

人工智能 | 技术能替代人眼吗？

原创：IDZ英特尔开发人员专区2017-10-13

数据不仅涉及数字

还是一种实施高效通信的工具，能够帮助我们做出明智的决策。

有史以来，伽利略对数据分析的理解最为透彻。1610年，他在观察昴星团时发现，用肉眼可以看到6颗明亮的星星，在黑夜，视力非常好的人最多能看到9颗星星。伽利略利用望远镜发现40多颗微弱的亮点散落在6个明亮的星星中间。他在星团图中记录了36颗星星的位置，在远古时代已经发现的星星周围标记了边框。



图 1.伽利略的昴星团绘制图

本质上，伽利略正在利用数据点和数据属性进行数据模式分析，将星星分为明亮的和暗淡的星星，最终将这些明亮的星星划分为星座。

我们现在研究的数据科学和机器学习采用了相同的方法，但是通过机器进行分类，需要采集更多的数据点，具有更复杂的属性。

数据科学热潮

数据科学，顾名思义就是研究数据：信息来源、信息意义和相关性方面的数据，以及对重要信息的潜在价值进行分析。

通过挖掘大量结构化和非结构化数据点，并确定其模式，企业可以获取较高的业务价值。数据分析和数据科学旨在从经验中学习。数据不仅能够帮助我们理解过去发生的事情，还能帮助我们分析其原因。在许多情况下，以上知识能够帮助我们预测未来，有效地管理未来。

数据科学如果利用得当，将是一项强大的工具，为企业带来巨大的竞争优势，从降低成本、提高效率到为业务扩展和客户终生价值带来新的机会。正因如此，全球的技术和业务组织都对数据科学的兴趣日益浓厚。

但是更好的分析需要更准确的推断，更准确的推断需要更完善的思考，反过来需要更出色的工具、更卓越的模式识别和更好的算法。

本文首先介绍了几个**数据科学的核心理念**、然后介绍了**图像分类器**（一种对数据科学原则的应用），还有**车辆分类器上专用的放大工具**，您可以利用这个工具解决类似复杂度的模式识别问题。

核心理念

在深入了解图像分类器之前，首先需要理解某些数据科学理念和术语。

1.分类器

分类器是一款数据挖掘工具，将需要分类的若干数据作为输入数据，尝试预测新数据的类别。

以包含一系列电子邮件的数据集为例，我们知道每封电子邮件都包含特定的属性，如发件人、日期、主题、内容等。根据这些属性，我们能将某人接收的邮件划分为两个或两个以上的类型（类别）。而且，我们能够根据电子邮件的属性预测接收的电子邮件类别。进行上述分类的数据科学算法被称为分类器。

分类器利用上述属性自动对用户的电子邮件进行分类。Gmail* 是自动分类电子邮件系统的绝佳示例，将邮件分为促销邮件、垃圾邮件、社交媒体更新或普通邮件。

2. 监督式学习

监督式学习算法分析训练数据，并生成推断函数，用于映射新的示例。

通过这种方法，利用标记训练数据推断数据模式。

为了训练模型，我们将数据集和相应的正确答案输入模型，模型训练完成后，可以预测未来的值。

利用多个现有的数据点来定义预测未来值的简单规则，又称启发法或基于规则的模型。

3. 无监督学习

数据建模的另一种方法是让数据自行决定，通过这种方法，将许多数据点输入设备，反过来基于复杂分析和模式识别技术，显示了潜在的信息集群。

通过学习来源于大数据的复杂决策规则，以程式化的方式制定明智决策的能力是机器学习的一个主要卖点。

利用机器学习，监督式学习技术（尤其是深度学习）的预测准确性提高了数倍。这种无需人类监督的学习技术目前生成了许多分类器，在许多任务的预测准确性方面优于人类。

这种无监督模型现在被用于各行各业，支持预测个人对电影的评分 (Netflix*)、对图像进行分类（如.com* 和 Facebook *）、进行语音识别 (Siri*)，还有更多其他功能！

