

## 特征提取前沿论文最新进展

2018.10.30 方建勇

- 在本文中，我们考虑二进制相似性问题，其中包括确定两个二元函数是否相似，只考虑它们的编译形式。这个问题在一些应用场景中是至关重要的，例如版权纠纷，恶意软件分析，漏洞检测等。该领域当前最先进的解决方案通过创建将二进制函数映射到向量的嵌入模型来工作。在  $R^n$ 。这种嵌入模型捕获二进制之间的句法和语义相似性，即，类似的二进制函数被映射到向量空间中接近的点。该策略具有许多优点，其中之一是可以预先计算几个二元函数的嵌入，然后将它们与简单的几何运算（例如，点积）进行比较。在[32]中，首先在由手工设计的特征构成的注释控制流图（ACFG）中转换函数，然后使用深度神经网络架构将图嵌入到矢量中。在本文中，我们提出并测试了几种计算注释控制流图的方法，这些方法使用无监督的特征学习方法，而不会产生人为偏差。我们的方法受到自然语言处理社区中使用的技术的启发（例如，我们使用 word2vec 来编码汇编指令）。我们表明我们的方法确实是成功的，并且它比以前最先进的解决方案带来更好的性能。此外，我们报告了函数嵌入的定性分析。我们发现了一些有趣的案例，其中嵌入根据原始二元函数的语义进行聚类。
- 本文提出了一种新的基于点云的场所识别系统，该系统采用深度学习的方法进行特征提取。通过使用在彩色图像上预训练的卷积

神经网络从范围图像中提取特征而无需对额外距离图像进行微调，与使用手工制作的特征相比，已经观察到显著的改进。得到的系统是照明不变的，旋转不变的并且对于与地点标识无关的移动物体是鲁棒的。除了系统本身，我们还为社区带来了一个新的地点识别数据集，其中包含点云和灰度图像，覆盖整个 360° 环境视图。此外，数据集的组织方式使得它有助于分别针对不相关的移动物体的旋转不变性或鲁棒性进行实验验证。

- 背景：Instagram 每天有数百万个帖子，可用于告知公共卫生监督目标和政策。然而，目前依赖于基于图像的数据的研究通常依赖于图像的手工编码，这是耗时且昂贵的，最终限制了研究的范围。自动图像分类中的当前最佳实践（例如，支持向量机（SVM），反向传播（BP）神经网络和人工神经网络）在其准确区分图像内的对象的能力方面受到限制。目的：本研究演示了如何使用卷积神经网络（CNN）提取图像中的独特特征，以及如何使用 SVM 对图像进行分类。方法：从 Instagram 收集水管或水烟（一种具有与香烟类似危害的新兴烟草产品）的图像并用于分析（ $n = 840$ ）。CNN 用于从识别出包含水管的图像中提取独特的特征。构建 SVM 分类器以区分具有和不具有水管的图像。然后比较图像分类的方法以显示 CNN + SVM 分类器如何提高准确性。结果：随着验证的训练图像的数量增加，提取的特征的总数增加。另外，随着 SVM 分类器学习的特征数量的增加，平均准确度增加。总体而言，分类的 420 幅图像中有 99.5% 被正确识别

为水烟或非水烟图像。与仅使用 SVM, CNN 或特征包 (BOF) 的早期方法相比, 这种准确度是一种改进。结论: CNN 提取了更多图像特征, 使 SVM 分类器能够更好地获得信息, 与提取较少特征的方法相比, 具有更高的准确性。未来的研究可以使用这种方法来扩大基于图像的研究范围。

- 状态监测是所有主要过程工业中的常规任务之一。电机, 齿轮, 轴承等机械部件是过程工业的主要部件, 其中的任何故障都可能导致整个过程完全停机, 这可能导致严重的损失。因此, 在发生之前预测任何接近的缺陷是至关重要的。为此目的存在几种方法, 并且正在进行许多研究以获得更好和有效的模型。然而, 它们中的大多数都是基于原始传感器信号的处理, 这是繁琐且昂贵的。最近, 基于特征的状态监测已经增加, 其中仅从原始信号中提取有用特征并且解释用于故障的预测。其中大多数是手工制作的功能, 其中这些功能是根据原始数据的性质手动获得的。这当然需要事先了解数据和相关过程的性质。这限制了特征提取过程。然而, 基于自动编码器的特征提取方法的最新发展提供了传统手工方法的替代方案;然而, 它们大多局限于图像和音频处理领域。在这项工作中, 我们开发了一种基于传统自动编码器和在线顺序极端学习机 (OSELM) 网络堆栈的在线状态监测的自动特征提取方法。该方法的性能与传统特征提取方法的性能相当。该方法可以达到 100% 的检测精度, 用于确定 NASA 轴承数据集的轴承健康状态。该方法的简单设计有望用于基于物联网 (IoT) 的预

