湖北汽车工业学院



**商务数据分析课程报告**

**专 业： 计算机科学与技术**

**班 级: 计算机172**

**姓 名： 李轩宇**

**成 绩：**

**完成日期: 2020年6月20日**

**任课教师: 龚家元**

# 一、深度学习概述

深度学习(DL, Deep Learning)是[机器学习](https://baike.baidu.com/item/%E6%9C%BA%E5%99%A8%E5%AD%A6%E4%B9%A0/217599" \t "https://baike.baidu.com/item/%E6%B7%B1%E5%BA%A6%E5%AD%A6%E4%B9%A0/_blank)(ML, Machine Learning)领域中一个新的研究方向，它被引入机器学习使其更接近于最初的目标——[人工智能](https://baike.baidu.com/item/%E4%BA%BA%E5%B7%A5%E6%99%BA%E8%83%BD/9180" \t "https://baike.baidu.com/item/%E6%B7%B1%E5%BA%A6%E5%AD%A6%E4%B9%A0/_blank)(AI, Artificial Intelligence)。 深度学习是学习[样本数据](https://baike.baidu.com/item/%E6%A0%B7%E6%9C%AC%E6%95%B0%E6%8D%AE/12726279" \t "https://baike.baidu.com/item/%E6%B7%B1%E5%BA%A6%E5%AD%A6%E4%B9%A0/_blank)的内在规律和表示层次，这些学习过程中获得的信息对诸如文字，[图像](https://baike.baidu.com/item/%E5%9B%BE%E5%83%8F/773234" \t "https://baike.baidu.com/item/%E6%B7%B1%E5%BA%A6%E5%AD%A6%E4%B9%A0/_blank)和声音等数据的解释有很大的帮助。它的最终目标是让机器能够像人一样具有分析学习能力，能够识别文字、图像和声音等数据。 深度学习是一个复杂的机器学习算法，在语音和图像识别方面取得的效果，远远超过先前相关技术。 深度学习在[搜索技术](https://baike.baidu.com/item/%E6%90%9C%E7%B4%A2%E6%8A%80%E6%9C%AF/1447197" \t "https://baike.baidu.com/item/%E6%B7%B1%E5%BA%A6%E5%AD%A6%E4%B9%A0/_blank)，[数据挖掘](https://baike.baidu.com/item/%E6%95%B0%E6%8D%AE%E6%8C%96%E6%8E%98/216477" \t "https://baike.baidu.com/item/%E6%B7%B1%E5%BA%A6%E5%AD%A6%E4%B9%A0/_blank)，机器学习，[机器翻译](https://baike.baidu.com/item/%E6%9C%BA%E5%99%A8%E7%BF%BB%E8%AF%91/411793" \t "https://baike.baidu.com/item/%E6%B7%B1%E5%BA%A6%E5%AD%A6%E4%B9%A0/_blank)，[自然语言处理](https://baike.baidu.com/item/%E8%87%AA%E7%84%B6%E8%AF%AD%E8%A8%80%E5%A4%84%E7%90%86/365730" \t "https://baike.baidu.com/item/%E6%B7%B1%E5%BA%A6%E5%AD%A6%E4%B9%A0/_blank)，[多媒体学习](https://baike.baidu.com/item/%E5%A4%9A%E5%AA%92%E4%BD%93%E5%AD%A6%E4%B9%A0/10528812" \t "https://baike.baidu.com/item/%E6%B7%B1%E5%BA%A6%E5%AD%A6%E4%B9%A0/_blank)，语音，推荐和个性化技术，以及其他相关领域都取得了很多成果。深度学习使机器模仿视听和思考等人类的活动，解决了很多复杂的模式识别难题，使得人工智能相关技术取得了很大进步，总的来说：深度学习就是一种基于无监督特征学习和特征层次结构的学习方法。

**深度学习的基本原理**：深度学习的概念源于人工神经网络的研究。深度学习通过组合低层特征形成更加抽象的高层表示属性类别或特征，以发现数据的分布式特征表示。

**深度学习于神经网络学习对比：（如图1所示）**

**相同点**:

二者均采用分层结构，系统包括输入层、隐层(多层)、输出层组成的多层网络，只有相邻层节点之间有连接，同一层以及跨层节点之间相互无连接， 每一层可以看作是一个对数回归模型。

**不同点:**

神经网络:采用BP算法调整参数，即采用迭代算法来训练整个网络。随机设定初值，计算当前网络的输出，然后根据当前输出和样本真实标签之间的差去改变前面各层的参数，直到收敛;

深度学习:采用逐层训练机制。采用该机制的原因在于如果采用BP机制，对于一个深层网络(7层以上)，残差传播到最前面的层将变得很小，出现所谓的梯度扩散。



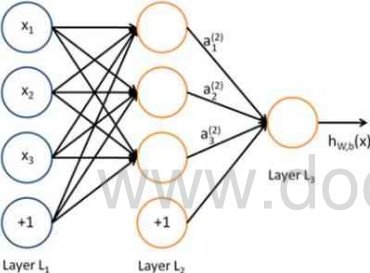


图1

# 深度学习训练过程

**下面以Wake- Sleep算法为例，**wake-sleep算法的主要目的是学习原始数据特征并能正确恢复原始数据。该算法主要分为两个阶段，即"wake"阶段与"sleep"阶段，其中"wake"阶段用来学习生成权重（generative weights），"sleep"阶段用来学习识别权重（recognition weights）。

