09+NeuralNetwork

2019 年 3 月 27 日

1 神经网络

1.1 一、概念

神经网络是时下最热的人工智能话题，而神经网络的历史也由来已久，近年来的算力大爆发使人工智能和神经网络发现了彼此。

神经网络通过神经元进行组织，数据从上一层神经元流向下一层神经元直到输出神经元，损失函数衡量预测和输出之间的差距，再通过反向传播更新各层神经元的参数。

神经网络由如下元素构成：

1. 输入层：数据从输入层进入模型
2. 隐藏层：数据在隐藏层中进行交互和组合
3. 输出层：输出层输出预测结果
4. 激活函数：各个神经元的对上一层的输入进行非线性处理的函数
5. 损失函数：衡量预测结果和实际结果的差距
6. 优化器：即以何种方式更新参数

1.2 二、符号说明

- X：输入数据，$X \in R^{p\times n}$，p 代表变量数，n 代表样本数
- $\hat{Y}$：输出数据，$\hat{Y} \in R^{k\times n}$，n 代表样本数，多分类时 k 代表分类数，二分类和回归时 k 为 1
- Y：实际结果，$Y \in R^{k\times n}$，n 代表样本数，多分类时 k 代表分类数，二分类和回归时 k 为 1
- p_i：第 i 层的神经元数，$p_0 = p$
- W_i：从第 i-1 层向第 i 层传播的矩阵，$W_i \in R^{p_i\times p_{i-1}}$，输入层为第 0 层时，$W_1 \in R^{p_1\times p}$
- $\alpha(z)$：激活函数，每一层每一个神经元的激活函数都可以不同，此处统一用
- $g(z)$：输出层的激活函数，通常和隐藏层的激活函数不同

- b_i：第 i 层的偏置项，$b_i \in R^{p_{i+1}}$
- Z_i：上一层激活函数的线性组合，$Z_i \in R^{p_i \times n}$
- A_i：线性组合的激活函数值，$A_i \in R^{p_i \times n}$
- * ：逐元素相乘

1.3 三、Feed Forward 前向传播

1.3.1 1. 从输入层到第一个隐藏层

首先是对输入数据的线性组合，由于偏执项是一个向量，对所有 n 个数据来说都相等。虽然此处维度按照线性代数并不能严格成立（因为 $W_1X \in R^{p_1 \times n}$，$b_1 \in R^{p_1 \times 1}$），但是由于 numpy 中的广播（broadcast）机制存在，在编程中以下公式是成立的。如果非要按照数学定义上成立可以对 b_1 乘上一个 $1 \times n$ 的值全为 1 的向量。

$$
\begin{aligned}
Z_1 &= W_1X + b_1 \quad \in R^{p_1 \times n} \\
\Leftrightarrow Z_1 &= W_1X + b_1 1^{1 \times n}
\end{aligned}
$$

然后是对第一层的各个神经元进行 ‘‘激活’’，对线性组合进行逐元素的函数计算

$$A_1 = \alpha(Z_1) \quad \in R^{p_1 \times n}$$

1.3.2 2. 从第 i-1 层到第 i 层

与输入不同，此时是将上一层的激活函数值进行线性组合：

$$Z_i = W_iA_{i-1} + b_i \quad \in R^{p_i \times n}$$

$$A_i = \alpha(Z_i) \quad \in R^{p_i \times n}$$

1.3.3 3. 从最后一个隐藏层到输出层

假设输入层是第 0 层，第 1——m-1 层是隐藏层，第 m 层是输出层。如果是二分类、回归等情况，则输出层只有一个神经元，若是多分类等情况则有多个神经元，将在后面介绍，暂时假定只有一个输出：

$$Z_m = W_mA_{m-1} + b_m \quad \in R^{k \times n}$$

$$\hat{Y} = A_m = g(Z_m) \quad \in R^{k \times n}$$

1.4 四、激活函数

激活函数有多种多样，本质上都是为了进行非线性组合，还有易于进行求导运算以便更新参数。此处简单介绍几种激活函数

1.4.1 1.sigmoid 函数

Sigmoid 函数已经在 logistic 回归中介绍过：

$$sigmoid(z) = \frac{1}{1+e^{-z}}$$

它是一种较早期的激活函数，现在多用于最后输出层的激活而不用在隐藏层中，这是因为当 x 远离原点时它的梯度会非常接近 0，会造成非常著名的 ‘‘梯度消失’’ 的现象。

考虑 sigmoid 函数的导数：

$$\frac{d}{dz}sigmoid(z) = \frac{e^{-z}}{(1+e^{-z})^2}$$

当 z=0 时其梯度最大为 0.25，当神经网络的层数变深时便是指数倍地降低，这便是 ‘‘梯度消失’’ 最直观和简洁的解释。

1.4.2 2.Relu（Rectified Linear Unit, 线性整流函数）

Relu 也曾是红极一时的激活函数，因其简洁的函数形式和导数形式（x 大于零导数为 1，其他情况为 0）使计算成本大大降低，但同时这也带来了神经元没有被激活的情况。这是因为当输入小于 0 时，输出和梯度都为 0，导致神经元 ‘‘死亡’’。

$$Relu(z) = max(0,z)$$

$$\frac{d}{dz}Relu(z) = \begin{cases} 1 & z > 0 \\ 0 & z \leq 0 \end{cases}$$

1.4.3 3.leaky Relu

leaky Relu 是我最喜欢的激活函数，因为它兼具了 Relu 的优点，且当输入小于零时不会出现神经元死亡的情况，k 通常的设置为 0.1。

$$leakyRelu(z,k) = max(kz,z)$$

$$\frac{d}{dz}leakyRelu(z) = \begin{cases} 1 & z > 0 \\ k & z \leq 0 \end{cases}$$

1.4.4 4.softmax

softmax 是专门用于多分类的输出层的激活函数，有两种等价形式，一种是针对 K 类有 K 个输出的线性相关的形式（即下式），另一个是针对 K 类有 K-1 个输出的线性无关的形式。

$$softmax(z) = \begin{bmatrix} \frac{e^{z_1}}{\sum_{i=1}^{k} e^{z_i}} \\ \frac{e^{z_2}}{\sum_{i=1}^{k} e^{z_i}} \\ \cdots \\ \frac{e^{z_j}}{\sum_{i=1}^{k} e^{z_i}} \\ \cdots \\ \frac{e^{z_k}}{\sum_{i=1}^{k} e^{z_i}} \end{bmatrix} = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \cdots \\ \hat{y}_i \\ \cdots \\ \hat{y}_k \end{bmatrix}$$

它的针对单一分量的偏导数形式和 sigmoid 函数极为相似：

1. 当分量出现在分母和分子上上时，我们用 a 表示和第 i 个分量无关的其他分量和：

$$\begin{aligned} \frac{d}{dz_i}\frac{e^{z_i}}{a+e^{z_i}} &= \frac{e^{z_i}(a+e^{z_i}) - e^{z_i}e^{z_i}}{(a+e^{z_i})^2} \\ &= \frac{ae^{z_i} + a^2 - a^2}{(a+e^{z_i})^2} \\ &= \frac{a(e^{z_i}+a) - a^2}{(a+e^{z_i})^2} \\ &= \frac{a}{a+e^{z_i}} - \left(\frac{a}{a+e^{z_i}}\right)^2 \\ &= \frac{a}{a+e^{z_i}}\left(1 - \frac{a}{a+e^{z_i}}\right) \\ &= \left(1 - \frac{e^{z_i}}{a+e^{z_i}}\right)\frac{e^{z_i}}{a+e^{z_i}} \\ &= (1-\hat{y}_i)\hat{y}_i \end{aligned}$$

2. 而当分量只出现在分母上时，我们用 b 表示分子上的第 j 个分量，用 a 表示与第 i、j 个分量无关的其他分量的和：

$$\begin{aligned} \frac{d}{dz_i}\frac{b}{a+b+e^{z_i}} &= \frac{-be^{z_i}}{(a+b+e^{z_i})^2} \\ &= \frac{-b(a+b+e^{z_i}) + ab + b^2}{(a+b+e^{z_i})^2} \\ &= \frac{-b}{a+b+e^{z_i}} + \frac{b(a+b)}{(a+b+e^{z_i})^2} \\ &= \frac{-b}{a+b+e^{z_i}} + \frac{b}{a+b+e^{z_i}}\left(1 - \frac{e^{z_i}}{a+b+e^{z_i}}\right) \\ &= -\hat{y}_j + \hat{y}_j(1-\hat{y}_i) \end{aligned}$$

按照矩阵的求导法则，$m \times 1$ 列向量对 $n \times 1$ 列向量求导的结果应该是 $mn \times 1$ 维向量，但是此时为了便于计算，我们将其改写成 $m \times n$ 的矩阵（或者 $n \times m$，看需求）则它的梯度为：

$$\nabla softmax(z) = \begin{bmatrix} \hat{y}_1(1-\hat{y}_1) & -\hat{y}_2+\hat{y}_2(1-\hat{y}_1) & \dots & -\hat{y}_k+\hat{y}_k(1-\hat{y}_1) \\ -\hat{y}_1+\hat{y}_1(1-\hat{y}_2) & \hat{y}_2(1-\hat{y}_2) & \dots & -\hat{y}_k+\hat{y}_k(1-\hat{y}_2) \\ \vdots & \vdots & \ddots & \vdots \\ -\hat{y}_1+\hat{y}_1(1-\hat{y}_k) & -\hat{y}_2+\hat{y}_2(1-\hat{y}_k) & \dots & \hat{y}_k(1-\hat{y}_k) \end{bmatrix}$$

1.5 五、损失函数

二分类和回归的损失函数不再赘述，和 logistic 回归和多元线性回归类似，这里介绍多分类的损失函数。

多分类的损失函数和二分类相同，也是通过似然函数进行定义：假设随机变量 Y 一共有 K 个取值，第 i 个样本对第 j 个取值的概率估计值为：

$$P(y_i = j) = \hat{y}_{ij} \ \ j = 1, 2, ..., k$$

则对 n 个样本，其似然函数为：

$$likelihood(Y, \hat{Y}) = \prod_{i=1}^{n}\prod_{j=1}^{k}\hat{y}_{ij}^{I(y_i=j)}$$

对其求自然对数，除以样本数进行标准化取负数：

$$loss(Y, \hat{Y}) = -\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{k}I(y_i = j)ln(\hat{y}_{ij})$$

这就是最终的损失函数。

1.6 六、Backward propagation 反向传播

反向传播是神经网络更新参数最经典也是最有效、最具有广泛性的算法。

反向传播的基础仍然是梯度下降法。

1.6.1 1. 输出层到最后一个隐藏层的反向传播

$$\frac{\partial}{\partial y_{ij}}loss = -\frac{1}{n}\frac{I(y_i = j)}{\hat{y}_{ij}}$$

由于输出是 $\hat{Y} \in R^{k\times n}$ 向量，损失对输出层的梯度和输出保持一致的维度：

$$\frac{\nabla loss}{\nabla \hat{Y}} = -\frac{1}{n}\begin{bmatrix} \frac{I(y_1=1)}{\hat{y}_{11}} & \dots & \frac{I(y_1=k)}{\hat{y}_{1k}} \\ & \ddots & \vdots \\ \frac{I(y_n=1)}{\hat{y}_{n1}} & & \frac{I(y_n=k)}{\hat{y}_{nk}} \end{bmatrix}^T \in R^{k\times n}$$

对输出层的激活函数有 $\hat{Y}_i = g(Z_i) = softmax(Z_i) \in R^{k\times 1}$，$Z_i \in R^{k\times 1}$

$$\frac{\nabla \hat{Y}}{\nabla Z_i} = \begin{bmatrix} \hat{y}_{i1}(1-\hat{y}_{i1}) & -\hat{y}_{i2}+\hat{y}_{i2}(1-\hat{y}_{i1}) & \dots & -\hat{y}_{ik}+\hat{y}_{ik}(1-\hat{y}_{i1}) \\ -\hat{y}_{i1}+\hat{y}_{i1}(1-\hat{y}_{i2}) & \hat{y}_{i2}(1-\hat{y}_{i2}) & \dots & -\hat{y}_{ik}+\hat{y}_{ik}(1-\hat{y}_{i2}) \\ \vdots & \vdots & \ddots & \vdots \\ -\hat{y}_{i1}+\hat{y}_{i1}(1-\hat{y}_{ik}) & -\hat{y}_{i2}+\hat{y}_{i2}(1-\hat{y}_{ik}) & \dots & \hat{y}_{ik}(1-\hat{y}_{ik}) \end{bmatrix}$$

将这个梯度矩阵乘以 $\frac{\nabla loss}{\nabla \hat{Y}}$ 与之对应的列，如果 $y_i = j$ 的话，这一列将是：

$$\begin{bmatrix} -\hat{y}_{i1} \\ \dots \\ -\hat{y}_{ij-1} \\ 1-\hat{y}_{ij} \\ -\hat{y}_{ij+1} \\ \dots \\ -\hat{y}_{ik} \end{bmatrix}$$

结合损失对估计值的梯度前的系数 $-\frac{1}{n}$ 于是恰巧有：

$$\frac{\nabla loss}{\nabla Z_m} = \frac{1}{n}(\hat{Y} - Y)$$

此时还未涉及到参数的更新，而 $Z_m = W_m A_{m-1} + b_m$ 中 $W_m \in R^{p_m = k\times p_{m-1}}$、$b_m \in R^{k\times 1}$ 均为参数。

$$\frac{\nabla Z_m}{\nabla W_m} = A_{m-1} \in R^{p_{m-1}\times n}$$

$$\frac{\nabla Z_m}{\nabla b_m} = 1^{1\times n} \in R^{1\times n}$$

将其和之前的梯度结合起来：

$$\frac{\nabla loss}{\nabla W_m} = \frac{\nabla loss}{\nabla Z_m}\left(\frac{\nabla Z_m}{\nabla W_m}\right)^T = \frac{\nabla loss}{\nabla Z_m} A_{m-1}^T \in R^{p_m = k\times p_{m-1}}$$
$$\frac{\nabla loss}{\nabla b_m} = \frac{\nabla loss}{\nabla Z_m} 1^{n\times 1} \in R^{k\times 1}$$

