

实验报告

By wind

1、概述：

本文用三种方式训练了手写数字识别的卷积网络 ConvNet, 并得到了不同的结果。

三种训练方式中, 每种方式训练周期为 100, 分为两个阶段, 各 50 周期

(1) 第一阶段(前 50 周期)用 0~4 训练网络, 第二阶段(后 50 周期)用 5~9 训练网络, 记为'0~4/5~9'

(2) 第一阶段用 0~4 训练网络, 第二阶段用 0~9 训练, 记为'0~4/0~9'

(3) 第一阶段、第二阶段都用 0~9 训练, 记为'0~9'

此外, 每种训练方式按训练数据集的大小分为[1/20, 1/10, 1/8, 1/6, 1/5, 1/4, 1/3, 1/2, 1]数据集, 并采取了训练 10 次取平均数的方法得到每种训练方式的结果

2、实验方法要点

2.1 分割训练集：

按照训练方式的不同, 本文训练集分为两部分, 用 train_A 表示训练集中 0~4 的部分, train_B 表示 5~9 的部分。全部训练集仍用 train_dataset 表示。 分割代码如下：

```
A = [0, 1, 2, 3, 4]
train_A = []
train_B = []
for i in range(len(train_dataset)):
    if train_dataset[i][1] in A:
        train_A.append(train_dataset[i])
    else:
        train_B.append(train_dataset[i])
```

运行 print(len(train_A), len(train_B)), 得到 30596, 29404. 即 train_A 和 train_B 中的数据量大致相等。

2.2 建立网络

本次实验所用卷积网络同 minst_convnet 中的 ConvNet, 从略

2.3 进行训练

在重循环下进行每次训练, 每次训练的两个阶段各有一个循环, 主要代码如下：

```
times = 10
fractions = [20, 10, 8, 6, 5, 4, 3, 2, 1]

for experiment in ['0~4/5~9', '0~4/0~9', '0~9']:
    for time in range(times):
        for fraction in fractions:
```

```

if experiment == '0~4/5~9':
    net = ConvNet()
    trainData = train_loaderA
    pass
if experiment == '0~4/0~9':
    net = ConvNet()
    trainData = train_loaderA
    pass
if experiment == '0~9':
    net = ConvNet()
    trainData = train_loader
    pass

for epoch in range(50):
    pass
for epoch in range(50, 100):
    pass

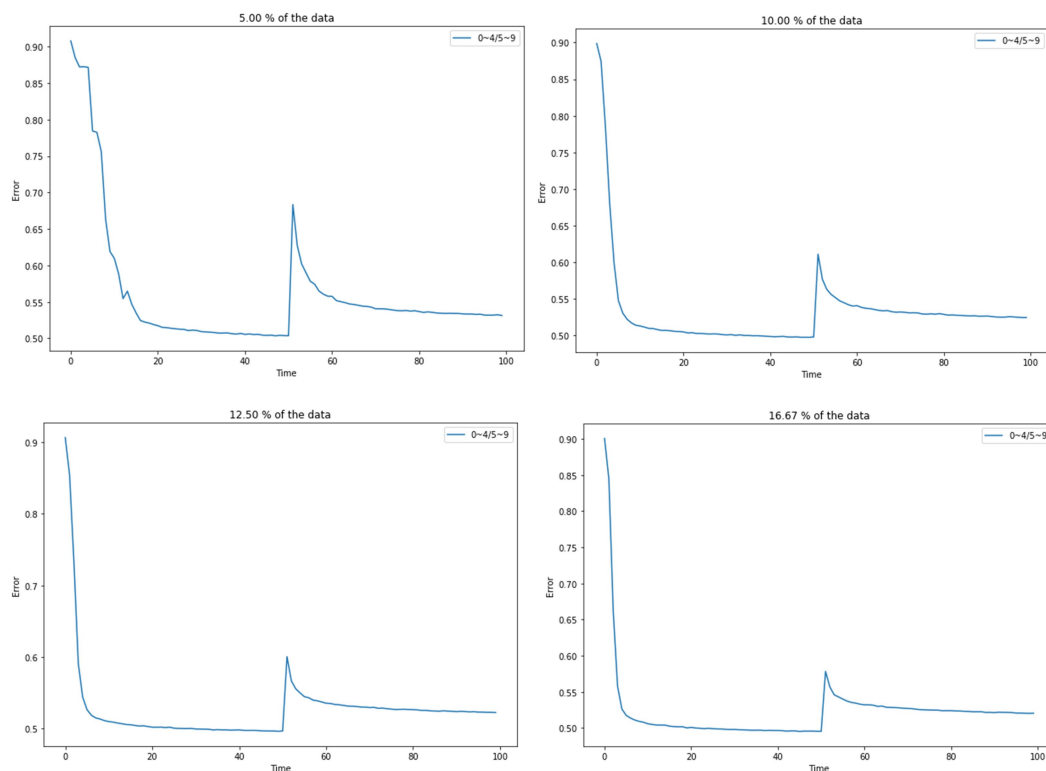
```