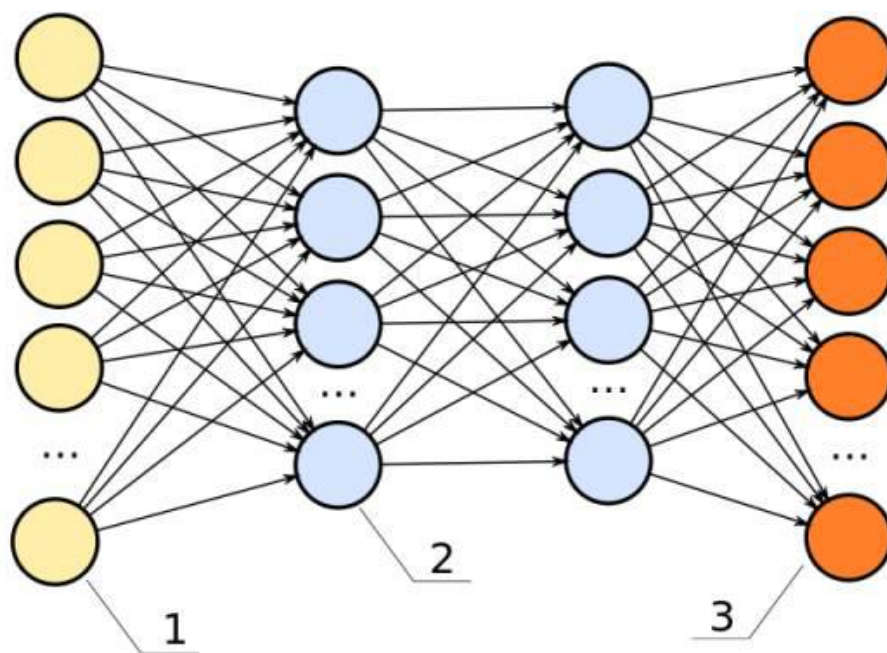
4.神经网络

神经网络是一种编程模式，受了人脑和神经系统（神经元）的结构和功能的启发，支持电脑根据观察性数据学习。

神经网络的目标是制定决策，然后像人脑一样地解决问题。现代神经网络项目通常处理数千至数百万个神经单元以及数百万个连接，但是其复杂性仍比人脑低出许多数量级，接近于一条蠕虫的计算能力。

Robert Hecht-Nielsen 教授是第一代神经计算机的发明人，他在《神经网络初级教程：第一部分》（Maureen Caudill，一名人工智能专家，1989年2月）中将神经网络描述为“.....一种由许多高度互联的简单处理单元构成的计算系统，处理外部输入的动态响应信息。”

神经网络划分为许多层，每一层都由大量的互连节点构成，每个节点都有激活功能。



- 通过输入层将模式传输至网络。
-
-
- 输入层与一个或多个隐藏层通信，在隐藏层中，通过加权连接系统进行实际的处理。
-
-
- 然后将隐藏层连接至输出层，后者负责输出结果。

5.神经元

神经元是人脑的一个基本工作单元，它是一种特殊的细胞，专门负责将信息传输至其他神经、肌肉或腺细胞。

在神经网络的编程世界中，在互联、交互的组件（称为节点或人工神经元）中对生物神经元进行模拟。节点接收输入数据，对数据执行简单的操作，并有选择地将结果传输至其他神经元或节点。

根据模型类型的不同，人工神经元的类型会有所不同。感知和 sigmoid 神经元是两种常用的神经元。

6.深度学习

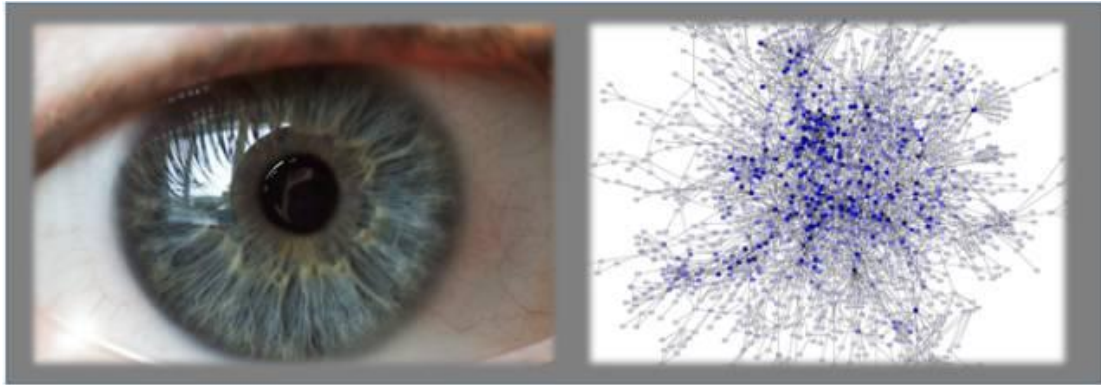
根据层数的不同，神经网络的深浅会有所差异。上图显示了一个输入层、两个隐藏层和一个输出层。神经网络也可以包含若干个隐藏层。和布尔电路的操作一样，利用中间层可以创建多个抽象层。

当我们创建多个抽象层时，深度学习是进行神经网络学习的一系列强大技术。

例如，如果您想针对视觉形状识别构建一个网络，第一层的神经元能够学习识别边缘，第二层的神经元能够学习识别角度，第三层将进一步识别复杂形状，如三角形、圆形或矩形，第四层将利用上述信息识别最终的物体，如滑板或圆形。