1.6.2 2. 第 i 层到第 i-1 层的反向传播

从第 i 层到第 i-1 层的反向传播和从输出层到最后一个隐藏层的推导相似：
假设 $\frac{\nabla loss}{\nabla Z_i} \in R^{p_i \times n}$ 已知，又 $Z_i = W_i A_{i-1} + b_i$，

$$\frac{\nabla Z_i}{\nabla A_{i-1}} = W_i^T \in R^{p_{i-1}\times p_i}$$

$$A_{i-1} = \begin{bmatrix} \alpha(z_{11}) & \dots & \alpha(z_{1p_i}) \\ & \ddots & \vdots \\ \alpha(z_{n1}) & & \alpha(z_{np_i}) \end{bmatrix} \in R^{p_{i-1}\times n}$$

$$\frac{\nabla A_{i-1}}{\nabla Z_{i-1}} = \begin{bmatrix} \alpha'(z_{11}) & \dots & \alpha'(z_{1p_i}) \\ & \ddots & \vdots \\ \alpha'(z_{n1}) & & \alpha'(z_{np_i}) \end{bmatrix} \in R^{p_{i-1}\times n}$$

其余部分和之前的相同

$$\frac{\nabla Z_{i-1}}{\nabla W_{i-1}} = A_{i-2} \in R^{p_{i-2}\times n}$$

$$\frac{\nabla Z_{i-1}}{\nabla b_{i-1}} = 1^{1\times n} \in R^{1\times n}$$

于是有：

$$\frac{\nabla loss}{\nabla Z_{i-1}} = W_i^T \frac{\nabla loss}{\nabla Z_i} * \frac{\nabla A_{i-1}}{\nabla Z_{i-1}} \in R^{p_{i-1\times n}}$$

$$\frac{\nabla loss}{\nabla W_{i-1}} = \frac{\nabla loss}{\nabla Z_{i-1}} A_{i-2}^T \in R^{p_{i-1}\times p_{i-2}}$$

$$\frac{\nabla loss}{\nabla b_{i-1}} = \frac{\nabla loss}{\nabla Z_{i-1}} 1^{n\times 1} \in R^{p_{i-1}\times 1}$$

1.6.3 3. 从第一层到输入层

从第 1 层到输入层的反向传播和从第 i 层到第 i-1 层的推导相似，区别在于输入是固定的数据，而不再是激活函数值：

假设 $\frac{\nabla loss}{\nabla Z_2} \in R^{p_2\times n}$ 已知，又 $Z_2 = W_2 A_1 + b_2$，$Z_1 = W_1 X + b_1$

$$\frac{\nabla loss}{\nabla Z_1} = W_2^T \frac{\nabla loss}{\nabla Z_2} * \frac{\nabla A_1}{\nabla Z_1} \in R^{p_{1\times n}}$$

$$\frac{\nabla loss}{\nabla W_1} = \frac{\nabla loss}{\nabla Z_1} X^T \in R^{p_1\times p_0}$$

$$\frac{\nabla loss}{\nabla b_1} = \frac{\nabla loss}{\nabla Z_1} 1^{n\times 1} \in R^{p_1\times 1}$$

1.7 七、优化器

优化器是指优化得到参数的方法，优化器基本都是基于梯度下降方法。如果你在线性回归中不用正规方程求解参数，而是用梯度下降，你会发现随着梯度不断下降，**梯度不断减小**。而这还不是最麻烦的问题，由于线性回归是凸优化，用梯度下降总会收敛到最小值，而神经网络多是非凸问题，梯度下降很可能会困在局部极值**无法收敛**。而且通常神经网络需要很多的数据进行训练，如果每次都像传统的梯度下降那样把所有数据都传入模型，则**计算成本很大**。

这里先介绍 **SGD**（Stochastic Gradient Descnet，随机梯度下降）优化器。

SGD 不再把所有的数据都用来进行梯度下降，而是只用小批量（**mini batch**）数据进行梯度下降，常见的选择是从 2 的 4 次方（16）到 2 的 10 次方之间，选用 2 的整数次方是根据计算机比特的特点决定的，而之前推导中梯度进行标准化时除以样本数，此时需要除以一批量的样本数。

控制梯度下降停止的条件也有所改变，由于神经网络强大的非线性组合能力，训练到收敛会造成过拟合，于是神经网络中用到最多的是早停法，也即小批量进行训练时将全部样本循环数遍（**epoch**）后就立即停下，避免过拟合。

1.8 八、应用

这次采用的是 minist 手写数字数据集，从 kaggle 的入门赛下载下来的训练数据集，有兴趣的可以把自己训练好的型跑一下 kaggle 上的测试数据集提交一下看看分数。（排名就不必了看了...）

In [1]:
```python
import pandas as pd
import numpy as np

train_data = pd.read_csv('data_set/minist.csv')
train_data.head()
```

Out[1]:
```
   label  pixel0  pixel1  pixel2  pixel3  pixel4  pixel5  pixel6  pixel7  \
0      1       0       0       0       0       0       0       0       0
1      0       0       0       0       0       0       0       0       0
2      1       0       0       0       0       0       0       0       0
3      4       0       0       0       0       0       0       0       0
4      0       0       0       0       0       0       0       0       0

   pixel8  ...  pixel774  pixel775  pixel776  pixel777  pixel778  pixel779  \
0       0  ...         0         0         0         0         0         0
1       0  ...         0         0         0         0         0         0
2       0  ...         0         0         0         0         0         0
3       0  ...         0         0         0         0         0         0
4       0  ...         0         0         0         0         0         0

   pixel780  pixel781  pixel782  pixel783
0         0         0         0         0
```

```
1        0        0        0        0
2        0        0        0        0
3        0        0        0        0
4        0        0        0        0

[5 rows x 785 columns]
```

```
In [2]: n = train_data.shape[0]
        np.random.seed(2099)
        index = np.random.permutation(n)
        train_index = index[0: int(0.7*n)]
        test_index = index[int(0.7*n): n]

        test_data = train_data.iloc[test_index]
        test_label = test_data['label']
        del test_data['label']
        test_data = np.array(test_data)
        test_label = np.array(test_label).reshape([n-int(0.7*n), 1])

        train_data = train_data.iloc[train_index]
        train_label = train_data['label']
        del train_data['label']
        train_data = np.array(train_data)
        train_label = np.array(train_label).reshape([int(0.7*n), 1])
```

```
In [3]: def to_category(label, num_classes):
            n = label.shape[0]
            tmp = np.zeros([n, num_classes])
            j = 0
            for i in label:
                tmp[j, i]=1
                j += 1
            return tmp
```

```
In [4]: def soft_max(z):
            """
            :param z: input, an p*n matrix
            :return: p*n matrix
            """
            e = np.exp(z)
            total = np.sum(e, axis=0, keepdims=True)
```

```
        weight = e / total
        return weight
```

```
In [5]: def accuracy(y, y_hat):
            y = np.argmax(y, axis=0)
            y_hat = np.argmax(y_hat, axis=0)
            return sum(y == y_hat)/len(y)
```

```
In [6]: def loss(y, y_hat):
            """
            :param y: the ture value p*n matrix
            :param y_hat: the predicted value
            :return: loss, it's quite computationally expensive
                        I once simplify it as np.sum(y*y_hat)
            """
            n = y.shape[1]
            tmp = y_hat**y
            tmp = -np.log(tmp.prod(axis=0)).sum()/n
            return tmp
```

```
In [7]: def leaky_relu(x, k=0.3):
            return (x > 0)*x + k*(x < 0)*x



        def d_leaky_relu(x, k=0.3):
            return (x > 0) + k*(x < 0)
```

```
In [8]: def init_w(b, a):
            w = np.random.randn(a * b)
            w = np.reshape(w, [b, a])
            return w



        def init_b(b):
            b = np.zeros([b, 1])
            return b
```

```
In [9]: def forward(x, parameter, cache):
            """
            :param x:input data p*n matrix
            :param parameter: a dict storing parameters
            :param cache: a dict storing computation result of each layer
```

```
        :return: the predicted value
        """
        cache['C1'] = np.dot(parameter['W1'], x) + parameter['b1']
        cache['A1'] = leaky_relu(cache['C1'])
        cache['C2'] = np.dot(parameter['W2'], cache['A1']) + parameter['b2']
        cache['A2'] = leaky_relu(cache['C2'])
        cache['C3'] = np.dot(parameter['W3'], cache['A2']) + parameter['b3']
        cache['A3'] = soft_max(cache['C3'])
        return cache
```

```
In [10]: def back_propagation(x, y, parameter, cache, step):
             """
             X 784*n / Y 10*n
             dW1 W1 800*784, db1 b1 800*1, A1 C1, 800*n
             dW2 W2 400*800, db2 b2 400*1, A2 C2, 400*n
             dW3 W3 10*400, db3 b3 10*1, A3 C3, 10*n
             :param y: true value
             :param parameter: dictionary storing all parameters
             :param cache: dictionary storing all the computation in process
             :param step: learning rate
             :return: updated parameters
             """
             number = y.shape[1]
             cache['dC3'] = cache['A3'] - y  # 10*n
             cache['dW3'] = np.dot(cache['dC3'], cache['A2'].T)/number  # 10*400
             cache['db3'] = np.sum(cache['dC3'], axis=1, keepdims=True)/number  # 10*1
             parameter['W3'] = parameter['W3'] - step*cache['dW3']  # 10*400
             parameter['b3'] = parameter['b3'] - step*cache['db3']  # 10*1

             cache['dC2'] = np.dot(parameter['W3'].T,
                                   cache['dC3'])*d_leaky_relu(cache['C2'])  # 400*n
             cache['dW2'] = np.dot(cache['dC2'], cache['A1'].T)/number  # 400*800
             cache['db2'] = np.sum(cache['dC2'], axis=1, keepdims=True)/number  # 400*1
             parameter['W2'] = parameter['W2'] - step*cache['dW2']  # 400*800
             parameter['b2'] = parameter['b2'] - step*cache['db2']  # 400*1

             cache['dC1'] = np.dot(parameter['W2'].T,
                                   cache['dC2'])*d_leaky_relu(cache['C1'])  # 800*n
             cache['dW1'] = np.dot(cache['dC1'], x.T)/number  # 800*784
             cache['db1'] = np.sum(cache['dC1'], axis=1, keepdims=True)  # 800*1
```

```
        parameter['W1'] = parameter['W1'] - step*cache['dW1']  # 800*784
        parameter['b1'] = parameter['b1'] - step*cache['db1']  # 800*1
        return cache, parameter
```

```
In [11]: def train(x, y, learning_rate=0.001, batch_size=128, epoch=5):
             """
             :param x: training data
             :param y: training label
             :param learning_rate: the length of a step
             :param batch_size: numbers of samples we train in a round
             :param epoch: rounds we train through training data
             :return: a trained set of parameters
             """
             parameter = dict()
             nx = x.shape[1]
             parameter['W1'] = init_w(800, 784)/100
             parameter['b1'] = init_b(800)
             parameter['W2'] = init_w(400, 800)/100
             parameter['b2'] = init_b(400)
             parameter['W3'] = init_w(10, 400)/100
             parameter['b3'] = init_b(10)

             index = np.array([], dtype='int')
             for i in range(0, nx, batch_size):
                 index = np.append(index, i)
             index = np.append(index, nx)

             cache = dict()
             for i in range(0, epoch):
                 for j in range(0, int(nx/batch_size)+1):
                     one_batch_x = x[:, index[j]:index[j+1]]
                     one_batch_y = y[:, index[j]:index[j+1]]
                     cache = forward(one_batch_x, parameter, cache)
                     prob = loss(one_batch_y, cache['A3'])
                     acc = accuracy(one_batch_y, cache['A3'])
                     print(str(i)+'--'+str(j)+'--'+str(index[j+1]))
                     print('loss: '+str(prob))
                     print('accuracy: '+str(acc))
                     [cache, parameter] = back_propagation(one_batch_x, one_batch_y,
                                                           parameter, cache, step=learning_rate)
```

```
            return cache, parameter
```

In [12]:
```
train_label = to_category(train_label, num_classes=10)
test_label = to_category(test_label, num_classes=10)

print(train_label.shape)
print(test_label.shape)
```

```
(29399, 10)
(12601, 10)
```