3、结果

本文根据实验结果的差异，先展示'0~4/5~9'的结果，再展示'0~4/0~9'和'0~9'的结果。

3.1 '0~4/5~9'的结果

(1) '0~4/5~9' 训练方式下，误差曲线如图：



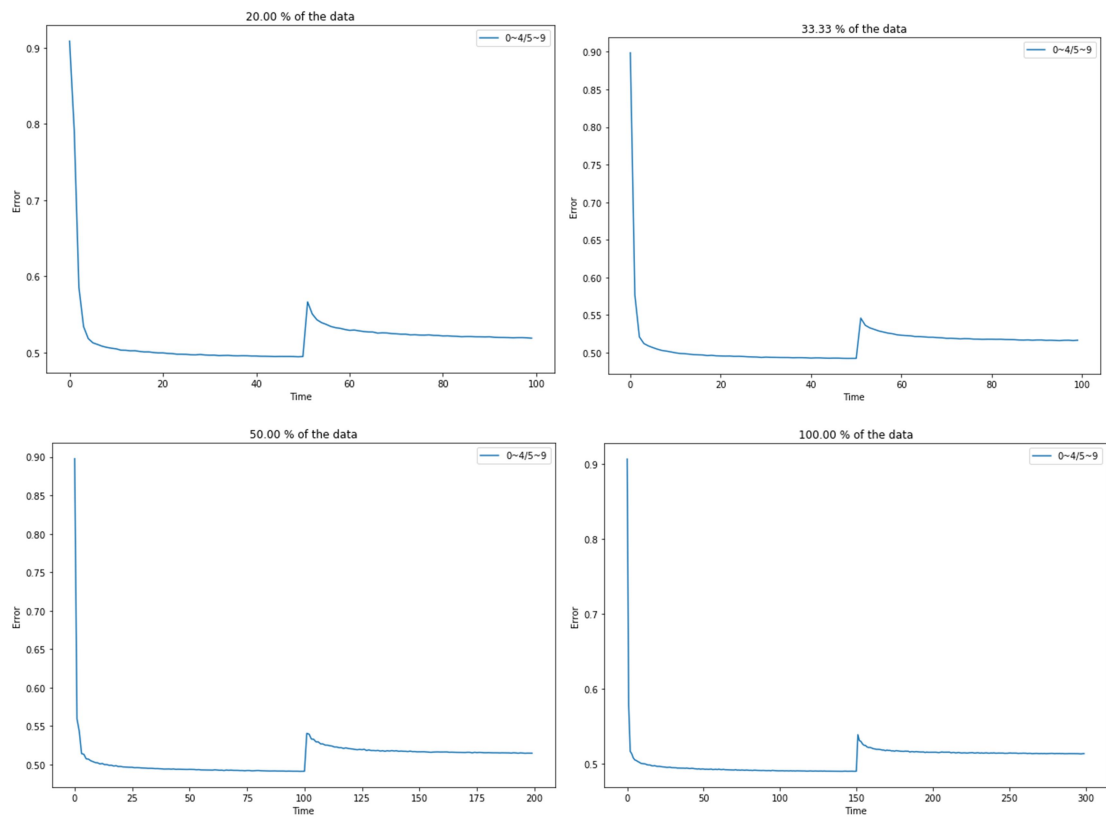


图1 ‘0~4/5~9’不同数据量下的误差曲线

可以看到采用 0~4 进行训练，第一阶段的错误率最多只能降到接近 0.5，在进入第二阶段训练（5~9）时，错误率有一个向上的突变，然后会降低，但最终仍在 0.5 以上，且略小于第一阶段的错误率。

