

3D 打印

3D 打印,相比于传统的减材加工制造技术,是一种增材制造。它是一种快速成形技术,以计算机三维模型为蓝本,通过软件分层离散和数控成型系统,利用激光束、热熔喷嘴等方式将金属粉末、陶瓷粉末、塑料、细胞组织等可粘合材料进行逐层堆积粘结,最终叠加成型,制造出实体产品。

减材加工是我们在很长一段时间内使用的材料加工方式,我们通过在一块完整的大型材料上去除不需要的部分,最终得到我们需要的结构。而增材制造则是相反的思路,我们对微小的单元进行组合、拼接,最后形成所需的整体。常见的减材加工包括木雕、石雕等,常见的传统增材制造包括绘画、泥塑、乐高积木等。

3D 打印的增材制造是通过建模->切片->打印->叠加的思路实现的。在通过建模软件完成对打印对象的建模后,切片引擎会将模型分为若干层的厚度确定的 2D 切片。这些切片的信息之后会被发送到 3D 打印机上,3D 打印机通过融化材料的方式将这些 2D 切片打印出来,然后逐层叠加,最后形成一个完整的模型。

技术分类

FDM

FDM,即 Fused Deposition Modeling,熔融沉积(又叫熔丝沉积)。这是市面上较为常见的一种 3D 打印机。将固态丝状热熔性材料加热融化,通过带有一个微细喷嘴的喷头挤出。热熔材料被喷出后,沉积在材料制作面板或前一层已固化的材料上,温度下降后固化完成,打印件所在面板整体下移,如此往复。

批注 [刘1]: 参考 :
<https://blog.csdn.net/fnqtyr45/article/details/103060064>。

批注 [刘2]: 参考 :
<https://www.bilibili.com/video/BV1Lb41127eF>。

FDM 打印成本相对较低，操作和原理也比较简单。但成品表面容易有明显的条纹，且打印速度较慢。

SLA

SLA，即 Stereo Lithography Apparatus，光固化成型技术。SLA 以液态光敏树脂为材料，通过以特定形状投射的紫外线照射使表层凝固成形。在表层成形后，打印件所在的支撑台整体下移，逐次往复此过程最终实现逐层固化。

批注 [刘3]: 参考：

https://www.bilibili.com/video/BV1Yb41127bD/?spm_id_from=333.788.videocard.0。

SLS

SLS,即 Selecting Laser Sintering ,选择性激光烧结。还有一个 SLM(Selecting Laser Melting, 选择性激光熔化) 与 SLS 的原理似乎很相似，就放在一起讲了。

SLS 以粉末状的材料为原料，将粉末材料预热到结晶融化的温度，然后铺到材料面板或前一层固化材料上，再用激光在层截面上扫描，使分模温度升至融化点，然后烧结形成粘接。接着下降打印件所在面板，继续铺粉、烧结，如此往复。

注意，SLS 中，打印件实际上是被埋在大量的材料粉末中的，具体可以看一下视频理解。

批注 [刘4]: 参考：

<https://www.bilibili.com/video/BV16x411o752?from=search&seid=8201389347781238852>，
<https://www.bilibili.com/video/BV17t4y1S7WW?from=search&seid=6965158707839997449>。

PolyJet

PolyJet，即聚合物喷射技术。通过喷射光敏聚合物，由 UV 紫外灯定向地对喷射出的光敏聚合物进行固化。完成一层固化后，下降打印件所在面板，逐层往复。

LOM

LOM，即 Laminated Object Manufacturing，叠层实体制造（也叫分层实体制造）。LOM 使用片状固态材料为基础材料，打印时将背面涂有热熔胶的片状材料加热，粘贴到前一层材料或基底上。然后用激光对材料进行切割，得到需要的轮廓。完成切割后，下降打印件所在面板，如此往复。

批注 [刘5]: <https://www.bilibili.com/video/BV1fk411T7DC?from=search&seid=12615724641333611951>。