测解决方案的简单硬件实现。

- 用于指导训练和评估数据集注释的注释指南可对机器学习模型的质量产生相当大的影响。在本研究中，我们探讨了注释指南对 app 特征提取模型质量的影响。作为主要结果，我们建议对现有注释指南进行一些更改，目标是使提取的应用程序功能对应用程序开发人员更有用和更丰富。我们通过模拟新注释指南的应用来测试所提出的更改，然后评估在使用初始和模拟指南注释的数据集上训练的受监督机器学习模型的性能。虽然与使用初始注释的数据集上训练的模型相比，自动应用程序特征提取的整体性能保持不变，但是使用模拟新注释在数据集上训练的模型提取的特征对于应用程序开发人员来说噪声较小且信息量较大。其次，我们感兴趣的是培训自动应用程序特征提取模型所需的注释训练数据类型。特别是，我们将探讨培训集是否应包含随后计划应用该模型的应用/应用类别的带注释的应用评论，或者是否足以从任何可用于培训的应用进行带注释的应用评论，即使这些与测试应用相比，应用来自非常不同的类别。我们的实验表明，虽然将它们纳入训练集有助于提高召回率，但不需要从测试应用程序中进行带注释的训练评论。此外，我们测试是否通过带注释的产品评论来扩充训练集有助于提高应用特征提取的性能。我们发现，在增强训练集上训练的模型可以提高召回率，但代价是精度下降。
- 问答（QA）研究是自然语言处理中一项重大且具有挑战性的任务。质量检查旨在从相关的文本片段或文档中提取确切的答案。

QA 研究背后的动机是使用最先进的搜索引擎的用户的需求。用户期望得到确切答案，而不是可能包含答案的文档列表。在本文中，为了从相关文档中成功地提取答案，需要提取几个有效的特征和关系。这些功能包括各种词法，句法，语义和结构特征。提出的结构特征是从问题和支持文档的依赖特征中提取的。实验结果表明，结合基本特征并采用依赖原则设计，结构特征提高了答案提取的准确性。提出的结构特征使用提取远距离关系的新设计原则。这种添加是提高整体答案提取精度的可能原因。

- 光学相干断层扫描允许眼科医生获得眼睛视网膜的横截面成像。辅以数字图像分析方法，可以进行有效的疾病检测。存在各种方法从 OCT 图像中提取特征。拟议的研究旨在比较手工和深度神经网络特征的有效性。评估的数据集由分布在四个类中的 32339 个实例组成，即 CNV，DME，DRUSEN 和 NORMAL。方法是定向梯度直方图（HOG），局部二值模式（LBP），DenseNet-169 和 ResNet50。因此，基于深度神经网络的方法优于手工制作的特征，DenseNet 和 ResNet 的准确率分别为 88% 和 89%，而 HOG 和 LBP 分别为 50% 和 42%。基于深度神经网络的方法也在代表性不足的类中表现出更好的结果。
- 本文提出了一种使用卷积神经网络（CNN）的城市情景摩托车分类系统。使用 CNN 已经实现了关于图像分类的显著结果，代价是用于数千甚至数百万个示例的训练的高计算成本。然而，可以从已经训练的 CNN 中提取特征。在这项工作中，包含在

CaffeNet 框架中的 AlexNet 用于从真实城市场景中的帧中提取特征。来自 CNN 的提取的特征用于训练支持向量机 (SVM) 分类器以将摩托车与其他道路使用者区分开。得到的结果分别对三类和五类分类任务的平均准确率分别为 99.40% 和 99.29%。对显示令人满意的分类的验证图像集进行进一步的实验

