

1、总结

这 7 天主要学习了计算机视觉方面的知识，并通过卷积神经网络模型了解了计算机在处理图像上时的流程。感觉学的东西稍微有些多且复杂抽象，接下来将做一个简要总结。

A、模型构建：

最开始学习的神经网络，只是一个简易的层次结构，仅包含了输入层、隐藏层和输出层，并实现一对多 logistic 回归模型，通过不断拟合，不断最小化代价函数，使得最终的映射关系可以给定任意输入得到输出，解决的主要问题包含非线性假设、过拟合和参数优化。

而目前所学的卷积神经网络是为了解决输入对象参数过大，所采取的解决措施，即通过卷积层缩小图片尺寸，深层化，池化层再进一步缩小 2D 尺寸，全连接层计算最终的输出，进而预测结果。由于网络结构过于复杂，且超参数的设置对于新手往往很难，因此推荐迁移学习，即使用开源项目为自己作铺垫。

B、图像处理

有了卷积神经网络作为模型，那么接下来要处理的就是实际的问题，比如滑动窗口进行图像定位、非极大值抑制解决图像定位的不准确。以及后来的三元损失函数解决人脸识别问题到内容、风格代价函数实现图片的风格迁移。

C、神经网络本质

神经网络的本质总结为一个词——拟合，不论是哪一种网络，一般的神经网络或是 CNN 卷积神经网络，都是不断拟合的过程，层次多少以及结构设置完全是为了迎合不同的业务需求，比如图片就需要大量特征，因此选择 CNN。根据结果的输出分为预测模型和匹配问题（目前接触到这两个）预测-即给定输入对象与标签，得到其中的映射关系，进而识别出对象的类别，最简单的就是识别分类问题；匹配-通过最小化代价函数，得到模型中的各个参数，实现业务功能，比如人脸识别，用的即是三元损失函数，而风格迁移用到的是内容代价函数和风格代价函数的最小化处理。其实预测问题也有代价函数，即输出标签与实际标签的差值。

所以，总体来讲，模型的定义完全取决于业务需求，代价函数又取决业务要求，是分类，还是识别还是风格迁移，最后一步使用梯度下降算法或是高级梯度下降算法本质上都是一致的。