打印材料

3D 打印的材料包括以下几种：

ABS 塑料，即 Acrylonitrile Butadiene Styrene Copolymer。

PLA 塑料，即 Polylactic Acid。

光敏树脂，是一种高分子胶状物质。成形产品外观平滑。

橡胶类材料，适合应用于要求防滑或有柔软表面的产品。

金属材料，通常要求金属粉末。常用的金属材料包括钛合金、钴铬合金、不锈钢和铝合金等。以钛合金为主的金属材料打印件被广泛应用于航天航空和生物医学领域。

陶瓷材料，硬度高、密度低、耐高温、耐腐蚀。在航空航天、汽车、生物医疗领域等有应用。但陶瓷打印所需技术较困难，成本较高，难以广泛推广。

复合型石膏材料，是一种全彩色的打印材料。

打印算法

切片算法

切片算法用于将复杂的三维模型转换为平面轮廓信息，以便让 3D 打印机能够逐层构建物体。

三维模型的表面一般由三角形组成。算法需要分析每个切平面和每个三角形是否相交，如果切平面和三角形有交点，就记录下交点。按一定的切平面高度，从上到下对整个模型进行交点记录，然后把相同高度每个高度的交点进行连接，就得到了模型的切平面。

路径规划算法

由于 FDM 方式的打印，喷嘴大小固定，因此它实际是一个用线条填充平面的过程。这

批注 [刘6]: 参考：

<https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CMFD&dbname=CMFD201602&filename=1016751287.nh&v=cYaU75rxtMaKNQbaDepCVCDkm%25mmd2FWvWuSiQjl9tZViOfGz5ibQ83r%25mmd2FLLc6WAudk6gQ>。

就要求必须由一个算法，确定打印时用线条填充平面的方式。

常见的路径规划包括单项扫描、多向扫描、十字网格扫描、螺旋扫描、Z 字型扫描、复合扫描。具体的内容可以看上面的参考论文。

电池

批注 [刘7]: 参考：

<https://www.zhihu.com/question/20507028>。

锂电池

批注 [刘8]: 参考：

<https://www.bilibili.com/video/BV19h411d7zu?t=23> ,
<https://www.bilibili.com/video/BV18K4y1a74Z?from=search&seid=17079580509588462186> ,

目前最常见的电池就是锂电池，其原因是锂于其他金属相比有着极高的比容量（ 3861mAh/g ），即每克锂能放出的电量为 3861mAh 。锂电池可以根据锂元素存在的形态分为金属锂电池（或锂原电池）和锂离子电池。目前，锂离子电池是被广泛使用的电池种类。

金属锂电池以锂作为负极，而锂离子电池则让锂以离子形态存在在电解质中。

锂离子电池根据其聚电解质不同还可分为液态锂离子电池（电解质为液态）和聚合物锂离子电池（电解质为凝胶状或固态）。锂离子电池的负极一般都为石墨，正极则有多种可能，因此一般都会用正极材料来对锂离子电池进行命名，如磷酸铁锂电池、钴酸锂电池、锰酸锂电池等。

苹果目前在 iPhone 和 apple watch 中使用的是锂离子电池，而在 ipad 和 MacBook 中使用的是锂聚合物电池。而现在常见的电动车的电池包括三元锂电池和磷酸铁锂电池（两者的区别主要体现在，三元锂电池在能量密度、快充能力上有优势，而磷酸铁锂在安全性、循环寿命、价格上有优势）。

柔性电池

批注 [刘9]: 参考：

<https://zhuanlan.zhihu.com/p/27655943> ,
<https://www.zhihu.com/question/20729824/answer/84438529>。

近些年柔性、可弯曲华成为了电子设备发展的潮流，可弯曲手环、手机等产品开始逐渐涌现。对于手机这类有一定厚度的产品尚可继续使用刚性电池，但对于更为轻薄的手环、手表等产品，就需要使用同样具有柔性的电池。

