.5-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (17.1 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 17.1/17.1 MB 93.7 MB/s eta 0:00:00
Downloading pandas-2.0.3-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (12.3 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 12.3/12.3 MB 111.7 MB/s eta 0:00:00
Downloading requests-2.31.0-py3-none-any.whl (62 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 62.6/62.6 kB 6.5 MB/s eta 0:00:00
Downloading scipy-1.10.1-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (34.4 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 34.4/34.4 MB 16.7 MB/s eta 0:00:00
Downloading pyparsing-3.0.9-py3-none-any.whl (98 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 98.3/98.3 kB 10.6 MB/s eta 0:00:00
Downloading qtconsole-5.6.0-py3-none-any.whl (124 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 124.7/124.7 kB 13.5 MB/s eta 0:00:00
Downloading QtPy-2.4.1-py3-none-any.whl (93 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 93.5/93.5 kB 8.7 MB/s eta 0:00:00
Using cached jedi-0.19.1-py2.py3-none-any.whl (1.6 MB)
Installing collected packages: requests, qtpy, pyparsing, numpy, matplotlib-inline, jedi, scipy, pandas, matplot
lib, qtconsole, jupyter, d2l
  Attempting uninstall: requests
    Found existing installation: requests 2.32.3
    Uninstalling requests-2.32.3:
      Successfully uninstalled requests-2.32.3
  Attempting uninstall: pyparsing
    Found existing installation: pyparsing 3.1.4
    Uninstalling pyparsing-3.1.4:
      Successfully uninstalled pyparsing-3.1.4
  Attempting uninstall: numpy
    Found existing installation: numpy 1.26.4
    Uninstalling numpy-1.26.4:
      Successfully uninstalled numpy-1.26.4
  Attempting uninstall: matplotlib-inline
    Found existing installation: matplotlib-inline 0.1.7
    Uninstalling matplotlib-inline-0.1.7:
      Successfully uninstalled matplotlib-inline-0.1.7
  Attempting uninstall: scipy
    Found existing installation: scipy 1.13.1
    Uninstalling scipy-1.13.1:
      Successfully uninstalled scipy-1.13.1
  Attempting uninstall: pandas
    Found existing installation: pandas 2.2.2
    Uninstalling pandas-2.2.2:
      Successfully uninstalled pandas-2.2.2
  Attempting uninstall: matplotlib
    Found existing installation: matplotlib 3.7.1
    Uninstalling matplotlib-3.7.1:
      Successfully uninstalled matplotlib-3.7.1
ERROR: pip's dependency resolver does not currently take into account all the packages that are installed. This
behaviour is the source of the following dependency conflicts.
albucore 0.0.16 requires numpy>=1.24, but you have numpy 1.23.5 which is incompatible.
albumentations 1.4.15 requires numpy>=1.24.4, but you have numpy 1.23.5 which is incompatible.
bigframes 1.19.0 requires numpy>=1.24.0, but you have numpy 1.23.5 which is incompatible.
chex 0.1.87 requires numpy>=1.24.1, but you have numpy 1.23.5 which is incompatible.
google-colab 1.0.0 requires pandas==2.2.2, but you have pandas 2.0.3 which is incompatible.
google-colab 1.0.0 requires requests==2.32.3, but you have requests 2.31.0 which is incompatible.
jax 0.4.33 requires numpy>=1.24, but you have numpy 1.23.5 which is incompatible.
jaxlib 0.4.33 requires numpy>=1.24, but you have numpy 1.23.5 which is incompatible.
mizani 0.11.4 requires pandas>=2.1.0, but you have pandas 2.0.3 which is incompatible.
plotnine 0.13.6 requires pandas<3.0.0,>=2.1.0, but you have pandas 2.0.3 which is incompatible.
xarray 2024.9.0 requires numpy>=1.24, but you have numpy 1.23.5 which is incompatible.
xarray 2024.9.0 requires pandas>=2.1, but you have pandas 2.0.3 which is incompatible.
Successfully installed d2l-1.0.3 jedi-0.19.1 jupyter-1.0.0 matplotlib-3.7.2 matplotlib-inline-0.1.6 numpy-1.23.5
pandas-2.0.3 pyparsing-3.0.9 qtconsole-5.6.0 qtpy-2.4.1 requests-2.31.0 scipy-1.10.1
```

In [3]:

```
import torch
from torch import nn
from torch.nn import functional as F
```

```python
from d2l import torch as d2l
```

In [4]:
```python
class Residual(nn.Module):
    """The Residual block of ResNet models."""
    def __init__(self, num_channels, use_1x1conv=False, strides=1):
        super().__init__()
        self.conv1 = nn.LazyConv2d(num_channels, kernel_size=3, padding=1,
                                   stride=strides)
        self.conv2 = nn.LazyConv2d(num_channels, kernel_size=3, padding=1)
        if use_1x1conv:
            self.conv3 = nn.LazyConv2d(num_channels, kernel_size=1,
                                       stride=strides)
        else:
            self.conv3 = None
        self.bn1 = nn.LazyBatchNorm2d()
        self.bn2 = nn.LazyBatchNorm2d()

    def forward(self, X):
        Y = F.relu(self.bn1(self.conv1(X)))
        Y = self.bn2(self.conv2(Y))
        if self.conv3:
            X = self.conv3(X)
        Y += X
        return F.relu(Y)
```

In [5]:
```python
blk = Residual(3)
X = torch.randn(4, 3, 6, 6)
blk(X).shape
```

Out[5]:
```
torch.Size([4, 3, 6, 6])
```

In [6]:
```python
blk = Residual(6, use_1x1conv=True, strides=2)
blk(X).shape
```

Out[6]:
```
torch.Size([4, 6, 3, 3])
```

In [7]:
```python
class ResNet(d2l.Classifier):
    def b1(self):
        return nn.Sequential(
            nn.LazyConv2d(64, kernel_size=7, stride=2, padding=3),
            nn.LazyBatchNorm2d(), nn.ReLU(),
            nn.MaxPool2d(kernel_size=3, stride=2, padding=1))
```

In [8]:
```python
@d2l.add_to_class(ResNet)
def block(self, num_residuals, num_channels, first_block=False):
    blk = []
    for i in range(num_residuals):
        if i == 0 and not first_block:
            blk.append(Residual(num_channels, use_1x1conv=True, strides=2))
        else:
            blk.append(Residual(num_channels))
    return nn.Sequential(*blk)
```

In [9]:
```python
@d2l.add_to_class(ResNet)
def __init__(self, arch, lr=0.1, num_classes=10):
    super(ResNet, self).__init__()
    self.save_hyperparameters()
    self.net = nn.Sequential(self.b1())
    for i, b in enumerate(arch):
        self.net.add_module(f'b{i+2}', self.block(*b, first_block=(i==0)))
    self.net.add_module('last', nn.Sequential(
        nn.AdaptiveAvgPool2d((1, 1)), nn.Flatten(),
        nn.LazyLinear(num_classes)))
    self.net.apply(d2l.init_cnn)
```

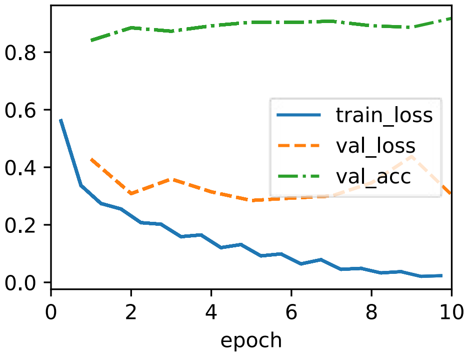
In [10]:
```python
class ResNet18(ResNet):
    def __init__(self, lr=0.1, num_classes=10):
        super().__init__(((2, 64), (2, 128), (2, 256), (2, 512)),
                       lr, num_classes)

ResNet18().layer_summary((1, 1, 96, 96))
```

```
Sequential output shape:         torch.Size([1, 64, 24, 24])
Sequential output shape:         torch.Size([1, 64, 24, 24])
Sequential output shape:         torch.Size([1, 128, 12, 12])
Sequential output shape:         torch.Size([1, 256, 6, 6])
Sequential output shape:         torch.Size([1, 512, 3, 3])
Sequential output shape:         torch.Size([1, 10])
```

In [11]:
```python
model = ResNet18(lr=0.01)
trainer = d2l.Trainer(max_epochs=10, num_gpus=1)
data = d2l.FashionMNIST(batch_size=128, resize=(96, 96))
```

```
model.apply_init([next(iter(data.get_dataloader(True)))[0]], d2l.init_cnn)
trainer.fit(model, data)
```



Discussion: 이 단원에서 배운 ResNet은, Residual connection(잔차 연결) 이라는 딥러닝에서 중요한 개념 중 하나를 도입한 네트워크이다. 깊은 신경망의 설계에서, 깊이가 너무 깊다면 학습이 어렵거나 성능이 떨어지는 문제점이 생기는데, 이럴 때 일부 층을 건너뛰는 방식인 잔차 연결을 이용하여 최소한의 성능을 보장하고 성능이 떨어지는 것을 막는다.이 방법은 네트워크가 입력과 출력의 잔차(residual)을 학습하게 된다.

이번 단원 역시 학습하는데 시간이 너무 걸려서, colab으로 진행하였다. 앞으로는 매번 pip install같은 것을 해야하는 문제가 있지만 gpu를 사용할 수 있는 colab을 사용해야 할 것 같다.