In [13]:
```
cache, parameter = train(x=train_data.T, y=train_label.T, epoch=5)
```

```
0--0--128
loss: 2.5144562796833654
accuracy: 0.0703125
0--1--256
loss: 2.217050238032897
accuracy: 0.1875
0--2--384
loss: 2.1658027472486903
accuracy: 0.234375
0--3--512
loss: 2.0168412770294104
accuracy: 0.296875
0--4--640
loss: 1.8946818751163836
accuracy: 0.40625
0--5--768
loss: 1.8244820246098055
accuracy: 0.421875
0--6--896
loss: 1.6413233158646316
accuracy: 0.515625
0--7--1024
loss: 1.6209324889136416
accuracy: 0.5546875
0--8--1152
loss: 1.5990883453285492
accuracy: 0.5390625
0--9--1280
```

```
loss: 1.385634792789608
accuracy: 0.6484375
0--10--1408
loss: 1.303833712288761
accuracy: 0.703125
0--11--1536
loss: 1.3310439935553542
accuracy: 0.6484375
0--12--1664
loss: 1.1805366474892716
accuracy: 0.7265625
0--13--1792
loss: 1.260705592177902
accuracy: 0.6640625
0--14--1920
loss: 1.1617022979770701
accuracy: 0.7109375
0--15--2048
loss: 1.1734547644587505
accuracy: 0.6875
0--16--2176
loss: 1.0308806471718963
accuracy: 0.734375
0--17--2304
loss: 1.0419973584182476
accuracy: 0.71875
0--18--2432
loss: 0.982214182089478
accuracy: 0.78125
0--19--2560
loss: 1.0488591875437283
accuracy: 0.7421875
0--20--2688
loss: 1.0342548244495975
accuracy: 0.6875
0--21--2816
loss: 1.0185086659398621
accuracy: 0.765625
0--22--2944
```

```
loss: 1.051002820444534
accuracy: 0.7109375
0--23--3072
loss: 0.9282233902726331
accuracy: 0.7109375
0--24--3200
loss: 0.8744240635188474
accuracy: 0.7890625
0--25--3328
loss: 0.8244569494972284
accuracy: 0.7890625
0--26--3456
loss: 0.7977779804797371
accuracy: 0.78125
0--27--3584
loss: 0.7676662447960514
accuracy: 0.8125
0--28--3712
loss: 0.7796647625908661
accuracy: 0.8046875
0--29--3840
loss: 0.7544081717094397
accuracy: 0.7890625
0--30--3968
loss: 0.8070503639703792
accuracy: 0.7421875
0--31--4096
loss: 0.7655405946902787
accuracy: 0.796875
0--32--4224
loss: 0.8616568699672724
accuracy: 0.75
0--33--4352
loss: 0.7552023431289551
accuracy: 0.8046875
0--34--4480
loss: 0.7433222356628701
accuracy: 0.84375
0--35--4608
```

```
loss: 0.7532108013910165
accuracy: 0.796875
0--36--4736
loss: 0.814190428647501
accuracy: 0.7890625
0--37--4864
loss: 0.6220072638473819
accuracy: 0.8125
0--38--4992
loss: 0.6470966796465027
accuracy: 0.84375
0--39--5120
loss: 0.7499685064324348
accuracy: 0.78125
0--40--5248
loss: 0.7315055664845771
accuracy: 0.8125
0--41--5376
loss: 0.5727038702302014
accuracy: 0.875
0--42--5504
loss: 0.6847071274302589
accuracy: 0.8203125
0--43--5632
loss: 0.7021457761788055
accuracy: 0.78125
0--44--5760
loss: 0.6151359921160389
accuracy: 0.8359375
0--45--5888
loss: 0.7613747692387356
accuracy: 0.8046875
0--46--6016
loss: 0.6592947467208227
accuracy: 0.8359375
0--47--6144
loss: 0.5879875441592406
accuracy: 0.8515625
0--48--6272
```

```
loss: 0.6835816107099395
accuracy: 0.84375
0--49--6400
loss: 0.6529306702483856
accuracy: 0.828125
0--50--6528
loss: 0.6633618750802753
accuracy: 0.8125
0--51--6656
loss: 0.5058631434980134
accuracy: 0.8828125
0--52--6784
loss: 0.6311525600106633
accuracy: 0.859375
0--53--6912
loss: 0.5339073062733845
accuracy: 0.84375
0--54--7040
loss: 0.5454118112789327
accuracy: 0.875
0--55--7168
loss: 0.6613946913247553
accuracy: 0.8203125
0--56--7296
loss: 0.5332989237884944
accuracy: 0.859375
0--57--7424
loss: 0.6223232868020128
accuracy: 0.8046875
0--58--7552
loss: 0.604196518083128
accuracy: 0.8515625
0--59--7680
loss: 0.613419671781042
accuracy: 0.7890625
0--60--7808
loss: 0.4894249692370229
accuracy: 0.875
0--61--7936
```

```
loss: 0.657183273231297
accuracy: 0.8046875
0--62--8064
loss: 0.5421928692676916
accuracy: 0.875
0--63--8192
loss: 0.5087934491958308
accuracy: 0.8828125
0--64--8320
loss: 0.5005148366749299
accuracy: 0.8671875
0--65--8448
loss: 0.4195915562480516
accuracy: 0.90625
0--66--8576
loss: 0.46682100140049426
accuracy: 0.8828125
0--67--8704
loss: 0.5070714440826076
accuracy: 0.84375
0--68--8832
loss: 0.5842310285878669
accuracy: 0.8125
0--69--8960
loss: 0.57634229921275
accuracy: 0.796875
0--70--9088
loss: 0.4528906176739569
accuracy: 0.8984375
0--71--9216
loss: 0.5432561850787224
accuracy: 0.8515625
0--72--9344
loss: 0.6901474553130629
accuracy: 0.75
0--73--9472
loss: 0.5315023767273102
accuracy: 0.8046875
0--74--9600
```

```
loss: 0.6398640322165798
accuracy: 0.8125
0--75--9728
loss: 0.49572381231156143
accuracy: 0.84375
0--76--9856
loss: 0.6074509097457453
accuracy: 0.8046875
0--77--9984
loss: 0.5324287706136872
accuracy: 0.8671875
0--78--10112
loss: 0.48572787122398214
accuracy: 0.875
0--79--10240
loss: 0.6006242921716833
accuracy: 0.8515625
0--80--10368
loss: 0.5951739834223229
accuracy: 0.8203125
0--81--10496
loss: 0.48336886031493415
accuracy: 0.8671875
0--82--10624
loss: 0.5392963149271843
accuracy: 0.828125
0--83--10752
loss: 0.45834187019552136
accuracy: 0.9140625
0--84--10880
loss: 0.46805463326844376
accuracy: 0.859375
0--85--11008
loss: 0.4195068700496333
accuracy: 0.8984375
0--86--11136
loss: 0.4227066246891506
accuracy: 0.8671875
0--87--11264
```

```
loss: 0.4943372348738676
accuracy: 0.859375
0--88--11392
loss: 0.5411924886369339
accuracy: 0.8671875
0--89--11520
loss: 0.45092554113525873
accuracy: 0.859375
0--90--11648
loss: 0.464912130932998
accuracy: 0.8828125
0--91--11776
loss: 0.4478056670775561
accuracy: 0.890625
0--92--11904
loss: 0.5370311291331553
accuracy: 0.8359375
0--93--12032
loss: 0.5630288479275366
accuracy: 0.8203125
0--94--12160
loss: 0.4171803691474189
accuracy: 0.859375
0--95--12288
loss: 0.46524645357264005
accuracy: 0.8515625
0--96--12416
loss: 0.45359123350166575
accuracy: 0.90625
0--97--12544
loss: 0.5152949188394549
accuracy: 0.859375
0--98--12672
loss: 0.3421918275751196
accuracy: 0.8828125
0--99--12800
loss: 0.4812717036524752
accuracy: 0.875
0--100--12928
```

```
loss: 0.43185703545621723
accuracy: 0.8671875
0--101--13056
loss: 0.4459955408035881
accuracy: 0.875
0--102--13184
loss: 0.3948813414634826
accuracy: 0.8984375
0--103--13312
loss: 0.5328456963084235
accuracy: 0.8359375
0--104--13440
loss: 0.5358278780806194
accuracy: 0.8359375
0--105--13568
loss: 0.5926010910582731
accuracy: 0.84375
0--106--13696
loss: 0.4149401024217997
accuracy: 0.890625
0--107--13824
loss: 0.47152444972215724
accuracy: 0.8203125
0--108--13952
loss: 0.2783783482785894
accuracy: 0.9296875
0--109--14080
loss: 0.4414253984259714
accuracy: 0.890625
0--110--14208
loss: 0.534016753920909
accuracy: 0.828125
0--111--14336
loss: 0.453277161279315
accuracy: 0.8984375
0--112--14464
loss: 0.4059201480843446
accuracy: 0.90625
0--113--14592
```

```
loss: 0.4527460669847329
accuracy: 0.8828125
0--114--14720
loss: 0.38021177738584777
accuracy: 0.8828125
0--115--14848
loss: 0.4659196956844336
accuracy: 0.84375
0--116--14976
loss: 0.46408500854516893
accuracy: 0.90625
0--117--15104
loss: 0.3973907757585733
accuracy: 0.890625
0--118--15232
loss: 0.4373244497774972
accuracy: 0.859375
0--119--15360
loss: 0.2861079983495005
accuracy: 0.9453125
0--120--15488
loss: 0.39743434843571407
accuracy: 0.8828125
0--121--15616
loss: 0.444039390421915
accuracy: 0.859375
0--122--15744
loss: 0.5383207455162855
accuracy: 0.8046875
0--123--15872
loss: 0.43430907616143255
accuracy: 0.890625
0--124--16000
loss: 0.42499198428230667
accuracy: 0.8984375
0--125--16128
loss: 0.33490809940522326
accuracy: 0.9140625
0--126--16256
```

```
loss: 0.39876944211092635
accuracy: 0.90625
0--127--16384
loss: 0.33841905219324436
accuracy: 0.890625
0--128--16512
loss: 0.43014093013278887
accuracy: 0.8671875
0--129--16640
loss: 0.44087146117967435
accuracy: 0.8984375
0--130--16768
loss: 0.4149286610319117
accuracy: 0.8984375
0--131--16896
loss: 0.5487121150878685
accuracy: 0.84375
0--132--17024
loss: 0.48033160960419186
accuracy: 0.8671875
0--133--17152
loss: 0.2827652008103738
accuracy: 0.9375
0--134--17280
loss: 0.5280190354445662
accuracy: 0.84375
0--135--17408
loss: 0.40136335386886063
accuracy: 0.90625
0--136--17536
loss: 0.43412777178162065
accuracy: 0.90625
0--137--17664
loss: 0.36986659277996203
accuracy: 0.890625
0--138--17792
loss: 0.36390656067434607
accuracy: 0.90625
0--139--17920
```

```
loss: 0.37277113417056407
accuracy: 0.875
0--140--18048
loss: 0.3281691140953412
accuracy: 0.921875
0--141--18176
loss: 0.30284969809562756
accuracy: 0.9375
0--142--18304
loss: 0.3724772248149252
accuracy: 0.8984375
0--143--18432
loss: 0.4380878591385008
accuracy: 0.8828125
0--144--18560
loss: 0.37506361842601255
accuracy: 0.8984375
0--145--18688
loss: 0.39088998228135197
accuracy: 0.90625
0--146--18816
loss: 0.35999429639693986
accuracy: 0.90625
0--147--18944
loss: 0.4114326298069638
accuracy: 0.890625
0--148--19072
loss: 0.41031210578035787
accuracy: 0.890625
0--149--19200
loss: 0.46009855033299535
accuracy: 0.8671875
0--150--19328
loss: 0.3394504977432337
accuracy: 0.8984375
0--151--19456
loss: 0.3800270390012476
accuracy: 0.8984375
0--152--19584
```

```
loss: 0.3947237096393335
accuracy: 0.859375
0--153--19712
loss: 0.3841021629213581
accuracy: 0.90625
0--154--19840
loss: 0.3205243187648902
accuracy: 0.890625
0--155--19968
loss: 0.4992907784343232
accuracy: 0.859375
0--156--20096
loss: 0.4204976871912581
accuracy: 0.8828125
0--157--20224
loss: 0.47473556478157025
accuracy: 0.890625
0--158--20352
loss: 0.3297899245674878
accuracy: 0.9140625
0--159--20480
loss: 0.5765076515144314
accuracy: 0.8515625
0--160--20608
loss: 0.5689105012752617
accuracy: 0.859375
0--161--20736
loss: 0.40266185415624056
accuracy: 0.9296875
0--162--20864
loss: 0.4079488638150572
accuracy: 0.890625
0--163--20992
loss: 0.40792975673851656
accuracy: 0.875
0--164--21120
loss: 0.2872802159388109
accuracy: 0.9140625
0--165--21248
```

```
loss: 0.28309465528930744
accuracy: 0.9140625
0--166--21376
loss: 0.4824761618666852
accuracy: 0.8671875
0--167--21504
loss: 0.44351047278586536
accuracy: 0.890625
0--168--21632
loss: 0.3832241958223705
accuracy: 0.90625
0--169--21760
loss: 0.3769097063798208
accuracy: 0.890625
0--170--21888
loss: 0.39200909359288477
accuracy: 0.8515625
0--171--22016
loss: 0.3641771654429299
accuracy: 0.875
0--172--22144
loss: 0.5023609432674898
accuracy: 0.8359375
0--173--22272
loss: 0.3721557937065077
accuracy: 0.8828125
0--174--22400
loss: 0.3576019482144596
accuracy: 0.90625
0--175--22528
loss: 0.35543354651101966
accuracy: 0.921875
0--176--22656
loss: 0.31797409572887086
accuracy: 0.9140625
0--177--22784
loss: 0.3133741169958173
accuracy: 0.921875
0--178--22912
```

```
loss: 0.3121597299498443
accuracy: 0.921875
0--179--23040
loss: 0.3400229507354354
accuracy: 0.9296875
0--180--23168
loss: 0.32361553850315117
accuracy: 0.9296875
0--181--23296
loss: 0.4994068685631748
accuracy: 0.859375
0--182--23424
loss: 0.3665033227083895
accuracy: 0.9140625
0--183--23552
loss: 0.3808076015139829
accuracy: 0.8671875
0--184--23680
loss: 0.41112087109827855
accuracy: 0.859375
0--185--23808
loss: 0.25294339447569847
accuracy: 0.9296875
0--186--23936
loss: 0.43550023073241034
accuracy: 0.859375
0--187--24064
loss: 0.5443954670852986
accuracy: 0.875
0--188--24192
loss: 0.3275130711997787
accuracy: 0.8828125
0--189--24320
loss: 0.4506671024664212
accuracy: 0.8828125
0--190--24448
loss: 0.36283389855458203
accuracy: 0.890625
0--191--24576
```

```
loss: 0.4840784045475447
accuracy: 0.84375
0--192--24704
loss: 0.3273309235119446
accuracy: 0.921875
0--193--24832
loss: 0.36902547357019627