目前发展的柔性电池大多为电解质为固体的电池。一般的思路是电池的正负极和电解质都使用极薄的固体。原本是脆性的陶瓷材料，在厚度降到毫米及以下后可以弯曲的，材料

会获得柔性，正是因此，以陶瓷作为固体电解质的电池具有弯曲性。

目前，柔性电池有着安全（全固体电池，不存在泄露）、轻薄、高效的优点。但其成本较高，制作工艺复杂，技术尚不成熟；且使用固态电池的柔性电池难以实现快充。

机器人

机器人涉及了之前提到过的大量知识，包括传感器、SLAM 定位、CV 算法等等，它是一个综合性极强的产业。

机器人的概念正在逐渐地泛化。以往我们提到机器人，往往会想起人形的机械设备。但现在，只要是帮助人类完成复杂或重复工作的机器设备都会被称为机器人。

机器人中较为常见的一个领域是让机器模仿人或其他生物的走路、跳跃或其他运动。它不仅包括在一般情况下的正常运动，还包括在受干扰情况下(被推动)或不良环境下(雪地、沙地、碎石表面)的运动。Boston Dynamic (波士顿动力公司) 就是一个专门从事这方面产品开发的公司，具体可以参考这个[链接](#)。

MSRR (Modular Self Reconfigurable Robot，模块化可重组机器人) 也是机器人的一大发展方向。这一机器人的设计思路就是电影《超能陆战队》中主角最开始设计的机器人的思路。MSRR 强调多个机器人单元组合所带来的可能性。MIT 的 ChainForm 可以通过小型单位的组合，将基础单元拼接成机器人的腿、游戏手柄、智能手环等。香港中文大学理工学院的 FreeBot 采用磁力实现球体与球体间的连接，多个球体单元就好像蚂蚁群一样，能够互相连接、互相运动，实现群体化的复杂运动。

批注 [刘10]: 参考：
<https://www.youtube.com/watch?v=uhND7Mvp3f4>。

批注 [刘11]: 参考：
<https://zhuanlan.zhihu.com/p/33852516>。

批注 [刘12]: 参考：
<https://mp.weixin.qq.com/s/NzLRS9Hhun1nba79GGH9jg>，
<https://www.bilibili.com/video/BV1HZ4y1V7YZ?from=search&seid=10167448553972230090>。

超级计算机

超级计算机是为国家高科技领域和尖端技术研究提供算力的设备。在科学领域，超算可以模拟大气、气候、海洋，预测地震、还笑等自然灾害；在生产领域，超算用于地下采煤、高空作业、爆破工作、石油勘探等领域的数据处理和分析；在生物领域，蛋白质结构模拟、基因测序等领域都需要超级计算机。可以说，一个国家的超级计算机提供的强大算力是这个国家先进科技发展的一大基础。

我国著名的超级计算机包括神威太湖之光（93pFLOPS。FLOPS，即 Float Point Operation per Second，每秒浮点运算次数。93pFLOPS 即每秒可进行 9.3 亿亿次浮点运算）、天河二号（33.9pFLOPS）等。

量子计算机

量子计算机是基于量子力学原理设计的计算机，它拥有并行计算的能力，在算力上远超现在的所有计算机。

传统的计算机信息单元 bit 只能固定地表示一个状态；而量子计算机中，基本信息单元 qubit 可以同时表示多种状态。因此，量子计算机的运算能力随着基本信息单元的增加而呈现指数型的上升。

由于对量子计算机的原理了解不深，因此这里不多赘述。

批注 [刘13]: 参考：
<https://www.bilibili.com/video/BV1Dx411B7sR>。

批注 [刘14]: https://www.bilibili.com/video/BV1LJ411C7sT/?spm_id_from=333.788.videocard.2 ,
<https://www.bilibili.com/video/BV1qx411z7eT> ,
<https://www.bilibili.com/video/BV1Eb411u7nK?from=search&seid=8610182263355054943>。