- 在本文中，我们从高空间分辨率遥感图像提出分布式特征提取工具。工具基于 Apache Hadoop 框架和 Hadoop 图像处理接口。考虑两个角点检测 (Harris 和 Shi-Tomasi) 算法和五个特征描述符 (SIFT, SURF, FAST, BRIEF 和 ORB)。在 LandSat-8 图像的特征提取任务中工具的稳健性在水平可伸缩性方面进行评估。
- 为低资源语言开发实用的语音识别器是具有挑战性的，不仅因为语言的（可能是未知的）属性，而且因为测试数据可能不是来自与可用训练数据相同的域。在本文中，我们关注使用基于序列的标准训练的系统的后一挑战，即域错配。我们展示了使用预训练的英语识别器的有效性，该识别器对于这种不匹配的条件是鲁棒的，作为低资源语言上的域规范化特征提取器。在我们的示例中，我们使用土耳其语会话语音和广播新闻数据。这使得能够快速开发用于新语言的语音识别器，其可以容易地适应任何领域。在各种跨域方案中进行测试，我们在音素错误率方面实现了约 25% 的相对改进，对于某些域，改进率约为 50%。
- 本文介绍了一种新的方法，将人脑连接组学与分割相结合，用于



脑肿瘤分割和生存预测。对于分割，我们利用 MNI152 1mm 空间中的现有大脑分割图谱并将该分割映射到每个单独的主题数据。我们使用深度神经网络架构和硬负挖掘来实现最终的体素级别分类。对于生存预测，我们提出了一种结合连通组学数据，脑分割信息和脑肿瘤掩模的特征的新方法。我们利用 Human Connectome Project 中的平均连接组信息，将每个主题脑容量映射到这个常见的连接空间。由此，我们计算描述由于脑肿瘤导致的潜在神经中断的细胞学特征。然后使用这些特征来预测受试者的总体存活。所提出的方法的主要新颖性是使用来自人类连接组项目的标准化大脑分割数据和纤维束成像数据来分析 MR 图像以进行分割和存活预测。在 BraTS2018 数据上报告了实验结果。

- 在本报告中，我们提出了皮肤病变图像的分类技术，作为我们提交的 ISIC 2018 皮肤病变分析中针对黑色素瘤检测的挑战的一部分。我们的数据来自 ISIC 2018：皮肤病变分析，针对黑色素瘤检测大挑战数据集。这些特征是通过卷积神经网络提取的，在我们的案例中是 ResNet50，然后使用这些特征我们训练一个具有级联层的 DeepForest 来对我们的皮肤病变图像进行分类。我们知道卷积神经网络是图像表示学习中最先进的技术，卷积滤波器学习通过反向传播从图像中检测特征。然后，这些特征通常被馈送到分类器，例如 softmax 层或用于分类任务的其他这样的分类器。在我们的例子中，我们不使用传统的反向传播方法并训练 softmax 层进行分类。相反，我们使用 Deep Forest，一种新颖

的决策树集合方法，其性能与广泛的任务中的深度神经网络具有高度竞争性。因此，我们使用 ResNet50 从皮肤病变图像中提取特征，然后使用深森林对这些图像进行分类。使用这种方法是因为深森林在只有小规模培训数据的地区被发现非常有效。此外，由于深林网络自身决定其复杂性，它也迎合了我们在这个问题中遇到的数据集不平衡问题。

- 作为无监督降维方法，主成分分析（PCA）已被广泛认为是高光谱图像（HSI）处理和分析任务的有效且有效的预处理步骤。它将每个频段作为一个整体，并全局提取最具代表性的频段。然而，不同的均匀区域对应于不同的物体，其光谱特征是多样的。通过对整个恒生指数的统一预测来降低维数显然是不合适的。在本文中，提出了一种简单但非常有效的超像素 PCA 方法，称为 SuperPCA，用于学习 HSI 的内在低维特征。与传统的 PCA 模型相比，SuperPCA 有四个主要特性。（1）与基于整个图像的传统 PCA 方法不同，SuperPCA 考虑了不同均匀区域的多样性，即不同区域应具有不同的投影。（2）大多数传统特征提取模型不能直接使用 HSI 的空间信息，而 SuperPCA 能够通过超像素分割将空间背景信息合并到无监督维数减少中。（3）由于通过超像素分割获得的区域具有均匀性，因此即使在噪声下，SuperPCA 也可以提取潜在的低维特征。（4）尽管 SuperPCA 是一种无监督的方法，但与监督方法相比，它可以实现竞争性能。由此产生的特征具有辨别力，紧凑性和抗噪性，从而提高了 HSI 分类性能。