（2）数据量的影响

由 1 中图可见，随着数据量逐渐增大，从第一阶段进入第二阶段的误差率突变量减小。

由数据量-误差率曲线可见，数据量增大时，误差率减小并趋近 0.5。

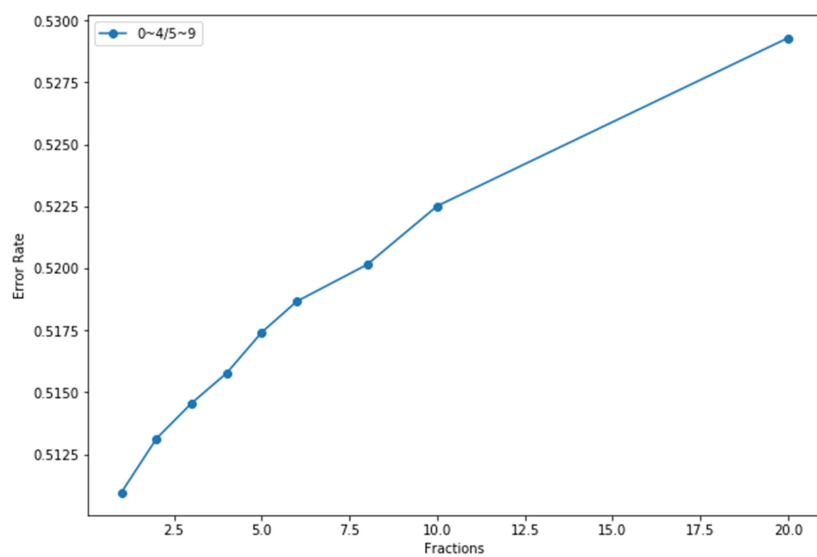


图2 ‘0~4/5~9’的数据量-误差曲线

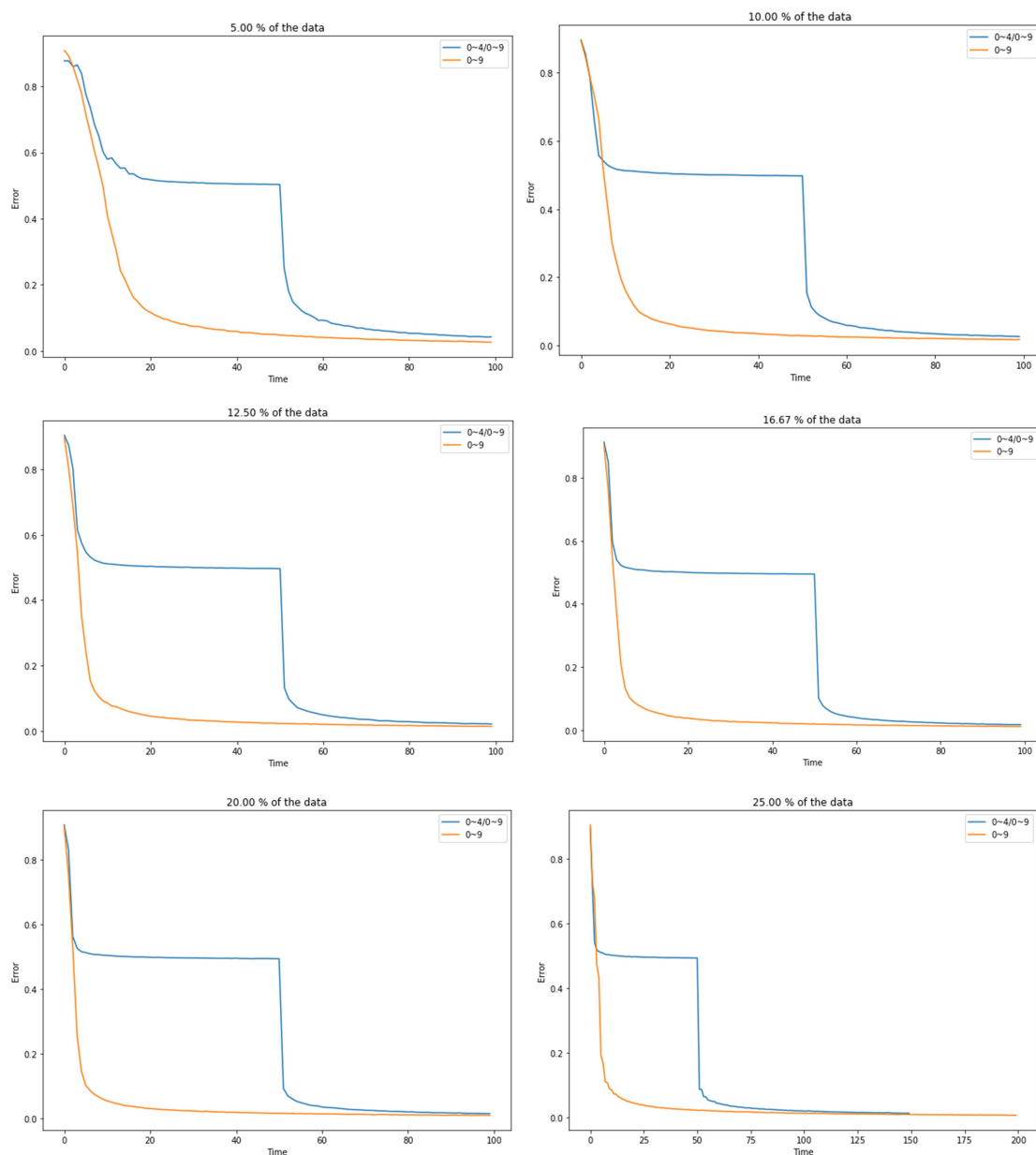
(3) 结果分析

从迁移学习的角度来看，网络从 $0 \sim 4$ 学到的特征对于 $5 \sim 9$ 的分类有一定的帮助，体现在面对全新的数据，第 51 周期的错误率与第 1 周期相比明显降低。数据量的增大使这种学习效果增强，但最多只能逼近网络对 $0 \sim 4$ 的学习效果而不会超过它。

3.2 ' $0 \sim 4/0 \sim 9$ ' 和 ' $0 \sim 9$ ' 的结果

(1) 这两种训练方式下，典型的误差曲线如图 3

在 ' $0 \sim 4/0 \sim 9$ ' 训练方式下，第一阶段网络的误差率降到 0.5 左右，第二阶段的误差率则进一步降到逼近 0。两个阶段的误差曲线形状相仿，在第一阶段、第二阶段交接处形成台阶型。' $0 \sim 9$ ' 训练方式则是一般的误差曲线形状。两种训练方式后期误差率都很小，但 ' $0 \sim 9$ ' 的误差曲线在 ' $0 \sim 4/0 \sim 9$ ' 下方。