算法

这里有很多内容在之前都详细讲过,因此这里更多地是一种结合我个人理解的概括性的描述。

CV

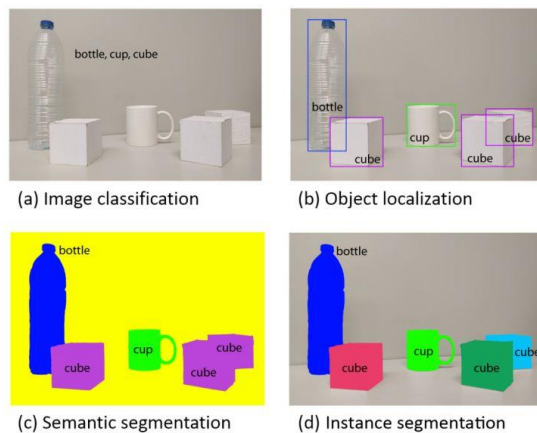
CV, 即 Computer Vision, 计算机视觉。计算机视觉的基本研究对象是图像, 主要工作是从**图像中识别出物体、场景和活动**。计算机视觉有 4 个主要任务: 图像分类、目标定位、语义分割、实例分割, 结合这四个任务, 还可以实现用一般的自然语言对图片内容的描述。在此基础上, 我们希望将计算机视觉拓展到视频中, 实现对目标对象的运动的跟踪以及实现对视频内容的语言描述。

图像分类, 要求就给定的图像, 给出其中所含有的物品类别。

目标定位, 确定各个类别的物体分别在图像中的什么位置, 用方框将其圈出。

语义分割, 确定每一个像素对应的具体的物体类别。

实例分割, 若某张图中存在同一类别的多个物体, 则需要对同一类别下的不同物体进行分别标注。



图像分类、目标定位、语义分割与实例分割

目标检测

目标检测算法的主要目的是将需要识别的目标从图像或视频中检测出来。

Faster RCNN

Faster RCNN 是一种 two-stage 的目标检测算法，即算法先通过一个神经网络对不同目标区域进行划分，再通过一个神经网络对区域内物体进行判断。

Two-stage 的优点在于检测精度高，缺点在于检测速度略慢。

YOLO

YOLO 是一种 one-stage 的目标检测算法，即直接由一个神经网络确定图片中的物体及其所属的类别。YOLO 也可被认为是 You Only Look Once 的缩写。

One-stage 的优点在于检测速度快，缺点在于精度略低。

手势识别

谷歌有一个开源的、跨平台的手势识别算法——MediaPipe Hands。

特征点匹配

SIFT

SIFT , Scale invariant Feature transform ,即尺度不变特征转换。SIFT 对旋转、缩放、光线特征有良好的鲁棒性,对视角变化也有一定的稳定性,被广泛应用于三维重构、图像拼接等领域。

SURF

SURF , Speed Up Robust Feature ,即加速稳健特征。它,且相比 SIFT 的速度更快。它善于处理模糊和旋转的图像,但不善于处理视角和光照变化。

光流算法

光流法是用于视频分析的,它用于检测视频中的某一对象在不同帧之间的运动,也可用于进行特征跟踪。

比较有名的光流算法有 Lucas Kanade 稀疏光流算法。它有是三个基本假设:目标的像素在不同帧之间的亮度、颜色不发生变化;图像上的点随时间移动缓慢,不存在剧烈变化;场景中相同表面的相邻点有相似的运动。

AI

AI,即 Artificial Intelligence,人工智能。机器学习是人工智能领域的一个算法分支,而深度学习和强化学习(一般提的比较少)又是机器学习下的一个领域。

机器学习与深度学习的差别

机器学习与深度学习实际上都是依赖特征来实现问题解决。两者的一个较大的差别在于,机器学习需要由人来给出特征,而深度学习则通过学习的过程自己实现特征的提取。

批注 [刘15]: 参考:

<https://www.bilibili.com/video/BV1Qb411W7cK?p=1>。

回归与分类

机器学习和深度学习最终都希望解决两类问题——回归和分类。回归就是指对某一个特征值进行量上的回答，如根据特定品种的小麦品种、天气状况、病虫害状况预测今年的小麦产量；分类则是对一个特征值质上的回答，如给出一张图片，判断图片中的对象是猫还是狗。

机器学习

机器学习的算法可以大概分为监督学习、非监督学习、强化学习三类。

监督学习

监督学习指使用已标记的数据来训练模型。模型在每次训练时都会根据自己预测的结果和实际给定的标记结果来确定自己的预测是否正确，从而对模型内部的参数进行调整。

监督学习常见的任务包括分类任务和回归任务。

监督学习的主要方法包括 K-邻近算法、决策树和支持向量机等。

K-邻近算法

K-邻近算法，又称 KNN，K-Nearest Neighbors。简单的说，KNN 算法通过确定新加入点周围的已有点的分类来判断新加入节点的分类。

具体地说，首先我们会有一组已确定分类的数据。在新加入一个点后，求这个点与剩余所有点在各个特征上的欧氏距离，然后取出其中的前 K 个点（K 由自己确定）。然后，对这 K 个点到新加入点的距离和他们所属的分类进行加权平均，确定新加入点可能的分类。

