

期末总结文档

17051533 张拓

17051536 郑贤港

论文题目

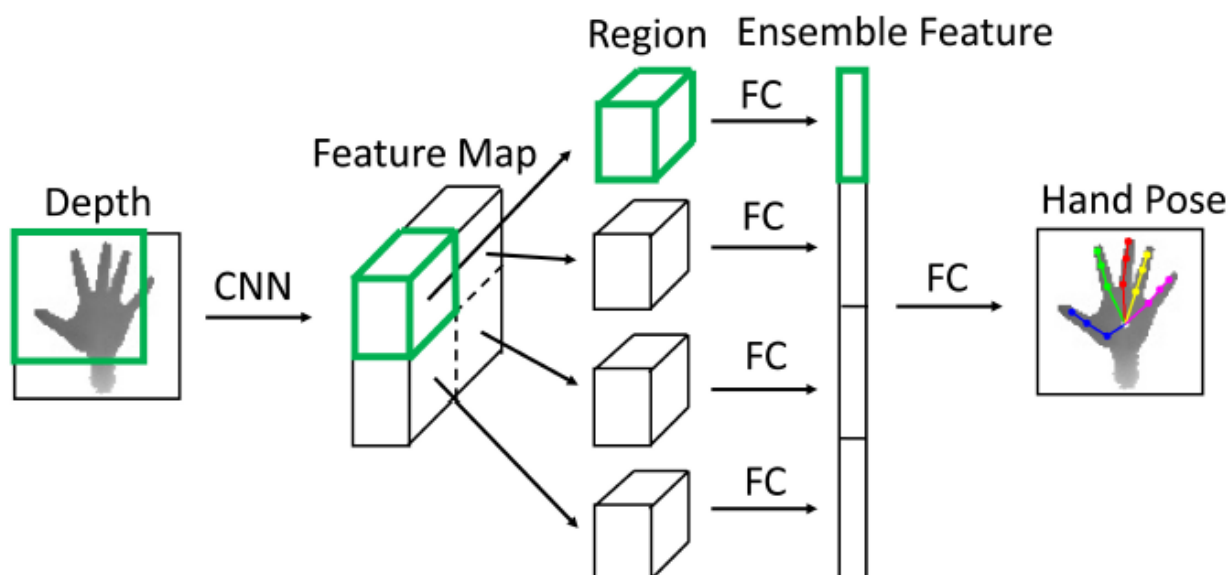
《REGION ENSEMBLE NETWORK: IMPROVING CONVOLUTIONAL NETWORK FOR HAND POSE ESTIMATION》

概述

这篇论文介绍了一个叫“Region Ensemble Net”的网络，翻译过来叫“区域集成网络”。它的主要步骤是，先将原图像通过CNN进行特征提取，然后将特征图通过REN网络进行训练，得到手势。如下图所示。