accuracy: 0.875
0--194--24960
loss: 0.38617889969368624
accuracy: 0.90625
0--195--25088
loss: 0.3240239176367957
accuracy: 0.9140625
0--196--25216
loss: 0.399083965780036
accuracy: 0.8828125
0--197--25344
loss: 0.44483693771873845
accuracy: 0.8828125
0--198--25472
loss: 0.5084765802822988
accuracy: 0.875
0--199--25600
loss: 0.3816273765758412
accuracy: 0.875
0--200--25728
loss: 0.33549340057174243
accuracy: 0.90625
0--201--25856
loss: 0.42596543164542394
accuracy: 0.84375
0--202--25984
loss: 0.23221470966040078
accuracy: 0.9453125
0--203--26112
loss: 0.3383040850454929
accuracy: 0.9296875
0--204--26240
```

```
loss: 0.28554657833485925
accuracy: 0.9296875
0--205--26368
loss: 0.467886174789014
accuracy: 0.859375
0--206--26496
loss: 0.34802149240598346
accuracy: 0.90625
0--207--26624
loss: 0.3078923210986365
accuracy: 0.9140625
0--208--26752
loss: 0.4087322367446533
accuracy: 0.8828125
0--209--26880
loss: 0.3673793402271704
accuracy: 0.8984375
0--210--27008
loss: 0.3055128566212699
accuracy: 0.9140625
0--211--27136
loss: 0.3459867854151793
accuracy: 0.8984375
0--212--27264
loss: 0.3914355415593235
accuracy: 0.9140625
0--213--27392
loss: 0.4220776143540002
accuracy: 0.875
0--214--27520
loss: 0.22968103295881812
accuracy: 0.9453125
0--215--27648
loss: 0.41668449968419236
accuracy: 0.8671875
0--216--27776
loss: 0.32720253762644524
accuracy: 0.890625
0--217--27904
```

```
loss: 0.31349499399099046
accuracy: 0.90625
0--218--28032
loss: 0.2527009781201379
accuracy: 0.921875
0--219--28160
loss: 0.3087098991516306
accuracy: 0.8984375
0--220--28288
loss: 0.43902029596589703
accuracy: 0.875
0--221--28416
loss: 0.3548413974404646
accuracy: 0.921875
0--222--28544
loss: 0.3450215101025608
accuracy: 0.890625
0--223--28672
loss: 0.24524699720009102
accuracy: 0.9296875
0--224--28800
loss: 0.46874000176229375
accuracy: 0.859375
0--225--28928
loss: 0.3578855945733159
accuracy: 0.890625
0--226--29056
loss: 0.3644711639459468
accuracy: 0.8828125
0--227--29184
loss: 0.42358075433792436
accuracy: 0.8828125
0--228--29312
loss: 0.3624121500369032
accuracy: 0.9140625
0--229--29399
loss: 0.3521769928408441
accuracy: 0.896551724137931
1--0--128
```

```
loss: 0.4784413933361167
accuracy: 0.8671875
1--1--256
loss: 0.2809984040435429
accuracy: 0.90625
1--2--384
loss: 0.5224073266557672
accuracy: 0.859375
1--3--512
loss: 0.3235940035634032
accuracy: 0.890625
1--4--640
loss: 0.41865132988449877
accuracy: 0.8515625
1--5--768
loss: 0.3722485682686894
accuracy: 0.8828125
1--6--896
loss: 0.3577078536245744
accuracy: 0.90625
1--7--1024
loss: 0.3102148202754932
accuracy: 0.90625
1--8--1152
loss: 0.2817566739136048
accuracy: 0.90625
1--9--1280
loss: 0.37031080057203847
accuracy: 0.8671875
1--10--1408
loss: 0.26918610721712893
accuracy: 0.9140625
1--11--1536
loss: 0.34609346389358664
accuracy: 0.875
1--12--1664
loss: 0.27582413647427456
accuracy: 0.921875
1--13--1792
```

```
loss: 0.27419080654011907
accuracy: 0.921875
1--14--1920
loss: 0.2939112637745688
accuracy: 0.90625
1--15--2048
loss: 0.3102518988493187
accuracy: 0.8828125
1--16--2176
loss: 0.33204120765980805
accuracy: 0.9453125
1--17--2304
loss: 0.2822052515498391
accuracy: 0.921875
1--18--2432
loss: 0.33787754391845404
accuracy: 0.921875
1--19--2560
loss: 0.3884173928934721
accuracy: 0.8671875
1--20--2688
loss: 0.28647193344995925
accuracy: 0.921875
1--21--2816
loss: 0.3095969068754494
accuracy: 0.8828125
1--22--2944
loss: 0.39489087785330207
accuracy: 0.875
1--23--3072
loss: 0.34880298762089224
accuracy: 0.8671875
1--24--3200
loss: 0.3037214476423902
accuracy: 0.921875
1--25--3328
loss: 0.28530917506398634
accuracy: 0.90625
1--26--3456
```

```
loss: 0.29314643782488764
accuracy: 0.9140625
1--27--3584
loss: 0.23507186548665399
accuracy: 0.9296875
1--28--3712
loss: 0.24446196276152982
accuracy: 0.96875
1--29--3840
loss: 0.300079059540142
accuracy: 0.90625
1--30--3968
loss: 0.3138964229494293
accuracy: 0.90625
1--31--4096
loss: 0.37485050138090525
accuracy: 0.8984375
1--32--4224
loss: 0.40481754889162824
accuracy: 0.875
1--33--4352
loss: 0.3097313623292942
accuracy: 0.921875
1--34--4480
loss: 0.35920632785927703
accuracy: 0.8984375
1--35--4608
loss: 0.38687265900328316
accuracy: 0.8984375
1--36--4736
loss: 0.41814952218396384
accuracy: 0.8515625
1--37--4864
loss: 0.24292502593632326
accuracy: 0.921875
1--38--4992
loss: 0.3298984284439357
accuracy: 0.90625
1--39--5120
```

```
loss: 0.35398397172634577
accuracy: 0.8984375
1--40--5248
loss: 0.39377776996960023
accuracy: 0.8984375
1--41--5376
loss: 0.2149024317936732
accuracy: 0.9453125
1--42--5504
loss: 0.30713244667517786
accuracy: 0.9140625
1--43--5632
loss: 0.2937702218335918
accuracy: 0.8828125
1--44--5760
loss: 0.31342263311723545
accuracy: 0.921875
1--45--5888
loss: 0.5137701880222669
accuracy: 0.8828125
1--46--6016
loss: 0.38857735940757243
accuracy: 0.8671875
1--47--6144
loss: 0.2625033273357935
accuracy: 0.921875
1--48--6272
loss: 0.3146359982184206
accuracy: 0.8984375
1--49--6400
loss: 0.32417084676958396
accuracy: 0.875
1--50--6528
loss: 0.4241690693642612
accuracy: 0.859375
1--51--6656
loss: 0.20548885540668363
accuracy: 0.9453125
1--52--6784
```

```
loss: 0.3527824271211092
accuracy: 0.921875
1--53--6912
loss: 0.23284593703537668
accuracy: 0.9140625
1--54--7040
loss: 0.2779823398313023
accuracy: 0.8984375
1--55--7168
loss: 0.41709150237123177
accuracy: 0.859375
1--56--7296
loss: 0.2623367208247557
accuracy: 0.9296875
1--57--7424
loss: 0.36704233540956954
accuracy: 0.921875
1--58--7552
loss: 0.3673391609373419
accuracy: 0.9375
1--59--7680
loss: 0.41090676359551226
accuracy: 0.875
1--60--7808
loss: 0.24343609902013907
accuracy: 0.9375
1--61--7936
loss: 0.4631960912363376
accuracy: 0.8828125
1--62--8064
loss: 0.30408090548921496
accuracy: 0.9140625
1--63--8192
loss: 0.3106892270517312
accuracy: 0.9375
1--64--8320
loss: 0.29241012925739635
accuracy: 0.90625
1--65--8448
```

```
loss: 0.21049285400482148
accuracy: 0.9375
1--66--8576
loss: 0.21814655111015233
accuracy: 0.953125
1--67--8704
loss: 0.22486617728784983
accuracy: 0.9375
1--68--8832
loss: 0.3442469318671014
accuracy: 0.890625
1--69--8960
loss: 0.31405442327854927
accuracy: 0.890625
1--70--9088
loss: 0.2512756437475721
accuracy: 0.9296875
1--71--9216
loss: 0.33934086320415857
accuracy: 0.8984375
1--72--9344
loss: 0.48646391609240675
accuracy: 0.8359375
1--73--9472
loss: 0.3472198928436054
accuracy: 0.8671875
1--74--9600
loss: 0.4607340927267417
accuracy: 0.90625
1--75--9728
loss: 0.3246488577845161
accuracy: 0.8984375
1--76--9856
loss: 0.35825968686244286
accuracy: 0.875
1--77--9984
loss: 0.3049351727502574
accuracy: 0.921875
1--78--10112
```

```
loss: 0.3121919562288198
accuracy: 0.8828125
1--79--10240
loss: 0.4295420232389461
accuracy: 0.8984375
1--80--10368
loss: 0.3983019736381338
accuracy: 0.8828125
1--81--10496
loss: 0.2665409958645021
accuracy: 0.9140625
1--82--10624
loss: 0.35328706955026573
accuracy: 0.890625
1--83--10752
loss: 0.2779423227207185
accuracy: 0.9140625
1--84--10880
loss: 0.2601522169079653
accuracy: 0.9375
1--85--11008
loss: 0.21878340370101992
accuracy: 0.9296875
1--86--11136
loss: 0.25071160511379764
accuracy: 0.9375
1--87--11264
loss: 0.3591139856502097
accuracy: 0.8671875
1--88--11392
loss: 0.39672080367074136
accuracy: 0.890625
1--89--11520
loss: 0.22977180467997094
accuracy: 0.9453125
1--90--11648
loss: 0.2868004728390251
accuracy: 0.9296875
1--91--11776
```

```
loss: 0.27743950034042697
accuracy: 0.921875
1--92--11904
loss: 0.3984753527672429
accuracy: 0.8984375
1--93--12032
loss: 0.3709176959356433
accuracy: 0.8984375
1--94--12160
loss: 0.2705289419573663
accuracy: 0.9140625
1--95--12288
loss: 0.2904986149067824
accuracy: 0.9375
1--96--12416
loss: 0.33194582380738397
accuracy: 0.9453125
1--97--12544
loss: 0.37430277906733767
accuracy: 0.8984375
1--98--12672
loss: 0.1999417758621406
accuracy: 0.9375
1--99--12800
loss: 0.37162875343143553
accuracy: 0.9296875
1--100--12928
loss: 0.28508558399260997
accuracy: 0.890625
1--101--13056
loss: 0.28808935179522466
accuracy: 0.9140625
1--102--13184
loss: 0.20992930406145704
accuracy: 0.9453125
1--103--13312
loss: 0.3735091988241605
accuracy: 0.8828125
1--104--13440
```

```
loss: 0.36639050533154693
accuracy: 0.8984375
1--105--13568
loss: 0.42127941663895585
accuracy: 0.8984375
1--106--13696
loss: 0.30563803393494354
accuracy: 0.90625
1--107--13824
loss: 0.3039179067194721
accuracy: 0.8984375
1--108--13952
loss: 0.17619169041092217
accuracy: 0.9453125
1--109--14080
loss: 0.31294737397913575
accuracy: 0.8984375
1--110--14208
loss: 0.3762742150902608
accuracy: 0.875
1--111--14336
loss: 0.3657364604865767
accuracy: 0.9140625
1--112--14464
loss: 0.276047693232034
accuracy: 0.921875
1--113--14592
loss: 0.315607276896158
accuracy: 0.8984375
1--114--14720
loss: 0.2736493509320972
accuracy: 0.8984375
1--115--14848
loss: 0.3034338647475618
accuracy: 0.859375
1--116--14976
loss: 0.3678895103481914
accuracy: 0.921875
1--117--15104
```

```
loss: 0.311873122562798
accuracy: 0.921875
1--118--15232
loss: 0.32489399741379676
accuracy: 0.890625
1--119--15360
loss: 0.18223088849386548
accuracy: 0.9609375
1--120--15488
loss: 0.2775342971148901
accuracy: 0.921875
1--121--15616
loss: 0.31650307958091417
accuracy: 0.9296875
1--122--15744
loss: 0.39002690437473786
accuracy: 0.8828125
1--123--15872
loss: 0.2918239983645352
accuracy: 0.921875
1--124--16000
loss: 0.2634624238654544
accuracy: 0.9375
1--125--16128
loss: 0.19652623676170858
accuracy: 0.953125
1--126--16256
loss: 0.2770593850140747
accuracy: 0.9375
1--127--16384
loss: 0.22288115505200634
accuracy: 0.9296875
1--128--16512
loss: 0.30825426253445504
accuracy: 0.90625
1--129--16640
loss: 0.3031483368258479
accuracy: 0.90625
1--130--16768
```

```
loss: 0.2897914942258314
accuracy: 0.9296875
1--131--16896
loss: 0.3910639523172185
accuracy: 0.8671875
1--132--17024
loss: 0.417093209706786
accuracy: 0.8671875
1--133--17152
loss: 0.16981852223661753
accuracy: 0.9765625
1--134--17280
loss: 0.4153660560946515
accuracy: 0.890625
1--135--17408
loss: 0.33546864347002825
accuracy: 0.9140625
1--136--17536
loss: 0.3138535351977386
accuracy: 0.921875
1--137--17664
loss: 0.27750742012206653
accuracy: 0.9140625
1--138--17792
loss: 0.2253598944403589
accuracy: 0.9375
1--139--17920
loss: 0.24571241740187483
accuracy: 0.921875
1--140--18048
loss: 0.2312282561599559
accuracy: 0.921875
1--141--18176
loss: 0.19433485944357415
accuracy: 0.953125
1--142--18304
loss: 0.30588490027446635
accuracy: 0.90625
1--143--18432
```

```
loss: 0.34284974860461637
accuracy: 0.9140625
1--144--18560
loss: 0.27273742700070447
accuracy: 0.9140625
1--145--18688
loss: 0.2811987983786788
accuracy: 0.9375
1--146--18816
loss: 0.26033479606459353
accuracy: 0.921875
1--147--18944
loss: 0.3268586097310512
accuracy: 0.921875
1--148--19072
loss: 0.3490945240203806
accuracy: 0.890625
1--149--19200
loss: 0.3790747820865246
accuracy: 0.90625
1--150--19328
loss: 0.21662839410198875
accuracy: 0.9375
1--151--19456
loss: 0.2603595533781914
accuracy: 0.9453125
1--152--19584
loss: 0.28275774294314265
accuracy: 0.90625
1--153--19712
loss: 0.3015639863915001
accuracy: 0.921875
1--154--19840
loss: 0.21570631567943796
accuracy: 0.9296875
1--155--19968
loss: 0.3939596911595955
accuracy: 0.890625
1--156--20096
```

```
loss: 0.29940220818314184
accuracy: 0.921875
1--157--20224