对三个公共数据集的实验表明，SuperPCA 模型明显优于传统的基于 PCA 的 HSI 分类维数降低基线。 Matlab 源代码可从 <https://github.com/junjun-jiang/SuperPCA> 获得

- 生成字符级特征是在各种自然语言处理任务中实现良好结果的重要步骤。为了减少产生手工制作特征的人工需求，已经提出了利用诸如卷积神经网络（CNN）或递归神经网络（RNN）之类的神经架构来自动提取这些特征的方法，并且已经显示出很好的结果。然而，CNN 生成与位置无关的特征，并且 RNN 很慢，因为它需要按顺序处理字符。在本文中，我们提出了一种使用密集连接网络自动提取字符级特征的新方法。所提出的方法不需要任何语言或任务特定的假设，并且显示出稳健性和有效性，同时比基于 CNN 或 RNN 的方法更快。在三个序列标记任务上评估此方法 – 时隙标记，词性（POS）标记和命名实体识别（NER） – 我们获得了最先进的性能，96.62 的 F1 得分和 97.73% 分别对插槽标记和 POS 标记的准确性，以及与 NER 上最先进的 91.13 F1 得分相当的性能
- 本文提出了一种利用线性偏微分方程（PDE）提取二值图像数据中几何形状特征的统一方法。 PDE 和功能的配方可以同时提取几何形状特征，包括厚度，形状方向和骨架。所提出的方法的主要优点是它没有关于距离的任何计算，它没有目标图像数据的拓扑约束，并且表面不必被区分到内部或外部。提供一维分析解决方案以验证所提出的方法。另外，显示了二维和三维数值示例以

证实所提出方法的有效性和有用性。

- 近年来，许多出版物表明基于卷积神经网络的特征可以具有优于工程特征的性能。然而，到目前为止，没有花费太多精力来为整个图像有效地提取局部特征。在本文中，我们提出了一种方法来有效地计算基于补丁的局部特征描述符，同时存在整个图像的池和跨越层。我们的方法是通用的，可以应用于几乎所有现有的网络架构。这包括用于所有本地特征提取任务的网络，例如相机校准，Patchmatching，光流估计和立体匹配。此外，我们的方法可以应用于其他基于补丁的方法，如滑动窗口对象检测和识别。我们使用基于 CNN 的流行特征提取方法的速度基准来完成我们的论文，该方法应用于整个图像，有或没有我们的加速，以及示例代码（对于 Torch），它显示了如何通过我们的方法轻松转换任意 CNN 架构。
- 许多品种的神经网络被吹捧为非常强大的机器学习工具，因为它们能够从不同形式的数据中提取大量信息，提取复杂的特征并实现强大的分类能力。在这项研究中，我们使用神经网络从图像和数字数据中提取特征，并使用这些提取的特征作为其他机器学习模型的输入，即支持向量机（SVM）和 k-最近邻分类器（KNN），以便看看神经网络提取的功能是否增强了这些模型的功能。我们以这种方式测试了 7 种不同的神经网络架构，4 种用于图像，3 种用于数值数据，每次训练各种不同的时间长度，然后将神经网络的结果独立地与数据上的 SVM 和 KNN 的结果进行

比较，最后将这些结果与使用通过神经网络架构提取的特征训练的 SVM 和 KNN 的模型进行比较。在 3 个不同的图像数据集和 2 个不同的数字数据集上重复该过程。结果表明，在许多情况下，与在原始特征上运行这些算法相比，使用神经网络提取的特征显著提高了 SVM 和 KNN 的能力，并且在某些情况下还单独超过神经网络的性能。这反过来表明，使用神经网络作为一种手段来提取某些数据集的其他机器学习模型进行分类的方法可能是合理的做法。

