2.4 Convolutional Neural Networks

This task is about using a convolutional neural network for classifying pictures of food into 11 classes. The food 11 dataset was used.

convolutional neural network

The layers used were:

```
nn.Sequential(OrderedDict([
  ("conv0", nn.Conv2d(3, 96, 5, 4)),
  ("gelu0", nn.GELU()),
  ("pool0", nn.MaxPool2d(2)),
  ("conv1", nn.Conv2d(96, 288, 3)),
  ("gelu1", nn.GELU()),
  ("pool1", nn.MaxPool2d(2)),
  ("conv2", nn.Conv2d(288, 576, 3)),
  ("gelu2", nn.GELU()),
  ("pool2", nn.MaxPool2d(2)),
  ("conv3", nn.Conv2d(576, 1152, 3)),
  ("gelu3", nn.GELU()),
  ("pool3", nn.MaxPool2d(2)),
  ("view", View(1152*2*2)),
  ("fc0", nn.Linear(4608, 2304)),
  ("gelu4", nn.GELU()),
  ("fc1", nn.Linear(2304, 576)),
  ("gelu5", nn.GELU()),
  ("fc2", nn.Linear(576, 11)),
  ("sigmoid", nn.Sigmoid())
]))
```

The network had 4 convolutions and 3 fully connected layers. I made the input of the CNN 256x256 pixels big. Also, on training data I used torchvision's AutoAugment on the default setting. The CNN was initialized randomly.

The results were somewhere around 62% accuracy on the test data.