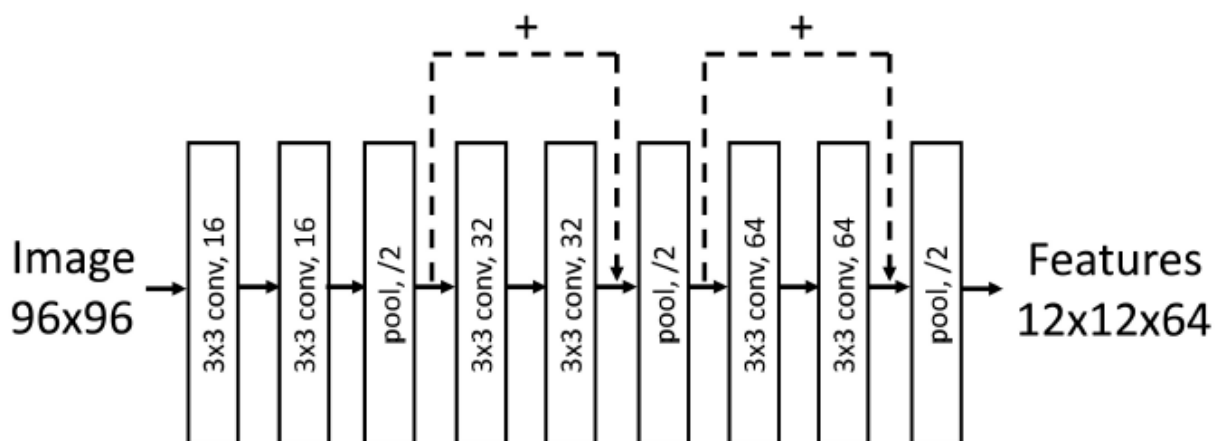


而其中的REN就是论文的主角，它是将特征图分为 $n \times n$ 的区域，然后将这 $n \times n$ 个区域分别通过全连接层，最后再将每个全连接层的输出连接起来通过一个全连接层的到我们需要的手势。如下图。