第一步:采用自下而上的无监督学习。

1)逐层构建单层神经元。

2)每层采用wake-sleep算法进行调优。每次仅调整一层，，逐层调整。

这个过程可以看作是一个特征学习的过程，是和传统神经网络区别最大的部分。

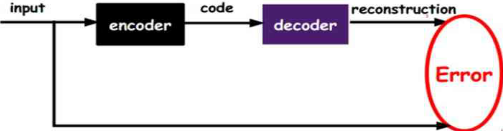
1)wake阶段:

认知过程，通过下层的输入特征( Input)和向上的认知(Encoder)权重产生每一层的抽象表示(Code) ，再通过当前生成(Decoder)权重产生一个重建信息( Reconstruction)，计算输入特征和重建信息残差，使用梯度下降修改层间的下行

生成(Decoder) 权重。也就是“如果现实跟我想象的不一样，改变我的生成权重使得我想象的东西变得与现实一样”。

2) sleep阶段:

生成过程，通过上层概念(Code) 和向下的生成(Decoder) 权重，生成下层的状态，再利用认知(Encoder) 权重产生一个抽象景象。利用初始上层概念和新建抽象景象的残差，利用梯度下降修改层间向上的认知(Encoder) 权重。也就是“如果梦中的景象不是我脑中的相应概念，改变我的认知权重使得这种景象在我看来就是这个概念”。如图2所示.

图2

第二步:采用自顶向下的监督学习。

这一步是在第一步学习获得各层参数进的基础上，在最顶的编码层添加一个分类器(例如logistic回归、 SVM等)，而后通过带标签数据的监督学习，利用梯度下降法去微调整个网络参数。深度学习的第步实质上是一个网络参数初始化过程。区别于传统神经网络初值随机初始化，深度学习模型是通过无监督学习输入数据的结构得到的，因而这个初值更接近全局最优，从而能够取得更好的效果。如图3所示。

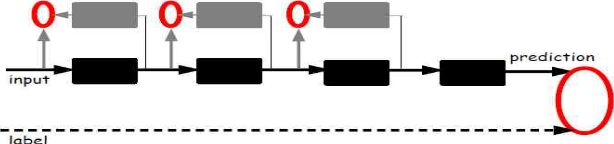


图3

# 深度学习常用框架

在开始深度学习项目之前，选择一个合适的框架是非常重要的，因为选择一个合适的框架能起到事半功倍的作用。研究者们使用各种不同的框架来达到他们的研究目的，侧面印证出深度学习领域百花齐放。在深度学习初始阶段，每个深度学习研究者都需要写大量的重复代码。为了提高工作效率，这些研究者就将这些代码写成了一个框架放到网上让所有研究者一起使用。接着，网上就出现了不同的框架。随着时间的推移，最为好用的几个框架被大量的人使用从而流行了起来。 全世界最为流行的深度学习框架有 TensorFlow 、Caffe、Theano、Keras、PyTorch、MXNet等，如下图4所示。这些深度学习框架被应用于计算机视觉、语音识别、自然语言处理与生物信息学等领域，并获取了极好的效果。

图4

# 深度学习框架实例

**以caffe为例介绍：**Caffe的全称是Convolutional Architecture for Fast Feature Embedding，它是一个清晰、高效的深度学习框架，核心语言是C++，它支持命令行、Python和MATLAB接口，既可以在CPU上运行，也可以在GPU上运行。

Caffe的优点是简洁快速，缺点是缺少灵活性。不同于Keras因为太多的封装导致灵活性丧失，Caffe灵活性的缺失主要是因为它的设计。在Caffe中最主要的抽象对象是层，每实现一个新的层，必须要利用C++实现它的前向传播和反向传播代码，而如果想要新层运行在GPU上，还需要同时利用CUDA实现这一层的前向传播和反向传播。这种限制使得不熟悉C++和CUDA的用户扩展Caffe十分困难。

Caffe凭借其易用性、简洁明了的源码、出众的性能和快速的原型设计获取了众多用户，曾经占据深度学习领域的半壁江山。但是在深度学习新时代到来之时，Caffe已经表现出明显的力不从心，诸多问题逐渐显现（包括灵活性缺失、扩展难、依赖众多环境难以配置、应用局限等）。尽管现在在GitHub上还能找到许多基于Caffe的项目，但是新的项目已经越来越少。