loss: 0.4052023444565691
accuracy: 0.9140625
1--158--20352
loss: 0.230451142536511
accuracy: 0.9453125
1--159--20480
loss: 0.46845843504096696
accuracy: 0.875
1--160--20608
loss: 0.49584770855640486
accuracy: 0.8984375
1--161--20736
loss: 0.3213300303503326
accuracy: 0.921875
1--162--20864
loss: 0.33239711506166025
accuracy: 0.921875
1--163--20992
loss: 0.309013077594275
accuracy: 0.8984375
1--164--21120
loss: 0.20684476378976668
accuracy: 0.9453125
1--165--21248
loss: 0.1828753456277153
accuracy: 0.9453125
1--166--21376
loss: 0.36682480166017384
accuracy: 0.8984375
1--167--21504
loss: 0.37334932863872355
accuracy: 0.8984375
1--168--21632
loss: 0.3130223202982156
accuracy: 0.9296875
1--169--21760
```

```
loss: 0.2600165076204858
accuracy: 0.9296875
1--170--21888
loss: 0.30039068884056314
accuracy: 0.90625
1--171--22016
loss: 0.26650323865358494
accuracy: 0.9140625
1--172--22144
loss: 0.3711276663502726
accuracy: 0.8984375
1--173--22272
loss: 0.2760758573351142
accuracy: 0.90625
1--174--22400
loss: 0.28228296683718684
accuracy: 0.9296875
1--175--22528
loss: 0.2760608740880042
accuracy: 0.9296875
1--176--22656
loss: 0.219883844296249
accuracy: 0.921875
1--177--22784
loss: 0.21697089647196888
accuracy: 0.921875
1--178--22912
loss: 0.23386858464233243
accuracy: 0.921875
1--179--23040
loss: 0.2778967611332675
accuracy: 0.9296875
1--180--23168
loss: 0.23043263862980054
accuracy: 0.9453125
1--181--23296
loss: 0.4145018362570708
accuracy: 0.9140625
1--182--23424
```

```
loss: 0.28478993154983884
accuracy: 0.9296875
1--183--23552
loss: 0.2686199481146821
accuracy: 0.9140625
1--184--23680
loss: 0.3194117983505595
accuracy: 0.890625
1--185--23808
loss: 0.1923340955850364
accuracy: 0.9453125
1--186--23936
loss: 0.35014091951697557
accuracy: 0.8984375
1--187--24064
loss: 0.4575791472506003
accuracy: 0.875
1--188--24192
loss: 0.266304389840751
accuracy: 0.90625
1--189--24320
loss: 0.37257096639163995
accuracy: 0.8984375
1--190--24448
loss: 0.29981049085964
accuracy: 0.9140625
1--191--24576
loss: 0.38319733110183973
accuracy: 0.8828125
1--192--24704
loss: 0.24811099374840245
accuracy: 0.9296875
1--193--24832
loss: 0.29674879569454726
accuracy: 0.890625
1--194--24960
loss: 0.32085245057571016
accuracy: 0.9140625
1--195--25088
```

```
loss: 0.22514362115236686
accuracy: 0.9453125
1--196--25216
loss: 0.35118332739436897
accuracy: 0.90625
1--197--25344
loss: 0.36331189238058564
accuracy: 0.8984375
1--198--25472
loss: 0.43987897576972745
accuracy: 0.8984375
1--199--25600
loss: 0.2992579234860539
accuracy: 0.9140625
1--200--25728
loss: 0.26666395603067566
accuracy: 0.921875
1--201--25856
loss: 0.33959145954241765
accuracy: 0.875
1--202--25984
loss: 0.18778093420870265
accuracy: 0.9453125
1--203--26112
loss: 0.2865569033984523
accuracy: 0.9296875
1--204--26240
loss: 0.21680094487446883
accuracy: 0.9296875
1--205--26368
loss: 0.40039280850551406
accuracy: 0.8828125
1--206--26496
loss: 0.2612819367944928
accuracy: 0.9296875
1--207--26624
loss: 0.2174146631443534
accuracy: 0.9140625
1--208--26752
```

```
loss: 0.34990426155293114
accuracy: 0.890625
1--209--26880
loss: 0.29917253759405366
accuracy: 0.9140625
1--210--27008
loss: 0.26558348395228615
accuracy: 0.921875
1--211--27136
loss: 0.24778017494325774
accuracy: 0.90625
1--212--27264
loss: 0.32988890825283457
accuracy: 0.921875
1--213--27392
loss: 0.34300200385577806
accuracy: 0.890625
1--214--27520
loss: 0.16635328808727912
accuracy: 0.9609375
1--215--27648
loss: 0.3487208999376741
accuracy: 0.8828125
1--216--27776
loss: 0.2690853334690969
accuracy: 0.9375
1--217--27904
loss: 0.2573507878034015
accuracy: 0.953125
1--218--28032
loss: 0.18163438002118937
accuracy: 0.953125
1--219--28160
loss: 0.2488713683519324
accuracy: 0.90625
1--220--28288
loss: 0.38752866962481713
accuracy: 0.8984375
1--221--28416
```

```
loss: 0.2919803677785158
accuracy: 0.9296875
1--222--28544
loss: 0.28305228358316903
accuracy: 0.921875
1--223--28672
loss: 0.1790467076772408
accuracy: 0.9375
1--224--28800
loss: 0.4062359630325349
accuracy: 0.875
1--225--28928
loss: 0.2904809504116944
accuracy: 0.9140625
1--226--29056
loss: 0.28241731071444065
accuracy: 0.9140625
1--227--29184
loss: 0.3828682030580231
accuracy: 0.8984375
1--228--29312
loss: 0.28646853430256464
accuracy: 0.9375
1--229--29399
loss: 0.264525579871738
accuracy: 0.9425287356321839
2--0--128
loss: 0.41297480227306166
accuracy: 0.8828125
2--1--256
loss: 0.2095900004345317
accuracy: 0.90625
2--2--384
loss: 0.4498871229439299
accuracy: 0.8515625
2--3--512
loss: 0.24945104584172287
accuracy: 0.921875
2--4--640
```

```
loss: 0.35414364952705124
accuracy: 0.8828125
2--5--768
loss: 0.2939646110547809
accuracy: 0.890625
2--6--896
loss: 0.3091939706757324
accuracy: 0.9140625
2--7--1024
loss: 0.2571216924935017
accuracy: 0.9375
2--8--1152
loss: 0.20645909197731965
accuracy: 0.9375
2--9--1280
loss: 0.3104414523823378
accuracy: 0.890625
2--10--1408
loss: 0.2184377474424078
accuracy: 0.953125
2--11--1536
loss: 0.29064085525491307
accuracy: 0.90625
2--12--1664
loss: 0.22491479480563942
accuracy: 0.921875
2--13--1792
loss: 0.21133533665545923
accuracy: 0.9296875
2--14--1920
loss: 0.2423380305127414
accuracy: 0.921875
2--15--2048
loss: 0.24890249562564684
accuracy: 0.90625
2--16--2176
loss: 0.2825886696565393
accuracy: 0.9453125
2--17--2304
```

```
loss: 0.21568460776294934
accuracy: 0.9453125
2--18--2432
loss: 0.2926194451140871
accuracy: 0.9375
2--19--2560
loss: 0.32324067560597225
accuracy: 0.8671875
2--20--2688
loss: 0.21445326906897427
accuracy: 0.921875
2--21--2816
loss: 0.24953987833615443
accuracy: 0.8984375
2--22--2944
loss: 0.3165146470961461
accuracy: 0.9140625
2--23--3072
loss: 0.2774181219834073
accuracy: 0.8828125
2--24--3200
loss: 0.24805744825849274
accuracy: 0.9296875
2--25--3328
loss: 0.23212184482675405
accuracy: 0.9140625
2--26--3456
loss: 0.24363453656942102
accuracy: 0.9140625
2--27--3584
loss: 0.17555438548800228
accuracy: 0.9375
2--28--3712
loss: 0.18427846449609042
accuracy: 0.96875
2--29--3840
loss: 0.24317040363624623
accuracy: 0.9140625
2--30--3968
```

```
loss: 0.23315855760406995
accuracy: 0.9296875
2--31--4096
loss: 0.3276961566327781
accuracy: 0.90625
2--32--4224
loss: 0.34578156299598417
accuracy: 0.890625
2--33--4352
loss: 0.25333003369876583
accuracy: 0.9453125
2--34--4480
loss: 0.304330179659014
accuracy: 0.9140625
2--35--4608
loss: 0.32914965068180224
accuracy: 0.8984375
2--36--4736
loss: 0.32612436895112773
accuracy: 0.8984375
2--37--4864
loss: 0.19844040141225544
accuracy: 0.921875
2--38--4992
loss: 0.30154239358182244
accuracy: 0.9140625
2--39--5120
loss: 0.29490817403302727
accuracy: 0.9140625
2--40--5248
loss: 0.33970352800268444
accuracy: 0.9140625
2--41--5376
loss: 0.18036859563125068
accuracy: 0.9375
2--42--5504
loss: 0.23739734169580795
accuracy: 0.9296875
2--43--5632
```

```
loss: 0.20872885661170665
accuracy: 0.921875
2--44--5760
loss: 0.26648870747809295
accuracy: 0.9296875
2--45--5888
loss: 0.48318595663357655
accuracy: 0.8828125
2--46--6016
loss: 0.33411854859954526
accuracy: 0.9140625
2--47--6144
loss: 0.22126630810781886
accuracy: 0.9296875
2--48--6272
loss: 0.24762541315395326
accuracy: 0.9140625
2--49--6400
loss: 0.2618460908693576
accuracy: 0.9140625
2--50--6528
loss: 0.38869727542848564
accuracy: 0.8828125
2--51--6656
loss: 0.15288891478661581
accuracy: 0.953125
2--52--6784
loss: 0.2977632557217401
accuracy: 0.9453125
2--53--6912
loss: 0.18113345898406846
accuracy: 0.9453125
2--54--7040
loss: 0.22483104816901783
accuracy: 0.9140625
2--55--7168
loss: 0.36130081463700087
accuracy: 0.8828125
2--56--7296
```

```
loss: 0.21099641213736348
accuracy: 0.9453125
2--57--7424
loss: 0.3081717897691191
accuracy: 0.921875
2--58--7552
loss: 0.32521651014020203
accuracy: 0.9296875
2--59--7680
loss: 0.3736985536824024
accuracy: 0.8984375
2--60--7808
loss: 0.19962879155877433
accuracy: 0.953125
2--61--7936
loss: 0.4169072655482665
accuracy: 0.8984375
2--62--8064
loss: 0.2502321834652378
accuracy: 0.921875
2--63--8192
loss: 0.2769833967055244
accuracy: 0.9375
2--64--8320
loss: 0.2508444807950252
accuracy: 0.90625
2--65--8448
loss: 0.17624664860032632
accuracy: 0.9609375
2--66--8576
loss: 0.1695451298105187
accuracy: 0.9609375
2--67--8704
loss: 0.16605490863698505
accuracy: 0.96875
2--68--8832
loss: 0.2890405121043577
accuracy: 0.9140625
2--69--8960
```

```
loss: 0.24748217026866132
accuracy: 0.9140625
2--70--9088
loss: 0.20861731371885953
accuracy: 0.9375
2--71--9216
loss: 0.2825250780616856
accuracy: 0.90625
2--72--9344
loss: 0.40437825715842696
accuracy: 0.859375
2--73--9472
loss: 0.2839297651552701
accuracy: 0.890625
2--74--9600
loss: 0.3913316358280359
accuracy: 0.9140625
2--75--9728
loss: 0.2796643443951746
accuracy: 0.90625
2--76--9856
loss: 0.28306431578429436
accuracy: 0.9140625
2--77--9984
loss: 0.2528199115950023
accuracy: 0.9453125
2--78--10112
loss: 0.2703128443962728
accuracy: 0.9140625
2--79--10240
loss: 0.37333851028590603
accuracy: 0.90625
2--80--10368
loss: 0.34076243920984
accuracy: 0.9140625
2--81--10496
loss: 0.20895457743901033
accuracy: 0.9296875
2--82--10624
```

```
loss: 0.3041723296770102
accuracy: 0.8984375
2--83--10752
loss: 0.22517690243272198
accuracy: 0.921875
2--84--10880
loss: 0.21302591309926117
accuracy: 0.9609375
2--85--11008
loss: 0.16478433959066557
accuracy: 0.9765625
2--86--11136
loss: 0.2121135236552824
accuracy: 0.9453125
2--87--11264
loss: 0.32445617370757185
accuracy: 0.875
2--88--11392
loss: 0.3528661070679706
accuracy: 0.90625
2--89--11520
loss: 0.1737366836425981
accuracy: 0.96875
2--90--11648
loss: 0.22992210437707247
accuracy: 0.953125
2--91--11776
loss: 0.22804092422966646
accuracy: 0.9375
2--92--11904
loss: 0.3633538410500681
accuracy: 0.890625
2--93--12032
loss: 0.31185641448693435
accuracy: 0.890625
2--94--12160
loss: 0.22168841598813654
accuracy: 0.9296875
2--95--12288
```

```
loss: 0.24210022962972927
accuracy: 0.9453125
2--96--12416
loss: 0.2852777315806002
accuracy: 0.953125
2--97--12544
loss: 0.3300314858428248
accuracy: 0.8984375
2--98--12672
loss: 0.1695380125611306
accuracy: 0.9609375
2--99--12800
loss: 0.3444474274836723
accuracy: 0.9296875
2--100--12928
loss: 0.23445092300056292
accuracy: 0.90625
2--101--13056
loss: 0.24135903284966692
accuracy: 0.9375
2--102--13184
loss: 0.15821932983029224
accuracy: 0.9609375
2--103--13312
loss: 0.31495741678562933
accuracy: 0.9140625
2--104--13440
loss: 0.31740745583236285
accuracy: 0.8984375
2--105--13568
loss: 0.35893248500338726
accuracy: 0.9140625
2--106--13696
loss: 0.27255210574826105
accuracy: 0.90625
2--107--13824
loss: 0.24901236937188254
accuracy: 0.921875
2--108--13952
```

```
loss: 0.14918541007422717
accuracy: 0.9453125
2--109--14080
loss: 0.2657316662022928
accuracy: 0.90625
2--110--14208
loss: 0.32465846459101516
accuracy: 0.890625
2--111--14336
loss: 0.33264665150686673
accuracy: 0.921875
2--112--14464
loss: 0.23596887446704928
accuracy: 0.9296875
2--113--14592
loss: 0.27232160899680324
accuracy: 0.9296875
2--114--14720
loss: 0.24184169560747548
accuracy: 0.9140625
2--115--14848
loss: 0.25579112019798844
accuracy: 0.9140625
2--116--14976
loss: 0.3346408837307868
accuracy: 0.9296875
2--117--15104
loss: 0.2865716975225735
accuracy: 0.9296875
2--118--15232
loss: 0.2842565280211326
accuracy: 0.921875
2--119--15360
loss: 0.15538020811969724
accuracy: 0.96875
2--120--15488