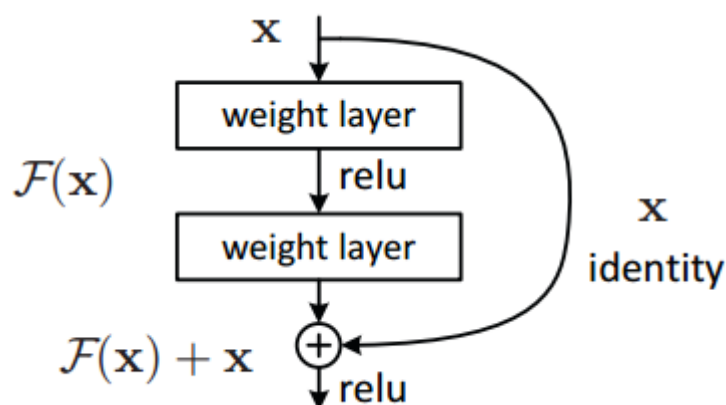


要点

在论文所给的CNN中，作者加入了两个残差层来减少冗余。

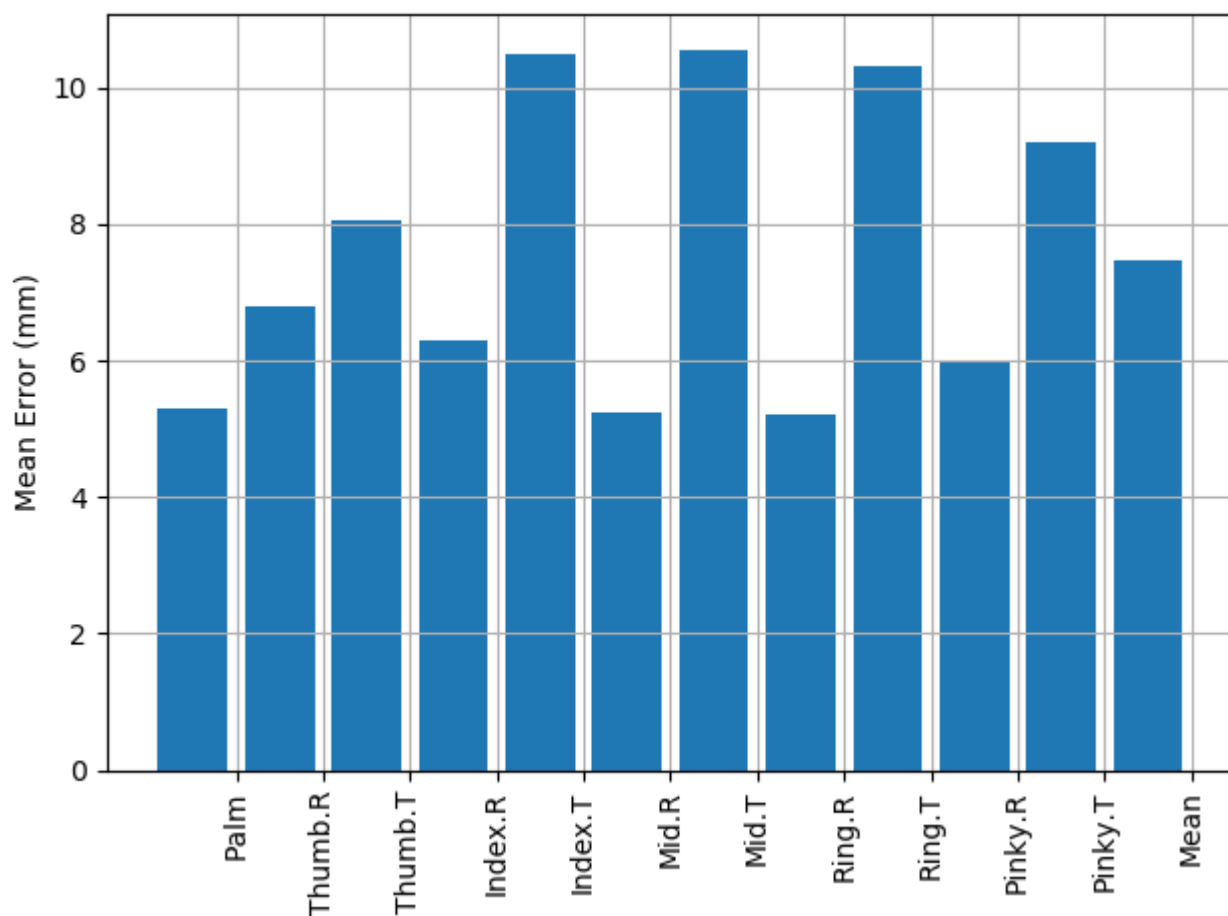


所谓残差网络，简单讲就是在第二层进行线性变化之后激活之前， $F(x)$ 加入了这一层输入值 x ，然后再进行激活后输出。这样做可以避免梯度消失，以及网络退化。



网络评估

以下是网络的评估（不同手指的准确度）。



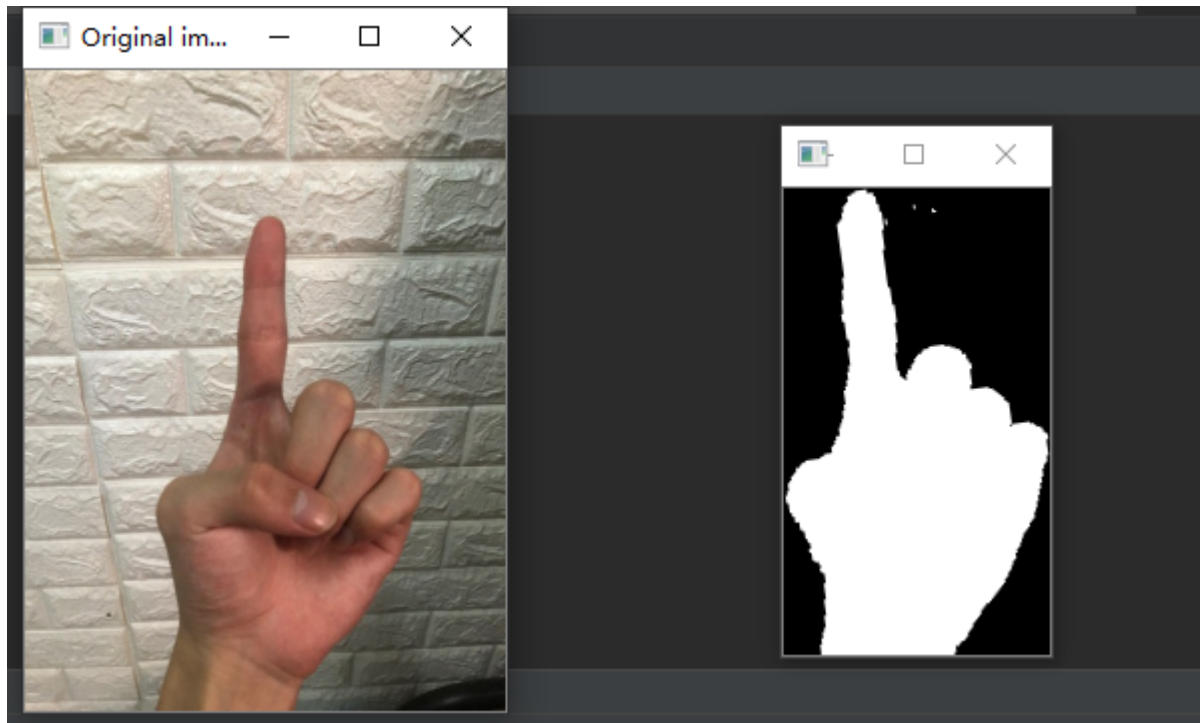
姿态识别效果



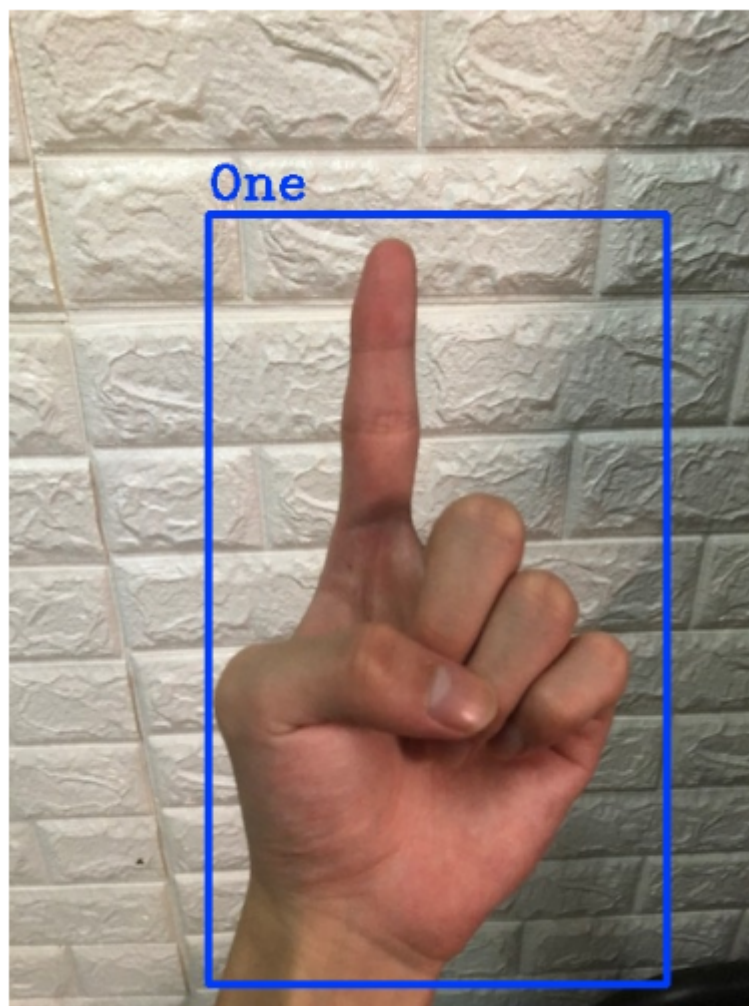
手势识别

手势识别我们采用了一个分类网络，目前只能分类1-5的手势。

我们从输入的图像通过背景去除的方法，将手部的提取出来，并做二值化，作为网络的输入。



得到的分类结果如下



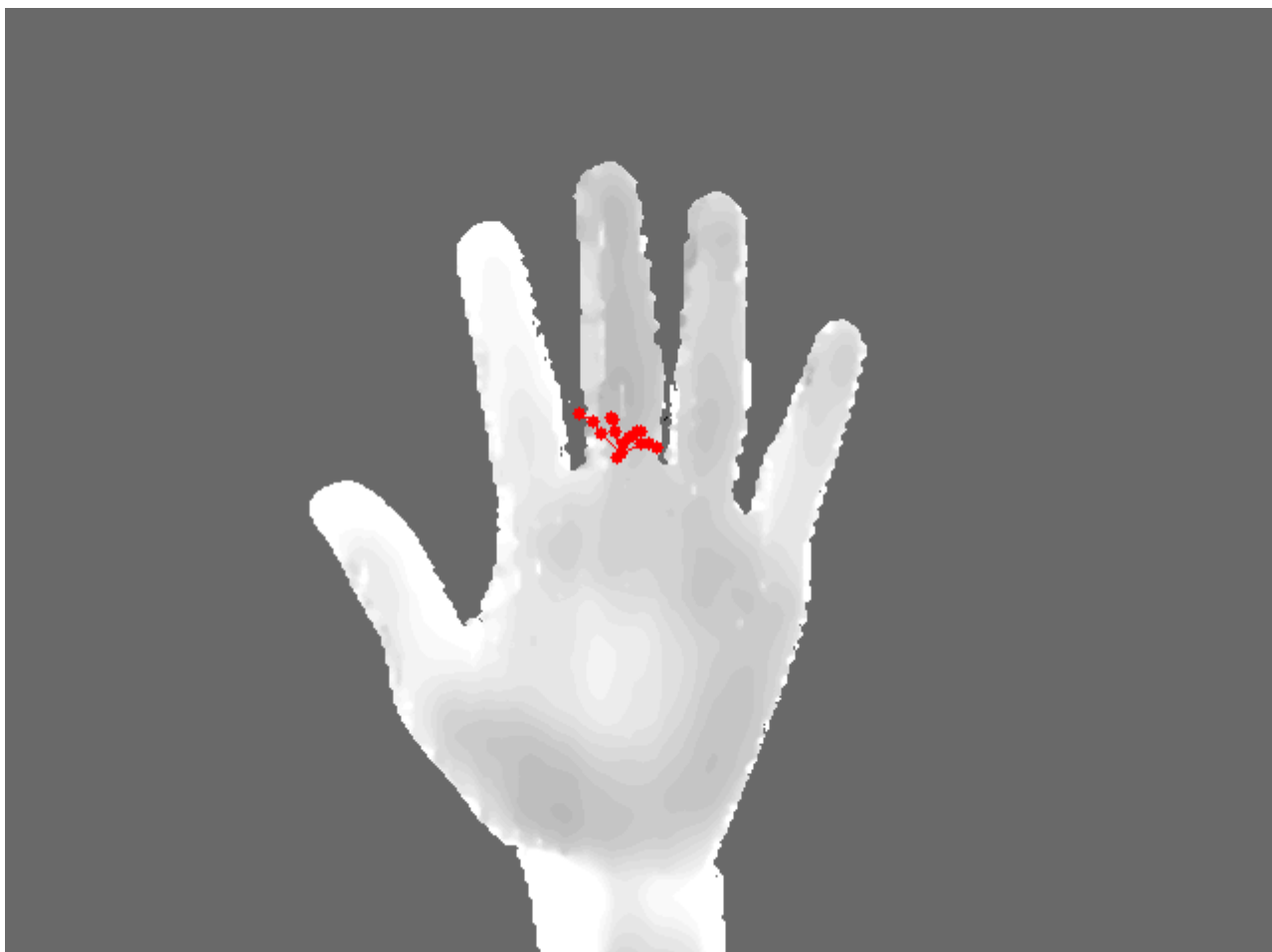
实验配置

深度学习框架：Pytorch

客户端后台框架：flask

数据集：icvl

遇到的问题



在用深度摄像头拍摄的图片作为输入时，发现它的效果很差劲，可能的原因是在输入之前的预处理可能和论文有所偏差，使得这个深度图在识别上出现问题。

小结

这个学期整体收获很大，其中很大一部分原因是原论文所需的caffe环境很难配置，使得我们不得不转用pytorch来复现这个论文的网络。因此我们通过这次复现过程，熟悉了pytorch的使用，并学会如何将一种框架的权重参数转换到另一个框架中，受益匪浅。从结果上来看，我们不算特别成功，但是从无到有的过程让我觉得收获巨大。