决策树

决策树将不同特征集中在一棵树中。当遇到新的样本时，算法将新样本的特征与树结构进行逐层比较，最终到达的叶节点就是新样本对应的分类。

批注 [刘16]: 参考：
<https://www.zhihu.com/question/27138263>。

支持向量机

支持向量机，又称 SVM，Support Vector Machine。它是一种二分类模型，以将样本投射到二维平面为例，它用于找出分割两组数据点的、尽可能宽的间隔（以两条边界线和一条中间线表示）。当引入新的样本时，也将其投射到同一个二维平面，通过确定新点所在的位置落在间隔的中间线的哪一侧，判断其所属的类别。

以上情况只是最基础的、能实现简单二分的情况。现实中很多情况是不能实现简单二分的，因此还会有加入了容错率考虑的 Soft Margin SVM 和用于将数据升维的核函数（kernel）。

批注 [刘17]: 参考：

<https://zhuanlan.zhihu.com/p/49331510>，
<https://www.bilibili.com/video/BV1A4411y7qK?from=search&seid=2312705551590001893>，
https://www.youtube.com/watch?v=Lpr_X8zuE8&feature=emb_logo。

自适应提升

自适应提升，即 Adaptive Boost，通常也会被写为 AdaBoost。AdaBoost 会先对模型进行训练，然后评估结果，减少对该模型做对的问题的注意力，增大对该模型做错的问题的注意力。在后续模型发展中，更专注与先前模型不能解决的困难问题。最后，当所有的模型被整合到一个框架下时，既有处理简单问题的模型，又有处理复杂问题的模型。

批注 [刘18]: 参考：

https://www.bilibili.com/video/BV1iA411e76Y/?spm_id_from=333.788.videocard.0，
https://blog.csdn.net/sinat_35821976/article/details/89790976。

非监督学习

非监督学习指使用没有标记的数据来训练模型。模型必须根据数据的特点来对数据自发地实现分析整理，在每次迭代中根据结果与样本的特征进行主动调整。

非监督学习的常见任务包括聚类（类似于分类任务）和数据降维（类似于回归，通过寻找函数关系的方式来实现降维）。

无监督学习的主要算法包括 K-means 算法等。

批注 [刘19]: 参考：

<https://www.bilibili.com/video/BV1KJ411U7PX?p=1>。

K-means

K-means 算法，又称 K 均值算法。K 均值算法用于将看似有规律的若干样本点分为 k 个不同的聚类（k 的数值需要自行确定）。

批注 [刘20]: 参考：

<https://www.naftaliharris.com/blog/visualizing-k-means-clustering/>。

以二维空间为例（即样本有两个特征值），K 均值算法的具体流程如下：

1. 首先需要确定 K 个初始点（确定方法有若干种，最简单的就是随机放置点），每个初始点对应一个簇类。
2. 接下来遍历所有的点，确定它们与初始点之间的距离，到哪一个初始点的距离最小，就将自己标注为距离最小的那个初始点的簇类。完成对所有点的遍历后，所有的点都会有自己的簇类。
3. 通过计算每一个簇类中的所有所属点的质心，将初始点移动到对应质心的位置。
4. 重复 2. 中的步骤，直到所有初始点不再移动或达到最大迭代次数。
5. 最后，每个点所属于的簇类就是 K 均值算法下对应样本的分类。

弱监督学习

由于无监督学习缺乏指定的标签，在实际应用中的性能往往有较大局限性。弱监督学习是指对数据进行不完全的标签，它可以指只对部分数据进行标签，也可以指对数据的标签不精确、不确切。弱监督学习可以分为半监督学习、迁移学习、强化学习三类。

半监督学习

半监督学习中，只有少量数据有标记。通过有标记的数据不足以训练出好的模型，但同时也存在着大量的无标记数据。半监督学习需要通过充分利用少量的有监督数据和大量的无监督数据来改善算法性能。