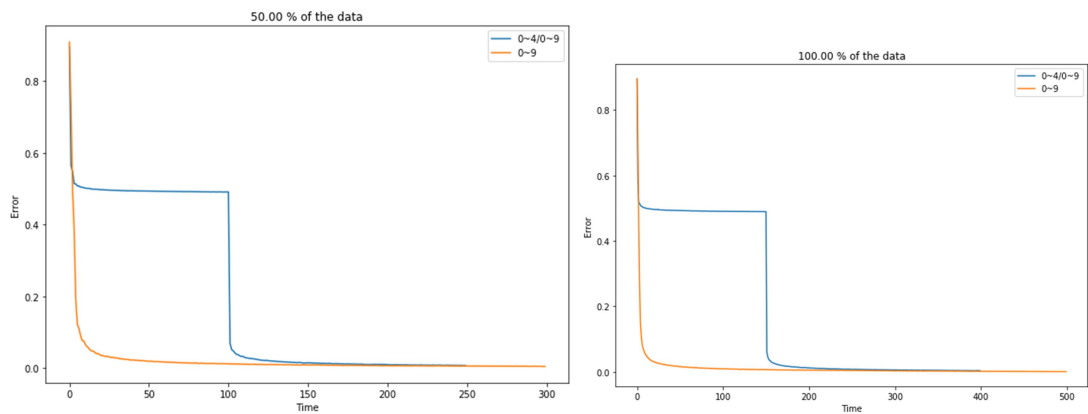


图3 ‘0~4/0~9’ 和 ‘0~9’ 不同数据量下的误差曲线

(2) 数据量的影响

两种训练方式下，误差率均在 0.05 以下，且与 (1) 中类似，随着数据量的增大，两种训练方式的误差率减小。使用 5% 的数据量时，误差率 0.03~0.04，使用 100% 的训练数据时，误差率在 0.01 以下。如图 4

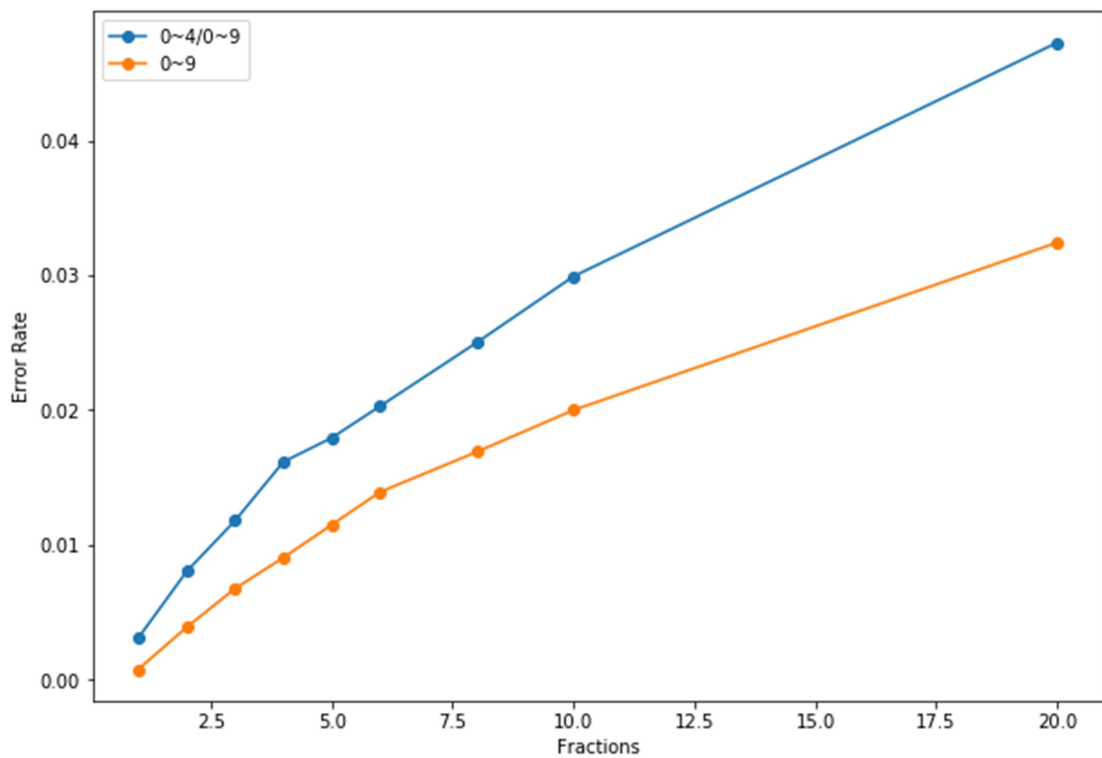


图4 ‘0~4/0~9’ 和 ‘0~9’ 不同数据量下的误差曲线

(3) 结果分析

从迁移学习的角度来看，从 0~4 学习的效果加速了学习 0~9 时的速率。体现在 ‘0~4/0~9’ 训练方式下 50 周期附近的误差曲线的斜率比 ‘0~9’ 训练方式的斜率（绝对值）更大。