loss: 0.23394176019223806
accuracy: 0.9453125
2--121--15616
```

```
loss: 0.27684753801984474
accuracy: 0.9453125
2--122--15744
loss: 0.34217822744201
accuracy: 0.8828125
2--123--15872
loss: 0.2476873090890031
accuracy: 0.9375
2--124--16000
loss: 0.20952110997057602
accuracy: 0.9453125
2--125--16128
loss: 0.16079385670478685
accuracy: 0.9609375
2--126--16256
loss: 0.23832031661395198
accuracy: 0.9375
2--127--16384
loss: 0.1819166619802945
accuracy: 0.9453125
2--128--16512
loss: 0.2632972538770256
accuracy: 0.9140625
2--129--16640
loss: 0.2538021127031236
accuracy: 0.921875
2--130--16768
loss: 0.24752862182365795
accuracy: 0.9296875
2--131--16896
loss: 0.3242916687207346
accuracy: 0.8984375
2--132--17024
loss: 0.38590803523100325
accuracy: 0.8671875
2--133--17152
loss: 0.13423707209333413
accuracy: 0.96875
2--134--17280
```

```
loss: 0.3726634636439494
accuracy: 0.9140625
2--135--17408
loss: 0.30706562614343175
accuracy: 0.921875
2--136--17536
loss: 0.2637006784032402
accuracy: 0.9296875
2--137--17664
loss: 0.2354397578666386
accuracy: 0.9296875
2--138--17792
loss: 0.18185277668011546
accuracy: 0.953125
2--139--17920
loss: 0.20712106106030054
accuracy: 0.9296875
2--140--18048
loss: 0.20028584516431272
accuracy: 0.9453125
2--141--18176
loss: 0.160319364638835
accuracy: 0.9609375
2--142--18304
loss: 0.28160381558916414
accuracy: 0.9296875
2--143--18432
loss: 0.3075902570563753
accuracy: 0.9140625
2--144--18560
loss: 0.23262041070688996
accuracy: 0.921875
2--145--18688
loss: 0.2307890658753357
accuracy: 0.9453125
2--146--18816
loss: 0.2332947331744768
accuracy: 0.9375
2--147--18944
```

```
loss: 0.29122206283057905
accuracy: 0.9375
2--148--19072
loss: 0.3226329447371223
accuracy: 0.9140625
2--149--19200
loss: 0.3484057641004431
accuracy: 0.90625
2--150--19328
loss: 0.18193152412620817
accuracy: 0.9375
2--151--19456
loss: 0.21459621045630867
accuracy: 0.9453125
2--152--19584
loss: 0.23706438327201323
accuracy: 0.921875
2--153--19712
loss: 0.26355196858347874
accuracy: 0.9296875
2--154--19840
loss: 0.17753591538695165
accuracy: 0.9609375
2--155--19968
loss: 0.33834344892620816
accuracy: 0.90625
2--156--20096
loss: 0.25595326780822314
accuracy: 0.9296875
2--157--20224
loss: 0.3601323478881924
accuracy: 0.9140625
2--158--20352
loss: 0.19457870534273003
accuracy: 0.953125
2--159--20480
loss: 0.4165983639589841
accuracy: 0.890625
2--160--20608
```

```
loss: 0.4490528555325763
accuracy: 0.8984375
2--161--20736
loss: 0.28328074415803906
accuracy: 0.921875
2--162--20864
loss: 0.30240956304893885
accuracy: 0.9375
2--163--20992
loss: 0.2568513982365574
accuracy: 0.9140625
2--164--21120
loss: 0.18131250985581687
accuracy: 0.953125
2--165--21248
loss: 0.14699429659741065
accuracy: 0.953125
2--166--21376
loss: 0.31706058902945033
accuracy: 0.90625
2--167--21504
loss: 0.3401551720161289
accuracy: 0.90625
2--168--21632
loss: 0.2844084192199285
accuracy: 0.9375
2--169--21760
loss: 0.21053867827663994
accuracy: 0.9296875
2--170--21888
loss: 0.25823802670581164
accuracy: 0.9296875
2--171--22016
loss: 0.22607249920934472
accuracy: 0.9140625
2--172--22144
loss: 0.3038543092175032
accuracy: 0.921875
2--173--22272
```

```
loss: 0.23870646083340874
accuracy: 0.9140625
2--174--22400
loss: 0.24846726691167115
accuracy: 0.9296875
2--175--22528
loss: 0.24470385664476388
accuracy: 0.9296875
2--176--22656
loss: 0.1806124687829641
accuracy: 0.9375
2--177--22784
loss: 0.18273405034756168
accuracy: 0.953125
2--178--22912
loss: 0.2027859933192368
accuracy: 0.9296875
2--179--23040
loss: 0.24986823784187218
accuracy: 0.9296875
2--180--23168
loss: 0.19533859854996083
accuracy: 0.9453125
2--181--23296
loss: 0.3677866376653006
accuracy: 0.9296875
2--182--23424
loss: 0.24554026759027503
accuracy: 0.9296875
2--183--23552
loss: 0.21762196959795094
accuracy: 0.921875
2--184--23680
loss: 0.2662105171886
accuracy: 0.921875
2--185--23808
loss: 0.1660840008670295
accuracy: 0.953125
2--186--23936
```

```
loss: 0.31122251628396386
accuracy: 0.8984375
2--187--24064
loss: 0.4162817839341352
accuracy: 0.8828125
2--188--24192
loss: 0.2463413744563606
accuracy: 0.9140625
2--189--24320
loss: 0.3367019328005427
accuracy: 0.9296875
2--190--24448
loss: 0.26800486477491714
accuracy: 0.9140625
2--191--24576
loss: 0.33900175576320246
accuracy: 0.90625
2--192--24704
loss: 0.21309262477867014
accuracy: 0.9296875
2--193--24832
loss: 0.26249411182650456
accuracy: 0.890625
2--194--24960
loss: 0.2911193561797826
accuracy: 0.921875
2--195--25088
loss: 0.18439977161928411
accuracy: 0.953125
2--196--25216
loss: 0.330365452474791
accuracy: 0.90625
2--197--25344
loss: 0.334585838882031
accuracy: 0.8984375
2--198--25472
loss: 0.4080992206256688
accuracy: 0.90625
2--199--25600
```

```
loss: 0.2539642070465923
accuracy: 0.9296875
2--200--25728
loss: 0.23080439146098
accuracy: 0.9140625
2--201--25856
loss: 0.29950389239497976
accuracy: 0.890625
2--202--25984
loss: 0.1683527605199734
accuracy: 0.9609375
2--203--26112
loss: 0.25820327049490455
accuracy: 0.9296875
2--204--26240
loss: 0.18991310727076
accuracy: 0.9296875
2--205--26368
loss: 0.36439684022918317
accuracy: 0.8984375
2--206--26496
loss: 0.2314942880110429
accuracy: 0.9453125
2--207--26624
loss: 0.1760921433105188
accuracy: 0.9453125
2--208--26752
loss: 0.31712116087917164
accuracy: 0.8984375
2--209--26880
loss: 0.2595149905503942
accuracy: 0.921875
2--210--27008
loss: 0.24620583498041496
accuracy: 0.921875
2--211--27136
loss: 0.20015425069101084
accuracy: 0.921875
2--212--27264
```

```
loss: 0.2982640786128716
accuracy: 0.9296875
2--213--27392
loss: 0.29850288034480765
accuracy: 0.90625
2--214--27520
loss: 0.1436075425045376
accuracy: 0.9609375
2--215--27648
loss: 0.3121079381468088
accuracy: 0.9140625
2--216--27776
loss: 0.2439864043706989
accuracy: 0.9453125
2--217--27904
loss: 0.23118648951350987
accuracy: 0.953125
2--218--28032
loss: 0.15068444648473175
accuracy: 0.953125
2--219--28160
loss: 0.22199905878627463
accuracy: 0.921875
2--220--28288
loss: 0.36430692142134097
accuracy: 0.9140625
2--221--28416
loss: 0.2617141676801218
accuracy: 0.9375
2--222--28544
loss: 0.2536757968269545
accuracy: 0.9375
2--223--28672
loss: 0.14866535987307758
accuracy: 0.9609375
2--224--28800
loss: 0.3690922334210863
accuracy: 0.8828125
2--225--28928
```

```
loss: 0.24808380212528044
accuracy: 0.921875
2--226--29056
loss: 0.24341947850466178
accuracy: 0.921875
2--227--29184
loss: 0.3593018392689412
accuracy: 0.890625
2--228--29312
loss: 0.2504936925562804
accuracy: 0.9453125
2--229--29399
loss: 0.22063736761852928
accuracy: 0.9425287356321839
3--0--128
loss: 0.3745320463317935
accuracy: 0.8828125
3--1--256
loss: 0.1765381238641272
accuracy: 0.921875
3--2--384
loss: 0.40850684569412765
accuracy: 0.8828125
3--3--512
loss: 0.20920549231952767
accuracy: 0.9296875
3--4--640
loss: 0.3209794164394881
accuracy: 0.8984375
3--5--768
loss: 0.2534020132463725
accuracy: 0.890625
3--6--896
loss: 0.28704157529416
accuracy: 0.9375
3--7--1024
loss: 0.22639839524472716
accuracy: 0.9375
3--8--1152
```

```
loss: 0.16613340746843647
accuracy: 0.953125
3--9--1280
loss: 0.27589608871360144
accuracy: 0.9140625
3--10--1408
loss: 0.18925423851357123
accuracy: 0.953125
3--11--1536
loss: 0.26580849483043123
accuracy: 0.921875
3--12--1664
loss: 0.19104179546796288
accuracy: 0.9296875
3--13--1792
loss: 0.17638583339019945
accuracy: 0.9375
3--14--1920
loss: 0.21448026487541083
accuracy: 0.9375
3--15--2048
loss: 0.22016247088891672
accuracy: 0.921875
3--16--2176
loss: 0.2527047391065319
accuracy: 0.9375
3--17--2304
loss: 0.17981911637562045
accuracy: 0.96875
3--18--2432
loss: 0.2669384703608221
accuracy: 0.9375
3--19--2560
loss: 0.28442579007151547
accuracy: 0.8828125
3--20--2688
loss: 0.17813960353638758
accuracy: 0.9140625
3--21--2816
```

```
loss: 0.21865110870211152
accuracy: 0.921875
3--22--2944
loss: 0.27509824083084006
accuracy: 0.921875
3--23--3072
loss: 0.23921142072095103
accuracy: 0.8984375
3--24--3200
loss: 0.21767698820698217
accuracy: 0.9296875
3--25--3328
loss: 0.2028690185466783
accuracy: 0.921875
3--26--3456
loss: 0.21659859824812155
accuracy: 0.9296875
3--27--3584
loss: 0.14589050644142126
accuracy: 0.9609375
3--28--3712
loss: 0.15476434958466345
accuracy: 0.9765625
3--29--3840
loss: 0.20963772021794147
accuracy: 0.921875
3--30--3968
loss: 0.19256686096944586
accuracy: 0.9375
3--31--4096
loss: 0.2935553955288488
accuracy: 0.9140625
3--32--4224
loss: 0.31115805380222594
accuracy: 0.890625
3--33--4352
loss: 0.22228210539054372
accuracy: 0.9453125
3--34--4480
```

```
loss: 0.27013937008115496
accuracy: 0.9296875
3--35--4608
loss: 0.2912872807869159
accuracy: 0.9140625
3--36--4736
loss: 0.26385767368979207
accuracy: 0.90625
3--37--4864
loss: 0.17417162311512116
accuracy: 0.9375
3--38--4992
loss: 0.2830441232612647
accuracy: 0.9296875
3--39--5120
loss: 0.2646953684743941
accuracy: 0.9296875
3--40--5248
loss: 0.3062922665270055
accuracy: 0.921875
3--41--5376
loss: 0.16164009720665318
accuracy: 0.953125
3--42--5504
loss: 0.2028410857948995
accuracy: 0.9453125
3--43--5632
loss: 0.16744005831067754
accuracy: 0.9609375
3--44--5760
loss: 0.23576960184583695
accuracy: 0.9453125
3--45--5888
loss: 0.46464843109209664
accuracy: 0.875
3--46--6016
loss: 0.29749890240877874
accuracy: 0.921875
3--47--6144
```

```
loss: 0.19931745766829084
accuracy: 0.953125
3--48--6272
loss: 0.207656166694414
accuracy: 0.9296875
3--49--6400
loss: 0.22317985066242266
accuracy: 0.9296875
3--50--6528
loss: 0.3678140432302166
accuracy: 0.8984375
3--51--6656
loss: 0.12557251065903696
accuracy: 0.953125
3--52--6784
loss: 0.26469874588097364
accuracy: 0.9453125
3--53--6912
loss: 0.15396057337516597
accuracy: 0.96875
3--54--7040
loss: 0.1939151494333302
accuracy: 0.9296875
3--55--7168
loss: 0.3240276571713446
accuracy: 0.8828125
3--56--7296
loss: 0.18379087969607352
accuracy: 0.953125
3--57--7424
loss: 0.2700015345309253
accuracy: 0.9296875
3--58--7552
loss: 0.2995463845850529
accuracy: 0.9296875
3--59--7680
loss: 0.34960616187099947
accuracy: 0.8984375
3--60--7808
```

```
loss: 0.1747026191435186
accuracy: 0.9609375
3--61--7936
loss: 0.38299777724865003
accuracy: 0.8984375
3--62--8064
loss: 0.21824070092788705
accuracy: 0.9296875
3--63--8192
loss: 0.25641544606764977
accuracy: 0.9375
3--64--8320
loss: 0.2244944078712317
accuracy: 0.9140625
3--65--8448
loss: 0.15681276479948175
accuracy: 0.96875
3--66--8576
loss: 0.14061251204877187
accuracy: 0.9765625
3--67--8704
loss: 0.13624363454139027
accuracy: 0.984375
3--68--8832
loss: 0.25480349136825803
accuracy: 0.921875
3--69--8960
loss: 0.2124945519907393
accuracy: 0.9296875
3--70--9088
loss: 0.18418825324679364
accuracy: 0.9375
3--71--9216
loss: 0.24583423344589847
accuracy: 0.9140625
3--72--9344
loss: 0.34465764876600485
accuracy: 0.8671875
3--73--9472
```

```
loss: 0.24346184984273492
accuracy: 0.8984375
3--74--9600
loss: 0.339775300604188
accuracy: 0.921875
3--75--9728
loss: 0.24977602972556884
accuracy: 0.9140625
3--76--9856
loss: 0.2405953563977826
accuracy: 0.921875
3--77--9984
loss: 0.22003559270214168
accuracy: 0.953125
3--78--10112
loss: 0.24085194417613984
accuracy: 0.921875
3--79--10240
loss: 0.3346375985492658
accuracy: 0.921875
3--80--10368
loss: 0.30321625324114737
accuracy: 0.9296875
3--81--10496
loss: 0.17295977350295555
accuracy: 0.9296875
3--82--10624
loss: 0.2758450290857206
