

循环神经网络实验报告

姓名： 黄昶玮

学号： 2112380

实验要求：

- 掌握 RNN 原理
- 学会使用 PyTorch 搭建循环神经网络来训练名字识别
- 学会使用 PyTorch 搭建 LSTM 网络来训练名字识别

报告内容：

- 老师提供的原始版本 RNN 网络结构（可用 print(net)打印，复制文字或截图皆可）、在名字识别验证集上的训练 loss 曲线、准确度曲线图以及预测矩阵图
- 个人实现的 LSTM 网络结构在上述验证集上的训练 loss 曲线、准确度曲线图以及预测矩阵图
- 解释为什么 LSTM 网络的性能优于 RNN 网络（重点部分）

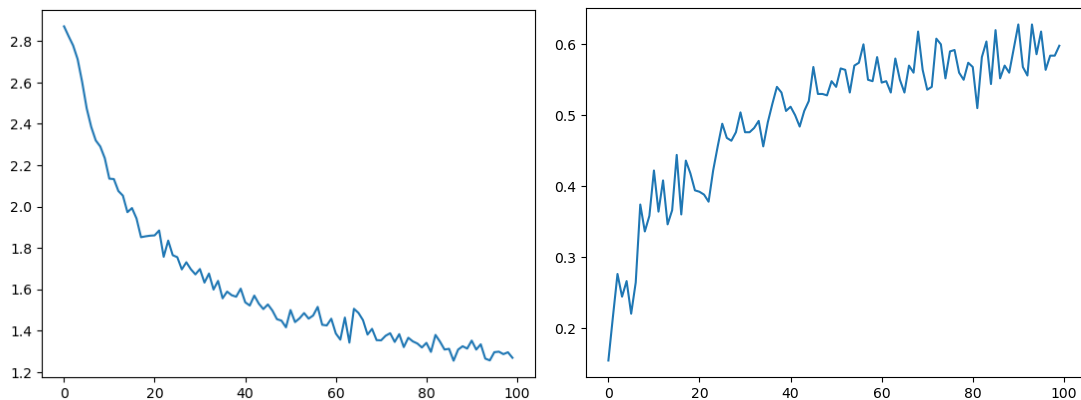
实验内容：

一 . RNN

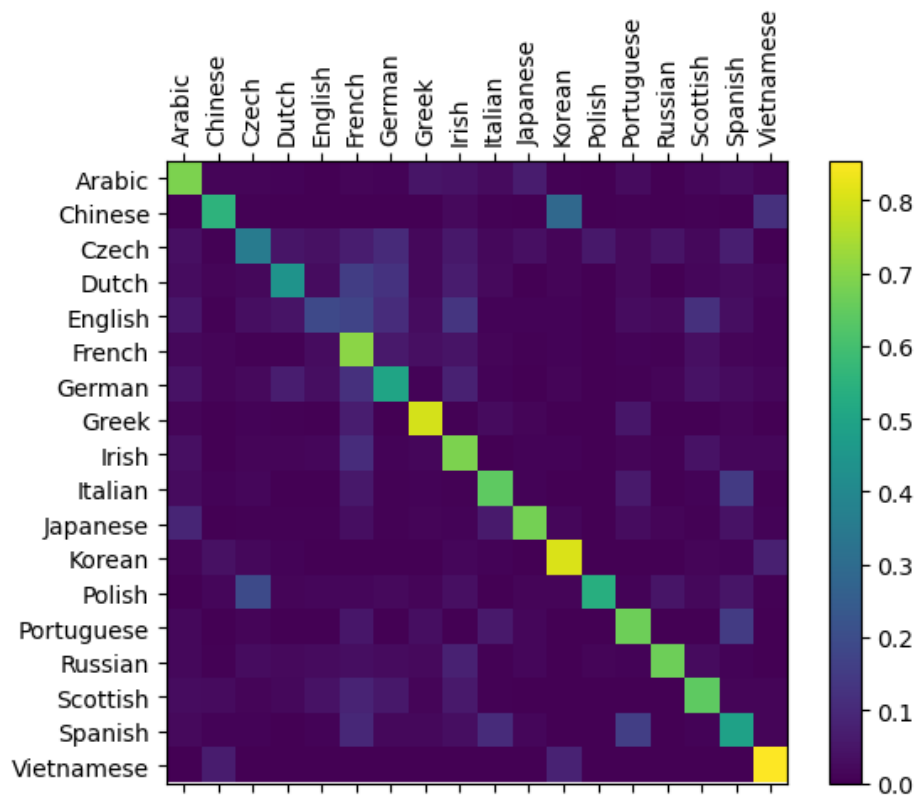
网络结构：

```
RNN(  
  (i2h): Linear(in_features=185, out_features=128, bias=True)  
  (i2o): Linear(in_features=185, out_features=18, bias=True)  
  (softmax): LogSoftmax(dim=1)  
)
```