迁移学习

迁移学习就是将在标注数据上训练得到的数据有效地迁移到未标注数据上。迁移学习可以充分利用既有模型的知识，使机器学习模型在面临新的任务时只需要进行少量的微调，即可完成相应的任务。

强化学习

强化学习没有训练数据告诉机器应当做哪个动作,它通过我们设置的、合适的奖励函数,不断尝试发现各个动作产生的结果,进而自主学习出不同情况下的策略。强化学习的目标就是研究在与环境交互过程中,学习到一种行为策略,以最大化得到的累积奖励。

在2016年于世界围棋冠军李世石交战的AlphaGo程序就是利用强化学习算法获胜的,目前强化学习在游戏、机器人、自动控制、调度、金融、网络通讯等领域都有发展。

深度学习

说到深度学习,就必然会与神经网络相关联,我们通常说的神经网络一般是指人工神经网络(ANN, Artificial Neural Network)。它是一个用大量简单处理单元,经广泛连接而组成的人工网络。

深度学习并不是万能的。深度学习的一些问题包括:它的整个过程是一个“黑盒子”,使用者并不能通过每一步骤所表现出的特征去理解神经网络为何实现决策;深度学习需要大量的数据支撑,如果数据不足则容易出现过拟合的问题(过拟合就是指过分追究细节贴合而忽视了整体的符合趋势,过拟合的模型通常无法很好地应对除了测试集以外的数据);训练神经网络需要极大的算力。

神经网络的构造

如同生物神经网络一样,人工神经网络基本单元也是神经元。神经元会接收与它相连接的上一层神经元的信息,然后根据这些信息判断自身是否要向下一层神经元发送信息(或发送什么样的信息)。简单地说,当前神经元的判断实际就是将之前神经元传递过来的值放入一个多元的函数中,根据函数结果确定继续向下传递的值是什么。

将大量的神经元集中,就得到了层(Layer)。同层间的神经元是相互独立的,不会相互连接。如果我们有两层A、B,将这两个层连接的方式就是将层A中的每个神经元都与层

批注 [刘21]: 参考:
<https://www.bilibili.com/video/BV1bx411M7Zx>。

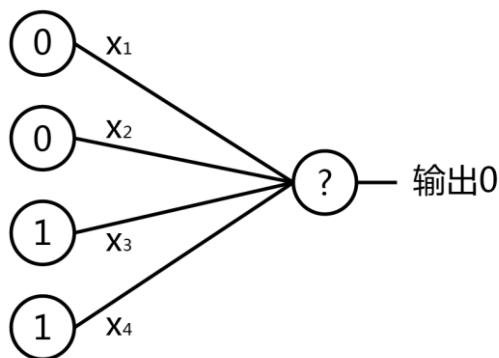
B 中的每个神经元相连接。此时，B 的任何一个神经元的输出都与 A 中的所有神经元输出的信息有关。这样连接的两层就形成了一个最简单的神经网络，其中 B 为输入，A 为输出。

我以一个极为简单的例子，描述一下上面的神经网络的建立和使用。

假设有一张正方形的黑白图片，它只有 4 个像素，我们以 1 表示像素为白，0 表示像素为黑。假设这张图片的左侧为黑，右侧为白，则这个图片可以表示为 $\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$ 。

现在我需要一个神经网络用于判断一张 4 个像素的图片左半边是否为白。这样的判断也可以用 0 和 1 表示，0 代表左半边不为白，1 代表左半边为白。

因此，我们可以构建一个神经网络，左侧为图片的四个像素的输入，右侧为一个 0 或 1 的输出，用于表示“左半边是否为白”这一结论是否为真。对于这四个像素，我们可以进行简单的处理，将第二列的像素直接移到第一行下面，变成一个四行一列的矩阵 $\begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$ （具体的处理方式并不重要，四列一行也行）。神经网络的具体构造如下。为了方便后面的说明，我把这些输入命名为 x_1 、 x_2 、 x_3 、 x_4 。



