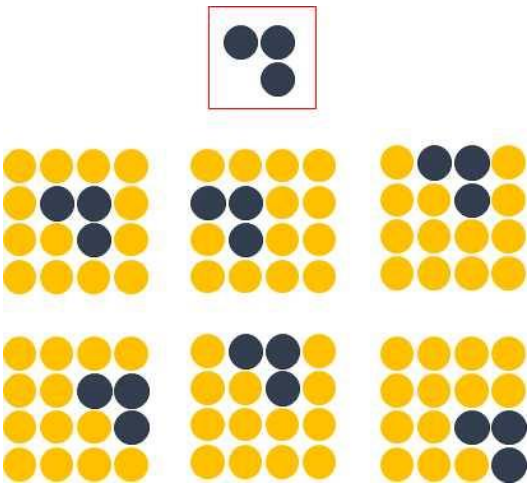
Report

Luyang Huang
14307110020

I. Convolutional Neural Network

In processing a picture, pictures represented by a 2D matrix are better than a 1D matrix, instinctively. Often, 2D matrix can keep the structure of a picture. If we want to know if a picture contains 'fold', as shown below. A full-connected neural network is inefficient because 'fold' on different place is stored in different place of weight matrix. To keep the invariance of a picture (including translation, rotation and illumination etc.), convolution is introduced into NN and thus leading to CNN.



i. Convolution

In digital image processing convolutional filtering plays an important role in many important algorithms in edge detection and related processes.

In optics, an out-of-focus photograph is a convolution of the sharp image with a lens function.

The math expression of a 2D convolution is:

$$y_{ij} = \sum_{u=1}^{m} \sum_{v=1}^{n} w_{uv} x_{i+u-1}\, y_{j+v-1}$$

By carefully checking this expression, we can find that if the weight matrix is symmetric, the rotation (by 90,180,270 degrees) and translation can be invariant. That is why convolution works well on processing image.

ii. Convolutional Layer

Normally, convolutional layer sweeps a 2D filter over a batch of images, applying the filter to each window of each image of the appropriate size. Often, in 2D convolutional layer, the input matrix contains four dimensions, batch size, height, width and depth. The stride and weight matrix are corresponding to these four sizes. Specially, in MNIST, the depth is one because it contains only

black-white pictures.

iii. Pooling Layer

The pooling layer sweeps a rectangular window over the input tensor, and returns a reduction operation (max, average, etc.). It is often used to epitomize the value of an area, for choosing features and reducing parameters.

II. Evaluation

I use ‘ReLU’ as activation function in convolution layers since they are often faster and do not influence the final accuracy.

Following Picture is the accuracy after 20000 batches, the final accuracy reaches about 0.99, which outperforms a simple dense.