输出结果：



预测矩阵图：

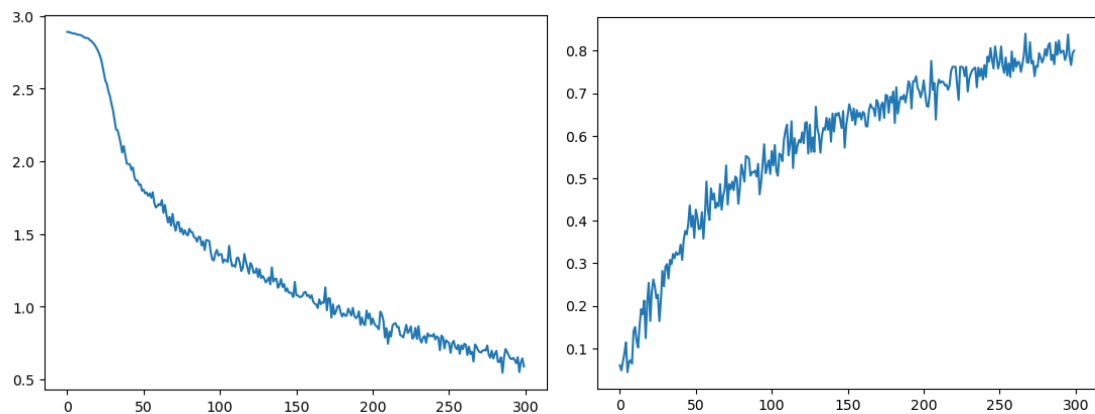


二 . LSTM

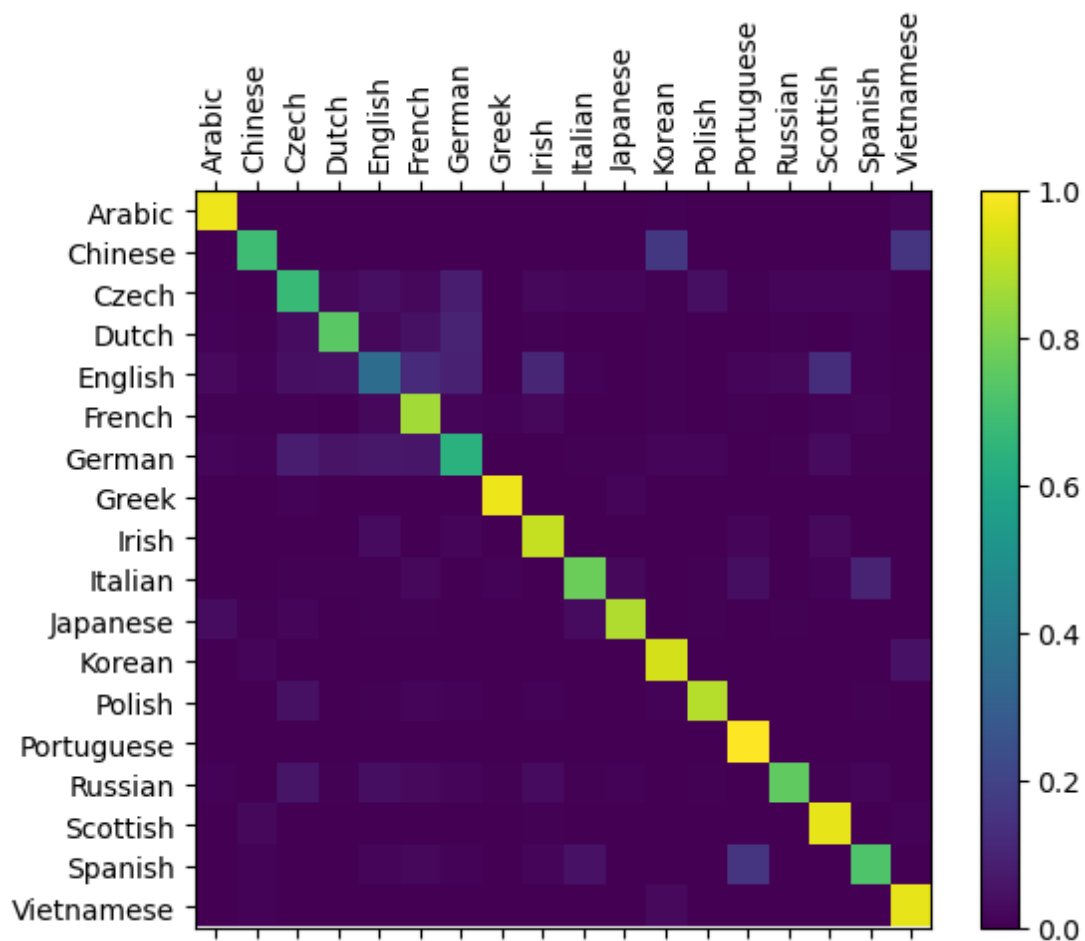
网络结构:

```
LSTM(
  (rnn): LSTM(57, 64)
  (out): Linear(in_features=64, out_features=18, bias=True)
  (softmax): LogSoftmax(dim=-1)
)
```

输出结果:



预测矩阵图：



三．实验总结和心得

为什么 LSTM 网络的性能优于 RNN 网络？

1. RNN 在训练过程中，随着时间步的增加，会遇到梯度消失或梯度爆炸的问题。这使得 RNN 难以学习和记住长时间步的数据依赖。LSTM 通过引入了门控机制（输入门、遗忘门和输出门），能够在长时间步内保持稳定的梯度，从而有效缓解梯度消失和梯度爆炸的问题。
2. LSTM 引入了一个记忆单元（Cell State），这个单元可以沿着序列进行传递，几乎不受梯度消失影响。这使得 LSTM 可以保留长期的信息，并在适当的时候丢弃不需要的信息。