神经网络构造示意

现在的一个问题是，如果希望输出为 0（说明左半边不为白），在输出位置的这个节点需要用什么样的函数处理之前输入的 x_1 、 x_2 、 x_3 、 x_4 。

一种可行的函数是 $y = \frac{x_1 + x_2 + 0 * x_3 + 0 * x_4}{2}$ 。由于主要关心的是左半边，因此将 x_1 和 x_2 的

系数设为 1, x_3 和 x_4 系数设为 0 (由于结果并不关心右半边的像素特征, 因此 x_3 和 x_4 的系数都为 0)。此时, 如果左半边为白, 则输出为 1, 左半边为黑, 则输出为 0。注意, 这里的系数就是指神经网络中对应的神经元节点的权重。

由此一来, 就形成了一个最简单的神经网络 (当然这里的特征是我们自行定义的, 严格意义上说并不符合深度学习的理念)。当我们输入一张左半边为白的照片 $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ 时, 网络会输出 1; 当我们输入一张左半边为黑的照片 $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ 时, 网络会输出 0。当照片 $\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ 左半边有黑也有白时, 网络会输出 0.5, 表示不确定的结果。

通常, 一般的神经网络在输入和输出层中还会有隐藏层, 简单的说就是在中间增加更多的神经元。这种基础的神经网络又被称为全连接神经网络。

神经网络分类

CNN

CNN, Convolutional Neural Network, 即卷积神经网络。CNN 可以用于图像分类、人脸识别、自动驾驶、语音识别、自然语言处理等。对于语音等内容进行识别时, 需要将语音内容转化为类似图的形式。

卷积神经网络一般可以分为卷积、池化 (下采样)、全连接这三个基础过程。

卷积是指将特定的卷积核矩阵在数据形成的矩阵上滑动, 记录每一个滑动位置的、根据卷积核及数据矩阵中数据计算出的值, 并将该值放到一个新的矩阵——特征图中。

池化是指在用一个像素代替原图上的若干像素, 在保留原图特征的同时压缩其大小。之所以需要池化, 是因为现在图像的大小越来越大, 而在特征上, 一张大图所表现出的特征和小图表现出的特征其实是类似的。因此, 需要将大图的特征进行保留, 同时将图像缩小。池化可以采用所有像素平均值来作为缩小后的像素值, 也可以采用所有像素中的最大值作为缩小后的像素值, 具体的方法会因情境而发生改变。

批注 [刘22]: 参考 :
<https://zhuanlan.zhihu.com/p/49184702>。

全连接是指将图像的行相连或列相连，作为神经网络的输入神经元。完成全连接后，就需要开始对神经网络的隐藏层进行调整。

RNN

RNN，Recurrent Neural Network，即循环神经网络。RNN 主要被用于语音识别、语言模型、机器翻译及时需分析等 NLP 领域。RNN 对处理具有序列特性的数据时非常有效，它能挖掘数据中的时序信息及语义信息（序列特性就是指输入的数据在时间上或逻辑上具有前后相关性）。

RNN 与普通神经网络的区别在于，普通神经网络的输出只依赖于输入及神经元的权重和偏移，而 RNN 的输出在此基础上还会依赖于一个由上一步的输入所决定的时间权重。具体内容在上面给出的参考中有详细的举例。

LSTM

LSTM，Long Short Term Memory，即长短期记忆网络。LSTM 是对 RNN 的拓展。由于多层神经网络很容易引起误差的逐步缩小或逐步放大，最终导致梯度消失或梯度爆炸。梯度消失和梯度爆炸的存在导致一个时刻的时间权重中包含了之前所有时刻的信息，这样不区分重要性的保存方式会导致 RNN 无法很好地将较长序列的前后端的元素进行联系。

因此，LSTM 就通过设计激活函数，只进行选择性的信息存储，保证序列前后端的重要信息能够相互联系。

GAN

GAN，Generative Adversarial Network，即生成对抗网络。GAN 与一般用于处理数据、得出结论的神经网络不同，它是用于实现数据生成的神经网络。现在极为流行的 AI 生成图片、低分辨率图片提升分辨率等就是依靠 GAN 实现的。