Caffe的作者从加州大学伯克利分校毕业后加入了Google，参与过TensorFlow的开发，后来离开Google加入FAIR，担任工程主管，并开发了Caffe2。Caffe2是一个兼具表现力、速度和模块性的开源深度学习框架。它沿袭了大量的 Caffe 设计，可解决多年来在 Caffe 的使用和部署中发现的瓶颈问题。Caffe2的设计追求轻量级，在保有扩展性和高性能的同时，Caffe2 也强调了便携性。Caffe2 从一开始就以性能、扩展、移动端部署作为主要设计目标。Caffe2 的核心 C++ 库能提供速度和便携性，而其 Python 和 C++ API 使用户可以轻松地在 Linux、Windows、iOS、Android ，甚至 Raspberry Pi 和 NVIDIA Tegra 上进行原型设计、训练和部署。

极盛的时候，Caffe占据了计算机视觉研究领域的半壁江山，虽然如今Caffe已经很少用于学术界，但是仍有不少计算机视觉相关的论文使用Caffe。由于其稳定、出众的性能，不少公司还在使用Caffe部署模型。

**Caffe的优势：**

1.上手快:模型与相应优化都是以文本形式而非代码形式给出。Caffe给出了模型的定义、最优化设置以及预训练的权重方便立即上手。

2.速度快:能够运行最棒的模型与海量的数据。Caffe与cuDNN结合使用，测试AlexNet模型，在K40上处理每张图片只需要1.17ms.

3.模块化:方便扩展到新的任务和设置上，可以使用Caffe提供的各层类型来定义自己的模型。

4.开放性:公开的代码和参考模型用于再现。

**Caffe-Mnist实例：**

手写数字图片：60000张训练库图片，10000张测试库图片。

获取代码： ./ get\_ mnist.sh

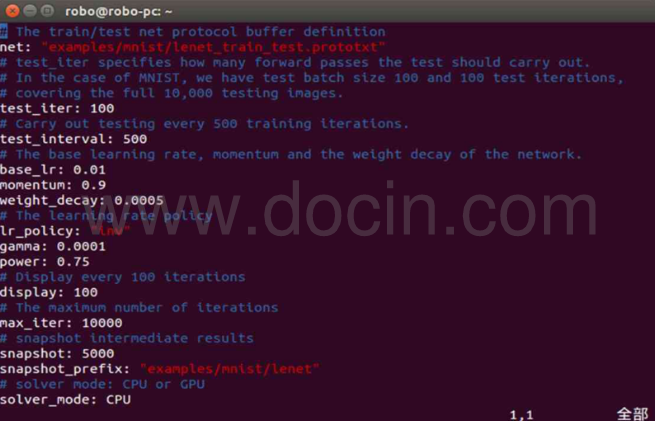
**Mnist数据集的格式转化**，如图5所示。

格式转换代码：./create\_ mnist.sh



图5

**全局参数配置(lenet solver.prototxt)** ，如图6所示

图6

**LeNet网络( lenet\_train\_ test. Prototxt)**

layer {

name: "mnist"

type: "Data

transform param {

scale: 0.00390625

}

data\_ param {

source: "mnist\_ train\_ Imdb"

backend: LMDB

batch size: 64

}

top: "data"

top: "label"

layer {

name: "conv1"

type: "Convolution'

param {Ir\_ mult: 1 }

param{Ir\_ mult: 2 }

convolution param

{

num\_ output: 20

kernel size: 5

stride: 1

weight fller {

type: "xavier"

bias\_ fller {

type: "constant"

}

bottom: "data"

top: "conv1"

}

**训练与测试：**如图7所示

训练测试代码：./train\_ lenet.sh

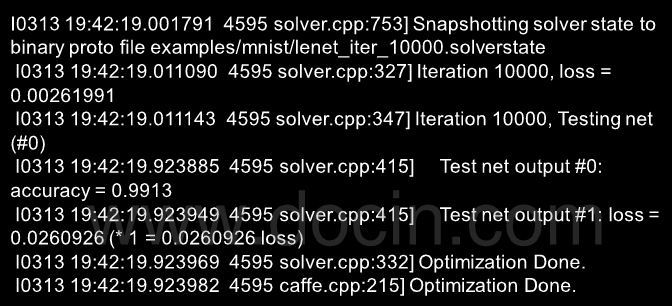


图7

# 总结体会

在计算机视觉领域，深度学习可以应用于图像分类（ResNet50），目标检测（Mask R-CNN），图像分割（UNET），图像恢复等领域。在自然语言处理领域，深度学习能够处理音频和文本数据，可应用于语言翻译（GMNT谷歌翻译）、语音识别、语音合成、情感分析以及自然语言生成等领域。在医学领域，深度学习可以应用于病变检测。比如检利用CT，MRI，造影图像判断病人是否患有某种疾病。深度学习框架给予了深度学习的更方便的模板，我们在进行深度学习的研究过程中要把深度学习框架放在首位。我们根据自己的需要，使用已有的模型，模型的参数也可以自由训练得到，在自己已有模型的基础上增加自己的layer，或者是在顶端选择自己需要的分类器和优化算法（比如常用的梯度下降法）。我相信，在未来我将深度学习真正掌握后，就可以致力于人工智能的周边研发中去，为人类进步贡献自己的力量。