accuracy: 0.90625
3--83--10752
loss: 0.19321882182395939
accuracy: 0.9375
3--84--10880
loss: 0.18686850469871336
accuracy: 0.96875
3--85--11008
loss: 0.13565538699071408
accuracy: 0.984375
3--86--11136
```

```
loss: 0.1873672214306457
accuracy: 0.9375
3--87--11264
loss: 0.30052036853406994
accuracy: 0.8828125
3--88--11392
loss: 0.32509252712755105
accuracy: 0.90625
3--89--11520
loss: 0.14436656144871424
accuracy: 0.9765625
3--90--11648
loss: 0.19741723433904124
accuracy: 0.9609375
3--91--11776
loss: 0.199206717368618
accuracy: 0.9375
3--92--11904
loss: 0.33926617092585776
accuracy: 0.890625
3--93--12032
loss: 0.2729901199704835
accuracy: 0.90625
3--94--12160
loss: 0.19162003928997215
accuracy: 0.9296875
3--95--12288
loss: 0.21215231893450287
accuracy: 0.953125
3--96--12416
loss: 0.25332895662748023
accuracy: 0.953125
3--97--12544
loss: 0.3003618677233149
accuracy: 0.8984375
3--98--12672
loss: 0.15554875667870727
accuracy: 0.96875
3--99--12800
```

```
loss: 0.3243784842657135
accuracy: 0.9375
3--100--12928
loss: 0.20262561838601315
accuracy: 0.9375
3--101--13056
loss: 0.21068553752962627
accuracy: 0.9453125
3--102--13184
loss: 0.13012513994354524
accuracy: 0.96875
3--103--13312
loss: 0.2770658745666306
accuracy: 0.9140625
3--104--13440
loss: 0.28400294694329287
accuracy: 0.90625
3--105--13568
loss: 0.31554719153324196
accuracy: 0.9140625
3--106--13696
loss: 0.24877068126284868
accuracy: 0.9140625
3--107--13824
loss: 0.2154044747494045
accuracy: 0.9296875
3--108--13952
loss: 0.131144160558513
accuracy: 0.9453125
3--109--14080
loss: 0.2373595195423747
accuracy: 0.9296875
3--110--14208
loss: 0.2922304045253362
accuracy: 0.90625
3--111--14336
loss: 0.30749746393031957
accuracy: 0.921875
3--112--14464
```

```
loss: 0.20976433386778454
accuracy: 0.921875
3--113--14592
loss: 0.24475433993601697
accuracy: 0.9375
3--114--14720
loss: 0.222466999229747
accuracy: 0.9140625
3--115--14848
loss: 0.23105961423305357
accuracy: 0.921875
3--116--14976
loss: 0.3116676254120633
accuracy: 0.921875
3--117--15104
loss: 0.26789192257455996
accuracy: 0.9296875
3--118--15232
loss: 0.25444517732221655
accuracy: 0.9296875
3--119--15360
loss: 0.13993826454596633
accuracy: 0.9609375
3--120--15488
loss: 0.20500716513045578
accuracy: 0.9453125
3--121--15616
loss: 0.2522686076300354
accuracy: 0.953125
3--122--15744
loss: 0.3120215876278157
accuracy: 0.890625
3--123--15872
loss: 0.22439185976795567
accuracy: 0.953125
3--124--16000
loss: 0.1799474290230641
accuracy: 0.953125
3--125--16128
```

```
loss: 0.14147879798266644
accuracy: 0.9609375
3--126--16256
loss: 0.2135019355738374
accuracy: 0.9453125
3--127--16384
loss: 0.15531874735301482
accuracy: 0.9453125
3--128--16512
loss: 0.23679262203443788
accuracy: 0.9375
3--129--16640
loss: 0.22197865983411122
accuracy: 0.9375
3--130--16768
loss: 0.21940627808132357
accuracy: 0.9296875
3--131--16896
loss: 0.28005867536360074
accuracy: 0.90625
3--132--17024
loss: 0.3578435426570293
accuracy: 0.8828125
3--133--17152
loss: 0.11338385456927327
accuracy: 0.9765625
3--134--17280
loss: 0.3415443389693035
accuracy: 0.9140625
3--135--17408
loss: 0.2887996571925874
accuracy: 0.9375
3--136--17536
loss: 0.2285472093088697
accuracy: 0.9453125
3--137--17664
loss: 0.20921180277170026
accuracy: 0.9296875
3--138--17792
```

```
loss: 0.15668103143074524
accuracy: 0.9609375
3--139--17920
loss: 0.1838888004431584
accuracy: 0.9375
3--140--18048
loss: 0.18002870963326909
accuracy: 0.9453125
3--141--18176
loss: 0.13937151315372437
accuracy: 0.96875
3--142--18304
loss: 0.2634648618923728
accuracy: 0.9296875
3--143--18432
loss: 0.28571718449796507
accuracy: 0.921875
3--144--18560
loss: 0.20686194032230174
accuracy: 0.9453125
3--145--18688
loss: 0.19763393788576372
accuracy: 0.953125
3--146--18816
loss: 0.2193465554697649
accuracy: 0.9296875
3--147--18944
loss: 0.26585812557861527
accuracy: 0.9375
3--148--19072
loss: 0.30112142216750043
accuracy: 0.9140625
3--149--19200
loss: 0.3243199087337086
accuracy: 0.9140625
3--150--19328
loss: 0.16382241998975583
accuracy: 0.9375
3--151--19456
```

```
loss: 0.18563720728314484
accuracy: 0.953125
3--152--19584
loss: 0.20656745037177093
accuracy: 0.9296875
3--153--19712
loss: 0.23810761953499693
accuracy: 0.9296875
3--154--19840
loss: 0.1548725005860651
accuracy: 0.9609375
3--155--19968
loss: 0.298470001184726
accuracy: 0.9140625
3--156--20096
loss: 0.2306427706670859
accuracy: 0.9296875
3--157--20224
loss: 0.3233230351495567
accuracy: 0.921875
3--158--20352
loss: 0.17361598648958626
accuracy: 0.953125
3--159--20480
loss: 0.38208914295336505
accuracy: 0.8984375
3--160--20608
loss: 0.4104934854413442
accuracy: 0.8984375
3--161--20736
loss: 0.25651673267690256
accuracy: 0.921875
3--162--20864
loss: 0.28110807373276336
accuracy: 0.9296875
3--163--20992
loss: 0.2199780238960099
accuracy: 0.9453125
3--164--21120
```

```
loss: 0.16732272763335393
accuracy: 0.9609375
3--165--21248
loss: 0.12567062033902904
accuracy: 0.953125
3--166--21376
loss: 0.281199237411468
accuracy: 0.9296875
3--167--21504
loss: 0.315100171083197
accuracy: 0.9296875
3--168--21632
loss: 0.26477824986694243
accuracy: 0.9375
3--169--21760
loss: 0.17936424944034757
accuracy: 0.9296875
3--170--21888
loss: 0.23142626705254665
accuracy: 0.953125
3--171--22016
loss: 0.20269827750146835
accuracy: 0.9453125
3--172--22144
loss: 0.2588965698420835
accuracy: 0.9296875
3--173--22272
loss: 0.21502885393132234
accuracy: 0.9375
3--174--22400
loss: 0.22419890984898766
accuracy: 0.9296875
3--175--22528
loss: 0.22314713745701986
accuracy: 0.9296875
3--176--22656
loss: 0.15580721978836176
accuracy: 0.9453125
3--177--22784
```

```
loss: 0.162110652372484
accuracy: 0.953125
3--178--22912
loss: 0.18279046493003864
accuracy: 0.9296875
3--179--23040
loss: 0.22865221778473802
accuracy: 0.9296875
3--180--23168
loss: 0.17347179831609896
accuracy: 0.9453125
3--181--23296
loss: 0.33143032935327177
accuracy: 0.9375
3--182--23424
loss: 0.2190145232734133
accuracy: 0.9375
3--183--23552
loss: 0.18728222620018842
accuracy: 0.9453125
3--184--23680
loss: 0.22847741936658794
accuracy: 0.9296875
3--185--23808
loss: 0.14818933356004754
accuracy: 0.953125
3--186--23936
loss: 0.28452760562922147
accuracy: 0.9140625
3--187--24064
loss: 0.38724581696704
accuracy: 0.90625
3--188--24192
loss: 0.23412607444498348
accuracy: 0.9140625
3--189--24320
loss: 0.31324863102184547
accuracy: 0.921875
3--190--24448
```

```
loss: 0.24131707773509908
accuracy: 0.9140625
3--191--24576
loss: 0.3110173913143684
accuracy: 0.9375
3--192--24704
loss: 0.18933270700318267
accuracy: 0.9296875
3--193--24832
loss: 0.23855157983007383
accuracy: 0.90625
3--194--24960
loss: 0.26874594511614863
accuracy: 0.921875
3--195--25088
loss: 0.15928946126189414
accuracy: 0.953125
3--196--25216
loss: 0.31588452324415006
accuracy: 0.9140625
3--197--25344
loss: 0.31726920248946183
accuracy: 0.890625
3--198--25472
loss: 0.3837290845948226
accuracy: 0.90625
3--199--25600
loss: 0.22038921926564425
accuracy: 0.9296875
3--200--25728
loss: 0.2062671181900289
accuracy: 0.9296875
3--201--25856
loss: 0.2740134276749661
accuracy: 0.890625
3--202--25984
loss: 0.15466762721039468
accuracy: 0.953125
3--203--26112
```

```
loss: 0.2358239745467462
accuracy: 0.9453125
3--204--26240
loss: 0.17278130579000855
accuracy: 0.9296875
3--205--26368
loss: 0.33698415925294767
accuracy: 0.8984375
3--206--26496
loss: 0.2169866154607956
accuracy: 0.9453125
3--207--26624
loss: 0.15128205002457126
accuracy: 0.953125
3--208--26752
loss: 0.29209547977912753
accuracy: 0.90625
3--209--26880
loss: 0.2283484310448265
accuracy: 0.921875
3--210--27008
loss: 0.22964551457613414
accuracy: 0.9296875
3--211--27136
loss: 0.17002546106690036
accuracy: 0.9375
3--212--27264
loss: 0.2741262246464584
accuracy: 0.9375
3--213--27392
loss: 0.2656441750577784
accuracy: 0.921875
3--214--27520
loss: 0.1285712413813624
accuracy: 0.9609375
3--215--27648
loss: 0.2859290627092792
accuracy: 0.921875
3--216--27776
```

```
loss: 0.22680448194467956
accuracy: 0.9453125
3--217--27904
loss: 0.2117715718789138
accuracy: 0.953125
3--218--28032
loss: 0.1314429630482407
accuracy: 0.9609375
3--219--28160
loss: 0.20354789307960947
accuracy: 0.9296875
3--220--28288
loss: 0.34624666484090116
accuracy: 0.9140625
3--221--28416
loss: 0.2404357052827938
accuracy: 0.9453125
3--222--28544
loss: 0.23452090810490678
accuracy: 0.9375
3--223--28672
loss: 0.13009297025790123
accuracy: 0.96875
3--224--28800
loss: 0.3404381906934405
accuracy: 0.8828125
3--225--28928
loss: 0.21532846926624505
accuracy: 0.9296875
3--226--29056
loss: 0.21589707960095458
accuracy: 0.9296875
3--227--29184
loss: 0.33927296560833353
accuracy: 0.90625
3--228--29312
loss: 0.22650685669148865
accuracy: 0.9609375
3--229--29399
```

```
loss: 0.19164046936831947
accuracy: 0.9540229885057471
4--0--128
loss: 0.3473088051415014
accuracy: 0.90625
4--1--256
loss: 0.15428242324710867
accuracy: 0.9375
4--2--384
loss: 0.37500307217663864
accuracy: 0.90625
4--3--512
loss: 0.1815322082835928
accuracy: 0.953125
4--4--640
loss: 0.2972646514343184
accuracy: 0.9140625
4--5--768
loss: 0.2269057940372407
accuracy: 0.90625
4--6--896
loss: 0.27017837621791474
accuracy: 0.9375
4--7--1024
loss: 0.20194329997101157
accuracy: 0.953125
4--8--1152
loss: 0.13871473440808663
accuracy: 0.96875
4--9--1280
loss: 0.25156409428341875
accuracy: 0.9375
4--10--1408
loss: 0.16917108262921082
accuracy: 0.9453125
4--11--1536
loss: 0.25106344389194196
accuracy: 0.921875
4--12--1664
```

```
loss: 0.16522813941787873
accuracy: 0.9375
4--13--1792
loss: 0.14974199139824937
accuracy: 0.9453125
4--14--1920
loss: 0.19369690187848487
accuracy: 0.9375
4--15--2048
loss: 0.2022024558474352
accuracy: 0.9375
4--16--2176
loss: 0.23041068018716826
accuracy: 0.9296875
4--17--2304
loss: 0.15476320409008287
accuracy: 0.96875
4--18--2432
loss: 0.24931002066930358
accuracy: 0.9375
4--19--2560
loss: 0.25491599937387277
accuracy: 0.8984375
4--20--2688
loss: 0.15476600160173345
accuracy: 0.9453125
4--21--2816
loss: 0.19643573223148397
accuracy: 0.9453125
4--22--2944
loss: 0.24517442339337248
accuracy: 0.921875
4--23--3072
loss: 0.21200875264688832
accuracy: 0.90625
4--24--3200
loss: 0.19505017818358517
accuracy: 0.9296875
4--25--3328
```

```
loss: 0.1826320411269322
accuracy: 0.9296875
4--26--3456
loss: 0.19641328210435197
accuracy: 0.9375
4--27--3584
loss: 0.12611286227673485
accuracy: 0.9765625
4--28--3712
loss: 0.13657313203172913
accuracy: 0.9765625
4--29--3840
loss: 0.18324998744253815
accuracy: 0.921875
4--30--3968
loss: 0.1664113700761874
accuracy: 0.953125
4--31--4096
loss: 0.26635131519878735
accuracy: 0.921875
4--32--4224
loss: 0.2853510030561251
accuracy: 0.90625
4--33--4352
loss: 0.20020656143892762
accuracy: 0.9453125
4--34--4480
loss: 0.24249131302136415
accuracy: 0.9375
4--35--4608
loss: 0.26238445519874126
accuracy: 0.921875
4--36--4736
loss: 0.21844266493189404
accuracy: 0.921875
4--37--4864
loss: 0.1561034364678347
accuracy: 0.9453125
4--38--4992
```

```
loss: 0.2677910279698471
accuracy: 0.9375
4--39--5120
loss: 0.2427759930482481
accuracy: 0.9375
4--40--5248
loss: 0.27953890523266095
accuracy: 0.921875
4--41--5376
loss: 0.1465806537936306
accuracy: 0.953125
4--42--5504
loss: 0.18123259151700336
accuracy: 0.953125
4--43--5632
loss: 0.1435829575107308
accuracy: 0.96875
4--44--5760
loss: 0.2125016978091369