GAN 包括生成器（Generator）、判别器（Discriminator）、训练集等单元。GAN 首先

批注 [刘23]: 参考：
<https://zhuanlan.zhihu.com/p/123211148>。

批注 [刘24]: 参考：
<https://www.bilibili.com/video/BV1Vx411j7xF?t=224>，
<https://zhuanlan.zhihu.com/p/76772734>。（梯度消失和梯度爆炸）

批注 [刘25]: 参考：
<https://www.bilibili.com/video/BV1Up411f7S7>。

训练辨别器,使其能够判断什么样的图片是“好”的图片,然后再训练生成器,让其生成“好”的图片。生成器会根据随机的噪声生成若干图片,然后将生成的图片交给训练器。训练器同时从训练集中挑选等量的图片,开始训练。此时,需要训练辨别器,使其能够区分什么图片是训练集中的“好图片”,什么图片是生成器生成的“好图片”。在经过若干次训练后,根据辨别器的训练结果传递给生成器,便于生成器对图片生成的调整。重复这样的步骤,直到辨别器最终无法区分生成器生成的图片与训练集中的图片,此时就说明生成器已经有良好的图片生成能力了。

国内的人工智能相关公司

国内从事深度学习应用的公司包括商汤科技、旷视科技、云从科技、阿里巴巴、百度、腾讯、微软等。

NLP

具体内容在“人机交互接口”中有,这边不多赘述,指把老师课上提到的补充一下。

主要任务

分词、词性标注、语义分析、句法分析。

词袋模型

词袋是指将所有可能词收集到一起的一个集合。词袋模型是指,在表达一句话时,将句子中出现的词汇在词袋中的值设为 1,将句子中未出现的词汇在词袋中的值设为 0。如此,便形成了一个与这个词有着对应关系的向量。

词袋模型的一大问题是维度过高。此外,词袋模型很难表示语义的关联。

批注 [刘26]: 商汤 : <https://www.sensetime.com/>。
旷视 : <https://www.megvii.com/>。
云从 : <https://www.cloudwalk.com/index>。
阿里 : <https://ai.aliyun.com/>。
百度 : <https://ai.baidu.com/>。
腾讯 : <https://ai.qq.com/>。
微软 : <https://azure.microsoft.com/zh-cn/overview/ai-platform/>。

词向量模型

词向量模型简单地说就是将一个普通的单词转化为一个若干维度（50-300 维）的向量进行表示。

大数据

大数据有 4V 特征，Volume（规模性）、Velocity（高速性）、Variety（多样性）、Value（价值性）。

大数据包括四个主要问题：数据采集、数据储存、数据处理、数据封装。（和 004_IOT 中提到的物联网对数据的操作有相似之处。）

主要问题

数据采集

数据采集可以根据互联网数据和物联网数据分为爬虫采集和流式采集。

数据储存

数据储存可以分为分布式文件系统（HDFS，Hadoop Distributed File System）、云存储系统、非关系数据库。

数据处理

数据处理的处理方式包括批处理、流处理、交互式处理、在线处理、图处理等。

数据封装

数据封装包括可视化、查询、统计分析。

推荐算法

搜索引擎主要有数据抓取子系统、内容索引子系统、链接结构分析子系统、内容检索子

批注 [刘27]: 参考：

<https://www.bilibili.com/video/BV1F7411x749?p=1>

•

系统构成。

推荐算法的依据包括：用户的显示反馈或隐式反馈数据、用户画像、项目内容数据、用户生成内容、用户历史数据库、社会关系和社会网络、情景感知（用户所处情景）。

经典的推荐算法有协同过滤和关键规则。

协同过滤主要针对用户的历史行为进行分析，找到与用户相类似的用户群体，依据这些群体的行为特征来对特定用户进行推荐。它适合复杂的非结构化数据，如电影、音乐等；善于发现用户形的兴趣点；不需要用户主动反馈就能直接实现推荐。

关键规则实际上是指某些商品之间存在着一定的联系。比如比较显性的网球拍和网球，也比如比较隐形的啤酒和尿布（买啤酒的男人通常也会买尿布，买尿布的男人通常也会买啤酒）。关联规则的数据源相对简单，只需要已有的交易信息就能进行分析；能直接地反应两商品间的联系，便于策略的拟定。

下面的内容不是非常重要，因此没有展开。

模式识别

CG

区块链

数字图像处理