- 计算机辅助诊断（CAD）系统旨在帮助临床医生完成各种任务，包括突出医学图像中的异常区域。常见的方法包括在从患者扫描中的正常和病理区域提取的一组特征向量上训练体素级二元分类器。然而，许多病理（例如癫痫）的特征在于可能位于大脑中任何位置的病变，具有各种形状，大小和纹理。这种异质性的充分表示需要大量的注释数据，这是医学领域中的主要问题。因此，我们建立在先前提出的将癫痫病变检测任务视为体素水平异常值检测问题的方法的基础上。它包括使用 El Azami 等人，2016 年的少量临床指导特征为大脑体积中的每个体素构建 oc-SVM 分类器。我们在本研究中的目标是通过替换手工制作的特征向前迈出一步。使用神经网络自动学习表示。我们提出了一种新型的暹罗网络，这些网络仅通过从健康患者的扫描中提取的补丁进行训练。该网络由作为子网的堆叠自动编码器组成，由补丁的重建误差规范化。它被设计用于学习表示以相对于所选度量（即余弦）“更接近”的

相同体素定位为中心的补丁。最后，子网的中间层表示被馈送到体素级的 oc-SVM 分类器。该方法在 3 例患者的 MRI 扫描中得到验证，确认癫痫病变并显示出有希望的表现。

- 该方法引入了学习动作类别的有效方式，而无需特征估计。该方法从低级值开始，与成功的 CNN 方法类似。但是，我们学习从原始视频数据中预测特定的视频表示，而不是提取一般图像特征。这种方法的好处是，在相同的计算费用下，它可以基于运动预测 2D 视频表示以及 3D 视频表示。所提出的模型依赖于判别式 Waldboost，我们为了学习视频表示而增强了多类式。在 UCF11 动作识别数据集上测试所提出的方法的适用性以及其时间效率。
- 分层时间记忆（HTM）是一种神经形态算法，模拟类似于新皮层工作原理的稀疏性，层次性和模块性。特征编码是创建稀疏二进制模式的重要步骤。这种稀疏性是由 HTM 的初始化阶段中的二进制权重和随机权重分配引入的。我们提出了 HTM 初始化阶段的替代确定性方法，该方法将 HTM 权重连接到输入数据并保留输入信息的自然稀疏性。此外，我们介绍了确定性方法的硬件实现，并将其与传统的 HTM 和现有的硬件实现进行了比较。我们在人脸识别问题上测试了所提出的方法，并表明它优于传统的 HTM 方法。
- 纹理分析和分类是图像处理科学家自 80 年代末以来一直备受关注的一些问题。如果纹理分析准确完成，它可以在很多情况下使用，例如对象跟踪，视觉模式识别和人脸识别。从现在开始，提

供了很多方法来解决这个问题。针对他们的技术差异，他们都使用相同的流行数据库来评估他们的性能，例如 Brodatz 或 Outex，这可能会使他们的性能偏向于这些数据库。在本文中，提出了一种方法来收集更有效的纹理图像数据库。拟议的方法包括两个阶段。第一个是基于灰度色差矩阵和局部二值模式特征开发特征表示，下一个是基于 K 均值聚类的创新算法，基于评估的特征收集图像。为了评估所提出方法的性能，收集纹理数据库并计算收集的一个和众所周知的数据库的费率。此外，基于提供的特征提取来评估纹理分类，并且通过一些现有技术的纹理分类方法来比较准确度。

- 先前看不见的条件会严重影响自动语音识别（ASR）系统的性能，这通常是由于训练和测试分布之间的不匹配。在本文中，我们通过研究域不变特征来解决鲁棒性问题，使域信息对 ASR 系统变得透明，解决了不匹配问题。具体来说，我们研究了一个名为分解分层变分自动编码器（FHVAE）的最新模型。FHVAE 学习如何在没有监督的情况下将序列级和段级属性分解为不同的潜在变量。我们认为包含段级信息的潜在变量集是我们期望的 ASR 域不变特征。在 Aurora-4 和 CHiME-4 上进行实验，分别在错配结构域上显示出 41% 和 27% 的绝对字错误率降低。
- 本文提出了一种检测时间序列预测中非平稳性的实用方法。该方法称为 SAFE，通过距离函数监视时间序列的频谱内容的演变。该方法旨在通过告知在线预测器在非平稳性存在时执行必要的适

应，与现有技术的机器学习方法结合实时工作。我们还提出了一种算法，在比较过程中按比例包含一些过去的的数据，以克服灾难性遗忘问题。为了验证我们的假设并测试我们方法的有效性，我们在涉及人工和现实数据集的方法的不同元素中提供了全面的实验。实验表明，所提出的方法能够在保持高预测性能的同时，在处理器或 GPU 周期方面显着节省计算资源。