直观来讲，深层网络有更多的隐藏层，在性能上优于浅层网络，然而，由于每个隐藏层的学习速度各不相同，训练深层网络的难度非常大。例如，在被称为梯度消失问题的现象中，前层神经元的學習速度比后层神经元的學習速度慢很多。因此，区分三角形、正方形或圆形相对简单，但是学习边缘和角度之间的差异会更难。

案例研究：图像分类器



图像识别

人脑非常强大，人眼同样强大！我们的视觉系统也许是最不受重视的奇迹了。

观察以下数字：

- 8754308
- **8754308**
- *8754308*
- 8754308

大多数人会轻松地认出数字 8754308 ,但是这种毫不费力具有一定的欺骗性。人类大脑的每一个半球均包含一个被称为 V1 的主要视皮层，包含 1.4 亿个神经元，神经元之间有数百亿个连接。此外，人类视觉不只包含一个视皮层，而是一系列视皮层，负责执行越来越复杂的图像处理。而且，人类视觉系统中的超级电脑经过了数亿年的训练（进化），能够适应我们的视觉世界。

这个被称为人眼的强大图像处理器能够分辨苹果和熊、狮子和美洲豹，阅读大量的标志，感知颜色以及进行各种模式识别。对大脑来说，这些任务看似简单；但是对计算机来说，实际上非常难以解决。

数据分析、数据科学和机器学习的进步是通过复杂的处理模型而实现的。有一种模型被称为深度卷积神经网络，能够在较难的视觉识别任务中实现合理的性能。

卷积神经网络

卷积神经网络是一种神经网络，利用相同神经元的多个相同副本，支持网络严格控制实际参数的数量，同时利用大量的神经元处理大型模型。

由于这种架构显著提高了神经网络的训练速度，尤其适合图像分类。更快的训练速度有助于训练深层（多层）网络，这也是进行图像识别和分类所必需的。

一款应用：车辆分类器

利用 TensorFlow* 创建模型，TensorFlow* 是一款利用数据流图进行数值计算的开源软件库。

训练集

创建图像分类器的第一步是创建训练数据集。创建训练数据时需要考虑以下方面：

- 图像的数量越多，经过训练的模型的准确性越高。每个对象类型应包含至少 100 张图像。
-

- 注意训练偏差。如果某张图像中一棵树位于花园中，另一张图像中一棵树位于路边，最终的训练预测将基于背景，而不是基于更为关键的特定对象特性，这是因为训练收集图像共同的特性。因此，收集的场景种类要尽量广泛，如不同的位置、时间、来自不同的设备等。

- 对训练数据进行精细分类，保持类别间的视觉差异，而不是分为包含许多不同外形的大类。否则，您可能最终得到无意义的抽象对象。

- 最后，请确保所有的图像均得以正确标记。

以下示例来自我们的数据集：



图 2：公共汽车的随机图像。



图 3：汽车的随机图像。



图 4：自行车的随机图像。



图 5：摩托车的随机图像。

为了创建车辆分类器，我们从 ImageNet* 为每个类别下载了大约 500 张图像，共有 4 种不同的车辆类别 - 汽车、自行车、公共汽车和摩托车。

训练模型

下一步，我们指定脚本，随机选择用于模型训练的图像。

由于对象识别模型具有数千个参数，有时训练会持续数周。传输学习技术通过获取一个经过完全训练的模型（针对一系列类别），如 ImageNet，然后利用现有的权重针对新类别进行重新训练，极大地缩短了训练时间。

我们采用这个技术训练模型。

在本篇案例研究中，我们只对最后一层进行重新训练，其他层保持不变。我们删除了原始的顶层，在车辆照片上训练了一个新的顶层，照片均不位于用于训练整个网络的原始 ImageNet 类别中。现在，利用这些传输技术训练下层，可以完成其他识别任务，几乎不需要对技术进行修改，而以前只能区分一系列对象。

验证模型

由于我们正训练模型，使它适合示例图像，还需要测试模型能否处理示例集以外的图像。毕竟，识别示例集以外的图像才是车辆分类器的全部目标！

因此，我们把图像分为 3 个不同的子集。

1.

训练集—用于训练网络的图像，其结果用于更新模型的权重。通常情况下，指派 80% 的图像用于训练。

2.

3.

验证集—在训练模型时，利用这些图像定期验证模型。通常，10% 的数据用于验证模型。

4.

5.

测试集—作为测试集，预测分类器实际性能的次数相对较少。

6.