accuracy: 0.9453125
4--45--5888
loss: 0.4481392480613742
accuracy: 0.8984375
4--46--6016
loss: 0.2702537618802683
accuracy: 0.921875
4--47--6144
loss: 0.18283551723862518
accuracy: 0.9453125
4--48--6272
loss: 0.17867901965222555
accuracy: 0.9375
4--49--6400
loss: 0.19651405392506305
accuracy: 0.9453125
4--50--6528
loss: 0.35061490405634477
accuracy: 0.9140625
4--51--6656
```

```
loss: 0.10809483273868442
accuracy: 0.9609375
4--52--6784
loss: 0.240513989807428
accuracy: 0.9453125
4--53--6912
loss: 0.13724038582854647
accuracy: 0.96875
4--54--7040
loss: 0.17101238811748826
accuracy: 0.9453125
4--55--7168
loss: 0.29441668060912407
accuracy: 0.90625
4--56--7296
loss: 0.16452451309253402
accuracy: 0.9609375
4--57--7424
loss: 0.24082398873956268
accuracy: 0.9296875
4--58--7552
loss: 0.2778889909123381
accuracy: 0.9375
4--59--7680
loss: 0.3298402637834492
accuracy: 0.8984375
4--60--7808
loss: 0.1553944237373288
accuracy: 0.9609375
4--61--7936
loss: 0.35485255000983607
accuracy: 0.90625
4--62--8064
loss: 0.1948582347072141
accuracy: 0.9296875
4--63--8192
loss: 0.23878496423497722
accuracy: 0.9375
4--64--8320
```

```
loss: 0.2045332184594949
accuracy: 0.9296875
4--65--8448
loss: 0.14264787525536632
accuracy: 0.96875
4--66--8576
loss: 0.11739757367912226
accuracy: 0.984375
4--67--8704
loss: 0.1172478313721293
accuracy: 0.984375
4--68--8832
loss: 0.22946253865868732
accuracy: 0.9296875
4--69--8960
loss: 0.18877786673706287
accuracy: 0.9296875
4--70--9088
loss: 0.16647309722415377
accuracy: 0.9375
4--71--9216
loss: 0.21686925469957624
accuracy: 0.921875
4--72--9344
loss: 0.3004506879534851
accuracy: 0.8984375
4--73--9472
loss: 0.21345709275904298
accuracy: 0.9140625
4--74--9600
loss: 0.299903373615383
accuracy: 0.9375
4--75--9728
loss: 0.2280047940999874
accuracy: 0.921875
4--76--9856
loss: 0.21340339241198036
accuracy: 0.9296875
4--77--9984
```

```
loss: 0.1946250906128487
accuracy: 0.953125
4--78--10112
loss: 0.21743654820372293
accuracy: 0.9375
4--79--10240
loss: 0.30261909374368307
accuracy: 0.9296875
4--80--10368
loss: 0.27454156512842964
accuracy: 0.9375
4--81--10496
loss: 0.1464114812107803
accuracy: 0.9296875
4--82--10624
loss: 0.25514153396439165
accuracy: 0.921875
4--83--10752
loss: 0.1709068042614178
accuracy: 0.9453125
4--84--10880
loss: 0.16743564711697892
accuracy: 0.96875
4--85--11008
loss: 0.11638199466205286
accuracy: 0.984375
4--86--11136
loss: 0.16856729568199957
accuracy: 0.9375
4--87--11264
loss: 0.2794145724214695
accuracy: 0.8828125
4--88--11392
loss: 0.3054642257222573
accuracy: 0.9140625
4--89--11520
loss: 0.12453490668140726
accuracy: 0.9765625
4--90--11648
```

```
loss: 0.17515790705420367
accuracy: 0.9609375
4--91--11776
loss: 0.17861904539316117
accuracy: 0.9375
4--92--11904
loss: 0.31995014502137664
accuracy: 0.8984375
4--93--12032
loss: 0.2436488524129315
accuracy: 0.90625
4--94--12160
loss: 0.1703556926923637
accuracy: 0.9296875
4--95--12288
loss: 0.18831394732996187
accuracy: 0.9609375
4--96--12416
loss: 0.22949011741845593
accuracy: 0.953125
4--97--12544
loss: 0.27616845829241937
accuracy: 0.921875
4--98--12672
loss: 0.14695358523738591
accuracy: 0.9765625
4--99--12800
loss: 0.30826419021214185
accuracy: 0.9375
4--100--12928
loss: 0.18105837965519433
accuracy: 0.9375
4--101--13056
loss: 0.18731779436385565
accuracy: 0.953125
4--102--13184
loss: 0.11130700355349205
accuracy: 0.984375
4--103--13312
```

```
loss: 0.24891680748961395
accuracy: 0.921875
4--104--13440
loss: 0.25547673533514453
accuracy: 0.90625
4--105--13568
loss: 0.2818721579774377
accuracy: 0.921875
4--106--13696
loss: 0.22940536771990566
accuracy: 0.9140625
4--107--13824
loss: 0.19020368644408786
accuracy: 0.9375
4--108--13952
loss: 0.11716078346317024
accuracy: 0.9453125
4--109--14080
loss: 0.21700612276952325
accuracy: 0.9375
4--110--14208
loss: 0.26972577169365375
accuracy: 0.921875
4--111--14336
loss: 0.2887339675476138
accuracy: 0.921875
4--112--14464
loss: 0.18881120132325035
accuracy: 0.9375
4--113--14592
loss: 0.22230948023186783
accuracy: 0.9375
4--114--14720
loss: 0.2090757623249052
accuracy: 0.921875
4--115--14848
loss: 0.2154870685738199
accuracy: 0.9375
4--116--14976
```

```
loss: 0.29120060812002957
accuracy: 0.921875
4--117--15104
loss: 0.2516875210804145
accuracy: 0.9375
4--118--15232
loss: 0.22774257295586295
accuracy: 0.9375
4--119--15360
loss: 0.12830747834828649
accuracy: 0.9609375
4--120--15488
loss: 0.18224703284315363
accuracy: 0.9453125
4--121--15616
loss: 0.23220465376461624
accuracy: 0.9609375
4--122--15744
loss: 0.2878933457801678
accuracy: 0.8984375
4--123--15872
loss: 0.20799687716837523
accuracy: 0.953125
4--124--16000
loss: 0.16005518551703846
accuracy: 0.96875
4--125--16128
loss: 0.12793902310279098
accuracy: 0.9609375
4--126--16256
loss: 0.1943362312528348
accuracy: 0.9453125
4--127--16384
loss: 0.13553740913317208
accuracy: 0.9453125
4--128--16512
loss: 0.21666334408277954
accuracy: 0.9453125
4--129--16640
```

```
loss: 0.19770677363120415
accuracy: 0.9453125
4--130--16768
loss: 0.19717726710284267
accuracy: 0.9609375
4--131--16896
loss: 0.24574046325169377
accuracy: 0.90625
4--132--17024
loss: 0.33301069539701794
accuracy: 0.890625
4--133--17152
loss: 0.09897151645755026
accuracy: 0.9765625
4--134--17280
loss: 0.31608486211452164
accuracy: 0.9140625
4--135--17408
loss: 0.27509729254516707
accuracy: 0.9375
4--136--17536
loss: 0.2021883310457841
accuracy: 0.953125
4--137--17664
loss: 0.19067923928799974
accuracy: 0.9296875
4--138--17792
loss: 0.13823823582555483
accuracy: 0.96875
4--139--17920
loss: 0.16702620005478624
accuracy: 0.9375
4--140--18048
loss: 0.16477472685298683
accuracy: 0.953125
4--141--18176
loss: 0.12279997343387833
accuracy: 0.96875
4--142--18304
```

```
loss: 0.24827728173044322
accuracy: 0.9453125
4--143--18432
loss: 0.26831408180612426
accuracy: 0.921875
4--144--18560
loss: 0.18633816460813352
accuracy: 0.9453125
4--145--18688
loss: 0.17259504479409854
accuracy: 0.96875
4--146--18816
loss: 0.20936105393304863
accuracy: 0.921875
4--147--18944
loss: 0.24438779063885963
accuracy: 0.9296875
4--148--19072
loss: 0.2821565693861163
accuracy: 0.9140625
4--149--19200
loss: 0.30178085789229325
accuracy: 0.9140625
4--150--19328
loss: 0.15150407132577826
accuracy: 0.9453125
4--151--19456
loss: 0.16357548720795784
accuracy: 0.96875
4--152--19584
loss: 0.18377383496196012
accuracy: 0.9453125
4--153--19712
loss: 0.21960075676874746
accuracy: 0.9375
4--154--19840
loss: 0.1391354621446731
accuracy: 0.9609375
4--155--19968
```

```
loss: 0.26802971224837496
accuracy: 0.9140625
4--156--20096
loss: 0.21165631545517294
accuracy: 0.9296875
4--157--20224
loss: 0.2920929382788837
accuracy: 0.9296875
4--158--20352
loss: 0.15893823133921198
accuracy: 0.9609375
4--159--20480
loss: 0.3540270337134941
accuracy: 0.90625
4--160--20608
loss: 0.3765765420653161
accuracy: 0.8984375
4--161--20736
loss: 0.2362343120438447
accuracy: 0.9296875
4--162--20864
loss: 0.2632897469588855
accuracy: 0.9296875
4--163--20992
loss: 0.19239745556935745
accuracy: 0.9453125
4--164--21120
loss: 0.15658280690668558
accuracy: 0.953125
4--165--21248
loss: 0.11018677510618216
accuracy: 0.96875
4--166--21376
loss: 0.25330075341720126
accuracy: 0.9375
4--167--21504
loss: 0.29482142117224003
accuracy: 0.9296875
4--168--21632
```

```
loss: 0.2496404014411239
accuracy: 0.9375
4--169--21760
loss: 0.15654102018619617
accuracy: 0.9375
4--170--21888
loss: 0.21115331076879118
accuracy: 0.9609375
4--171--22016
loss: 0.186894443692652
accuracy: 0.9453125
4--172--22144
loss: 0.224789668422549
accuracy: 0.9609375
4--173--22272
loss: 0.1973216498892326
accuracy: 0.9453125
4--174--22400
loss: 0.2036453410062981
accuracy: 0.9375
4--175--22528
loss: 0.2050423967411078
accuracy: 0.9375
4--176--22656
loss: 0.13756878484257085
accuracy: 0.953125
4--177--22784
loss: 0.14588562538751845
accuracy: 0.953125
4--178--22912
loss: 0.16834311299517724
accuracy: 0.9296875
4--179--23040
loss: 0.20978379717515488
accuracy: 0.9375
4--180--23168
loss: 0.1581502732882275
accuracy: 0.9609375
4--181--23296
```

```
loss: 0.30142223475813135
accuracy: 0.9375
4--182--23424
loss: 0.20020181201133075
accuracy: 0.9453125
4--183--23552
loss: 0.166217960783842
accuracy: 0.9453125
4--184--23680
loss: 0.2004790303811757
accuracy: 0.9296875
4--185--23808
loss: 0.13488663351531902
accuracy: 0.9609375
4--186--23936
loss: 0.2642442750992303
accuracy: 0.9140625
4--187--24064
loss: 0.36344703042870746
accuracy: 0.90625
4--188--24192
loss: 0.22454976873230106
accuracy: 0.921875
4--189--24320
loss: 0.29377585742803103
accuracy: 0.9375
4--190--24448
loss: 0.21915623527578887
accuracy: 0.9453125
4--191--24576
loss: 0.2888051909100393
accuracy: 0.9375
4--192--24704
loss: 0.1700751531583608
accuracy: 0.9296875
4--193--24832
loss: 0.2198196906714845
accuracy: 0.9140625
4--194--24960
```

```
loss: 0.25078148984770054
accuracy: 0.9296875
4--195--25088
loss: 0.14153178801175986
accuracy: 0.9609375
4--196--25216
loss: 0.3023061913760594
accuracy: 0.921875
4--197--25344
loss: 0.3033680329707076
accuracy: 0.8984375
4--198--25472
loss: 0.36265597128156274
accuracy: 0.90625
4--199--25600
loss: 0.19358156279830613
accuracy: 0.9296875
4--200--25728
loss: 0.1876284098145668
accuracy: 0.9296875
4--201--25856
loss: 0.254522099241759
accuracy: 0.90625
4--202--25984
loss: 0.1448301971377703
accuracy: 0.9609375
4--203--26112
loss: 0.21639550644067523
accuracy: 0.953125
4--204--26240
loss: 0.16026514121428076
accuracy: 0.953125
4--205--26368
loss: 0.31387583246728834
accuracy: 0.90625
4--206--26496
loss: 0.20744355656488156
accuracy: 0.9453125
4--207--26624
```

```
loss: 0.13474342046189247
accuracy: 0.96875
4--208--26752
loss: 0.26879003751392144
accuracy: 0.9140625
4--209--26880
loss: 0.20254818356045182
accuracy: 0.9296875
4--210--27008
loss: 0.21398362656411749
accuracy: 0.9453125
4--211--27136
loss: 0.1491441271215847
accuracy: 0.9609375
4--212--27264
loss: 0.25480025944582774
accuracy: 0.9375
4--213--27392
loss: 0.23771575672541972
accuracy: 0.921875
4--214--27520
loss: 0.11636209717052189
accuracy: 0.9609375
4--215--27648
loss: 0.2646070282105851
accuracy: 0.921875
4--216--27776
loss: 0.21227390437637356
accuracy: 0.9453125
4--217--27904
loss: 0.19590521896421656
accuracy: 0.9609375
4--218--28032
loss: 0.11751416633383338
accuracy: 0.9609375
4--219--28160
loss: 0.18846191980064797
accuracy: 0.9296875
4--220--28288
```

```
loss: 0.33051941285664543
accuracy: 0.921875
4--221--28416
loss: 0.22392232273574922
accuracy: 0.9453125
4--222--28544
loss: 0.219323078742186
accuracy: 0.9296875
4--223--28672
loss: 0.1165998121947568
accuracy: 0.96875
4--224--28800
loss: 0.3165181200996261
accuracy: 0.8984375
4--225--28928
loss: 0.18950658569351692
accuracy: 0.9453125
4--226--29056
loss: 0.19469315971257084
accuracy: 0.9453125
4--227--29184
loss: 0.3202240651737227
accuracy: 0.921875
4--228--29312
loss: 0.20897876309324015
accuracy: 0.9609375
4--229--29399
loss: 0.1699488396563522
accuracy: 0.9655172413793104
```

```
In [14]: hat_label = forward(test_data.T, parameter, cache)
         hat_label.keys()

Out[14]: dict_keys(['C1', 'A1', 'C2', 'A2', 'C3', 'A3', 'dC3', 'dW3', 'db3', 'dC2', 'dW2', 'db2',

In [15]: hat_label = hat_label['A3']
         hat_label.shape

Out[15]: (10, 12601)

In [16]: loss(test_label.T, hat_label)
```

Out[16]: 0.2211812376862138

In [ ]: