

第一周 Alexnet

1. 文字回答： ImageNet 数据集与 ILSVRC 之间的关系是什么？

ImageNet datasets is a meta image database where more than 15 million images were built by ontological sequences upon backbone of WordNet structure. ILSVRC is abbreviation of ImageNet Large Scale Visual Recognition Challenge (ILSVRC) which uses indexed clean and annotated images from ImageNet to evaluate visual recognition algorithms. ILSVRC uses ground truth annotations from ImageNet to allow researchers to compare progresses in object detection.

ILSVRC 使用 ImageNet 的一个子集，总共有大约 120 万个训练图像，50,000 个验证图像，以及 150,000 个测试图像；1000 类别标记。

Large Scale Visual Recognition Challenge 即 ILSVRC(2012~2017)比赛，是基于该数据集的 1000 个类别的比赛。训练集 120 万张图片。从 2010 年起，每年 ImageNet 的项目组织都会举办一场基于 ImageNet 的大规模视觉识别竞赛（ImageNet Large Scale Visual Recognition Challenge, ILSVRC）。在 ILSVRC 竞赛中诞生了许多成功的图像识别方法，其中很多是深度学习方法，它们在赛后又会得到进一步发展与应用。

ImageNet 数据集和 ILSVRC 竞赛大大促进了计算机视觉技术，乃至深度学习的发展，在深度学习的浪潮中占有举足轻重的地位。

2. 文字回答： AlexNet 训练过程的数据增强中，针对位置采取什么方法将一张图片有可能变成 2048 张不一样的图片（1 个像素不一样，也叫做不一样）？

Alexnet adopts Random Cropping 224x224 patches from original images, flipping them horizontally, and changing the density of RGB channels using PCA color augmentation. This data augmentation helped reduce overfitting and decrease the error rates by more than 1%. 原图像是 256*256，我们从中抽取 224*224 的部分，如此一张图片产生 $32*32=1024$ 张图片。然后进行水平翻转变换，一张图片总计产生 2048 张图片。

3. 文字回答： Alexnet 使用的 dropout,在使用过程中需要注意什么？

Dropout 的引用是为了克服 overfitting，而 overfitting 产生的原因可能是训练数据集太少，参数过多，特征维度过高，模型过于复杂，噪声多这几种原因。Alexnet 在使用 dropout 过程中要注意在测试过程中不能使用 dropout。其中在使用 tf 过程中默认在测试层中不使用 dropout 随机失活。

Dropout addresses the question of overfitting caused by small dataset, large number of parameters, high dimensionality, complex models, much noises. Dropout remains effective during training but cannot be used for testing, default setting for tf is not to use dropout during testing process.

AlexNet 最后 2 个全连接层中使用了 Dropout，因为全连接层容易过拟合，而卷积层不容易过拟合。

1. 随机删除网络中的一些隐藏神经元，保持输入输出神经元不变；
2. 将输入通过修改后的网络进行前向传播，然后将误差通过修改后的网络进行反向传

播；

3.对于另外一批的训练样本，重复上述操作 1

4. 文字回答：读完该论文，对你的启发点有哪些？

启发很多，总结有几点：首先学会了几种提升网络性能的方法，比如数据增强，比如随机失活，比如网络结构是由五个卷积层和两个全连接层组合在一起的。这些技巧对之后构建深度神经网络起到非常关键的作用，对于提升网络的性能有着决定性的作用。

There are a lot of inspirations, and a few points to summarize: First, I learned several methods to improve neural network performance, such as data augmentation, dropouts, the Alexnet structure is composed of five convolutional layers and two fully connected layers. These skills played critical roles in building a deep neural network in the future, and applying these techniques result in improved performance of the network.

答：1.深度与宽度可决定网络能力。

2.更强大的 GPU 及更多数据可进一步提高模型性能。

3.图片缩放细节，对短边先缩放。

4.ReLU 不需要对输入进行标准化来防止饱和现象，即说明 sigmoid/tanh 激活函数有必要对输入进行标准化。

5.卷积核学习到频率、方向和颜色特征。

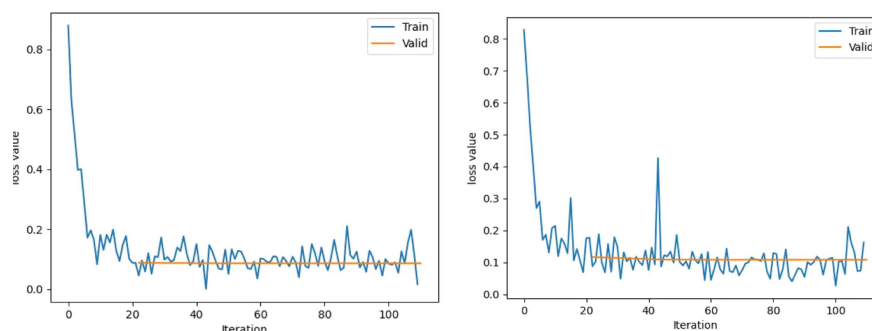
6.相似图片具有“相近”的高级特征。

7.图像检索可基于高级特征，效果应该优于基于原始图像。

8.网络结构具有相关性，不可轻易移除某一层。

9.采用视频数据，可能有新突破。

5. 代码实践：在猫狗数据集上，对比采用预训练模型和不采用预训练模型这两种情况，训练曲线有何种差异，截图打卡。



文字回答：本篇论文的学习笔记以及总结

ImageNet Classification with Deep Convolutional Neural Network

基于卷积神经网络的图像分类

作者：Alex Krizhevsky

单位：加拿大多伦多大学

发表会议及时间：NIPS 2012

在我们的数据分析中，人去做特征工程，人去做分类。

在机器学习中，人去做特征工程，机器学习模型（如决策树、Adaboost、xgboost）去做分类

在深度学习中，卷积神经网络做特征工程的，softmax 或 svm 帮我们做分类。

学习目标：

1. 模型结构设计：ReLU、LRN、Overlapping Pooling、双 GPU 训练
2. 减轻过拟合：图像增强、Dropout
3. 卷积核可视化
4. ILSVRC：大规模图像识别挑战赛

研究背景：

-- ImageNet 数据集包含 21841 个类别，14197122 张图片其通过 WordNet 对类别进行分组，使数据集的语义信息更合理，非常适合图像识别。ILSVRC-2012 从 ImageNet 中挑选 1000 类的 1200000 张作为训练集。

-- 强大计算资源-GPU：高性能计算资源使得大型神经网络可以快速训练。

研究成果：

-- AlexNet 在 ILSVRC-2012 以超出第二名 10.9 个百分点夺冠。

研究意义：

-- 拉开卷积神经网络统治计算机视觉的序幕。

-- 加速计算机视觉应用落地。

机器学习领域：深度学习领域：

特征提取->特征筛选->输入分类器 2012 年 里程碑 AlexNet 特征工程和分类集成于一体神经网络不可解释，黑盒子。

应用领域:安防领域的人脸识别、行人检测、智能视频分析、行人跟踪、交通领域的交通场景物体识别、车辆计数、逆行检测、车辆检测与识别，以及互联网领域的基于内容的图像检索、相册自动分类。

摘要：

1. 在 ILSVRC-2010 的 120 万张图片上训练深度卷积神经网络，获得最优结果，top-1 和 top-5 error 分别为 37.5%，17%。
2. 该网络 (AlexNet) 由 5 个卷积层和 3 个全连接层构成，共计 6000 万参数，65 万个神经元。
3. 为加快训练，采用非饱和激活函数-ReLU,采用 GPU 训练
4. 为减轻过拟合，采用 Dropout。
5. 基于以上模型及技巧，在 ILSVRC-2012 以超出第二名 10.9 个百分点成绩夺冠。

了解了 AlexNet 网络结构及参数计算

ReLU 在所有的全连接层使用。它的优点：

- 1.使网络训练更快。
- 2.防止梯度消失（弥散）。
- 3.使网络具有稀疏性。

LRN(Local Response Normalization)局部相应标准化：有助于 AlexNet 泛化能力的提升，受真实神经元侧抑制启发，在第一个卷积层和第二个卷积层之后。

4) Overlapping Pooling（重叠池化），相对于传统的非重叠池化，采用重叠池化不仅可以提升预测精度，同时一种程度上可以减缓过拟合。

5) 了解了 AlexNet 网络特色及训练技巧，降低过拟合的方法：数据增强和 Dropout（随机

失活，在 FC 层使用，在训练的时候用，测试阶段所有神经元都是激活状态。)

6) 对实验结果分析有了一定的理解。

课程相关链接：

Imagenet 官方网站：

<http://www.image-net.org/>

猫狗数据集：

<https://www.kaggle.com/c/dogs-vs-cats-redux-kernels-edition/data>

alexnet 预训练模型：

<https://download.pytorch.org/models/alexnet-owt-4df8aa71.pth>

Imagenet:

Classnames:

<https://github.com/ashwin/imagenet1000>

Imagenet classnames 中文：

https://befree2008.github.io/2018/10/05/20181005_ImageNet1000_分类名称和编号