TensorFlow 命令

我们通过以下命令实现 TensorFlow 的传输学习功能。

前提条件

注：*Bazel* 帮助我们通过命令行运行 *TensorFlow*。

训练模型

文件夹结构如下所示：

每一个类别（1 类和 2 类）代表一个车辆类别，如汽车、公共汽车等。每一个文件包含该类别的图像。此处，将 *TensorFlow* 模型指向图像文件夹。

通过 *Bazel* 进行测试

为了进行测试，我们使用了另一个称为 *Label** 的 *TensorFlow* 程序，它是一个位于 *TensorFlow* 目录中的 C++ 程序。

结果

通过几个图像对模型进行了测试。随附每个类别的示例图像和模型得出的结果。



`Itensorflow/examples/label_image/main.cc:205]` 汽车 (0) :

0.729847

Itensorflow/examples/label_image/main.cc:205] 摩托车 (1) :
0.140029

Itensorflow/examples/label_image/main.cc:205] 自行车 (2) :
0.0864567

Itensorflow/examples/label_image/main.cc:205] 公共汽车 (3) :
0.0436665

结果显示模型预测图像为汽车的可信度为 72%。



**Itensorflow/examples/label_image/main.cc:205] 公共汽车 (3) :
0.933695**

Itensorflow/examples/label_image/main.cc:205] 汽车 (0) :
0.0317426

Itensorflow/examples/label_image/main.cc:205] 摩托车 (1) :
0.0192131

Itensorflow/examples/label_image/main.cc:205] 自行车 (2) :
0.0153493

从上图我们可以看出，模型预测图像为公共汽车的可信度为 93%；预测为摩托车的可信度只有 1.9%。



Itensorflow/examples/label_image/main.cc:205] 自行车 (2) :

0.999912

Itensorflow/examples/label_image/main.cc:205] 汽车 (0) :

4.71345e-05

Itensorflow/examples/label_image/main.cc:205] 公共汽车 (3) :

2.30646e-05

Itensorflow/examples/label_image/main.cc:205] 摩托车 (1) :

1.80958e-05



I tensorflow/examples/label_image/main.cc:205] 摩托车 (1) :
0.979943

Itensorflow/examples/label_image/main.cc:205] 自行车 (2) :
0.019588

Itensorflow/examples/label_image/main.cc:205] 公共汽车 (3) :
0.000264289

Itensorflow/examples/label_image/main.cc:205] 汽车 (0) :
0.000204627

新的可能

互联网的发展使每个人都能接触到知识，社交媒体让我们能够随时随地保持彼此之间的联系。现在，随着机器学习的发展，新一轮的创新大潮在许多领域日益凸显。

我们来看一下车辆图像识别的应用！

收费桥



面向汽车的自动化收费已经不足为奇了。但是，您是否注意到收费桥上的一列列卡车长队？

汽车、SUV、卡车和其他车辆类型的收费各不相同。技术使利用传感器自动检测收费成为了可能。为何不对收费的金额进行分类，进一步提高便利性？汽车收费 5 美元，卡车收费 25 美元.....全部实现自动化。

自动驾驶汽车

Google* 和 Uber* 正积极研发自动驾驶汽车。想象一下，未来的道路上全是自动驾驶汽车，道路上有一些人，至少我们可以依靠我们的视觉系统来谨慎驾驶，穿梭于无人驾驶汽车和鲁莽的驾驶者之间。

但是当具有超级计算能力的人眼不再关注其它车辆时，我们需要一台车辆分类器，帮助自动驾驶汽车识别不同的道路交通类型！

我们的安全

如今，FBI、警察局和许多安全公司保护我们和国家的安全。



即使在日常生活中，我们看到琥珀警戒，显示被绑架的小孩在高速公路上的汽车内，以及其他情况。



想象一下在科幻电影里，以完全自动化的形式进行搜索，现在通过 google 地图的聚合、基于数据科学的图像识别和车辆分类算法，这一切都有了现实的可能性。



停车

停车是一个日益严峻的问题，尤其对城市而言。借助强大的车辆分类器，我们将建造完全自动化的高效停车方式，特定的区域和楼层用于停放特定的车辆类型。跑车需要较低的高度，而运动型卡车身型庞大。每层车库的大门上需要安装车辆检测传感器，只允许特定类型的车辆进入停车区。高度上节省了许多空间，确保在有限的空间内停放更多的汽车！

结论

这些示例只是一个开始，我们非常期待收到您的反馈，了解您所作的改进和处理的复杂问题，释放隐藏在数据中的巨大潜力。

无论如何，基于本文涵盖的话题和引用的示例，希望您能和我们一样相信，技术（尤其是模拟人脑和人眼的技术）正在快速地向发展，可能在不久的将来会替代人眼。自动化将替代人工任务，尤其是涉及可视化和分类的重复性任务。

尽管如此，人脑结构和复杂的视觉系统仍有许多谜团尚未解开，随着研究的深入，我们将掀开它们的神秘面纱。