- 本文提出了一种适用于心电图（ECG）信号的特征提取算法（Pan Tompkins 算法）的适当和有效的实现，用于四种心脏疾病的检测和分类：睡眠呼吸暂停，心律失常，室上性心律失常和长期心房颤动（AF）和通过使用 pan Tompkins RR 检测，然后进行特征提取以进行分类，将它们与正常心跳区分开来。本文还提出了一种利用现有神经网络分类器进行信号分类的新方法。
- 随着数据的高维度，特征提取变得越来越重要。Autoencoder 作为一种基于神经网络的特征提取方法，在生成高维数据的抽象特征方面取得了巨大成功。但是，它没有考虑数据样本的关系，这可能会影响使用原始和新功能的实验结果。在本文中，我们提出了一个考虑数据特征及其关系的关系自动编码器模型。我们还将其扩展到与其他主要自动编码器模型一起使用，包括稀疏自动编码器，去噪自动编码器和变分自动编码器。所提出的关系自动编码器模型在一组基准数据集上进行评估，实验结果表明，与自动编码器的其他变体相比，考虑数据关系可以生成更强大的特征，从而实现更低的构造损失，然后在进一步分类中降低错误率。



- 文本挖掘是一个旨在从文本数据中提取信息的字段。这种研究领域的挑战之一来自预处理阶段，其中应从非结构化数据中提取向量（和结构化）表示。公共提取创建大且稀疏的向量，表示每个术语对文档的重要性。因此，这通常会导致困扰大多数机器学习算法的维数诅咒。为了解决这个问题，在本文中，我们提出了一种新的监督特征提取和约简算法，名为 DCDistance，它根据文档与每个类标签代表之间的距离创建特征。因此，所提出的技术可以减少超过原始集合的 99% 的特征。此外，与传统和最先进的特征选择算法相比，该算法还能够提高一组基准数据集的分类准确度。
- 在用于关键短语提取的监督方法中，候选短语用一组手工制作的特征编码，并且训练机器学习算法以区分关键短语和非关键短语。虽然手动设计的功能在实践中表现良好，但特征工程是一个困难的过程，需要专业知识，通常不能很好地概括。在本文中，我们介绍了 SurfKE，这是一个功能学习框架，它利用文本本身自动发现关键短语所展示的模式。我们的模型将文档表示为图形，并自动学习短语的特征表示。与强基线相比，该模型在性能方面取得了显著进步。
- 特征提取和特征选择是输入日志预处理的首要任务，以便在利用机器学习的同时检测网络安全威胁和攻击。当涉及从不同来源获得的异构数据的分析时，发现这些任务是耗时的并且难以有效地管理。在本文中，我们提出了一种处理特征提取和特征选择的方法。

法，用于对来自不同网络传感器的异构数据进行安全分析。该方法在 Apache Spark 中实现，使用其名为 pyspark 的 python API。

- EMFET 是一种开源且灵活的工具，可用于从任何电子邮件语料库中提取大量功能，并以 EML 格式保存电子邮件。提取的特征可以分为三个主要组：标题特征，有效负载（主体）特征和附件特征。该工具的目的是帮助从业者和研究人员构建可用于训练机器学习模型以进行垃圾邮件检测的数据集。到目前为止，可以使用 EMFET 提取 140 个特征。EMFET 是可扩展且易于使用的。EMFET 的源代码可在 GitHub 上公开获取（<https://github.com/WadeaHijawi/EmailFeaturesExtractio>
- 在本文中，我们提出了一种基于人工神经网络（ANN）的高光谱图像光谱空间特征提取和分类框架。对于有限的标记样本，仅利用光谱信息进行训练，并且在测试阶段将空间背景集成在后面。利用人脸识别的最新进展，采用结合 softmax 损失和中心损失的联合监督符号来训练所提出的网络，通过该网络收集类内特征，同时扩大类间变化。基于所学习的架构，提取的基于频谱的特征由中心分类器分类。此外，为了融合光谱和空间信息，开发了自适应光谱 - 空间中心分类器，其中同时考虑多尺度邻域，并且使用自适应投票策略确定最终标签。最后，三个著名数据集的实验结果验证了所提方法与最先进方法相比的有效性。
- 在本文中，我们通过采用多级特征分析提出了一种实时的对象识别方法，并展示了将特征提取适用于朴素贝叶斯分类框架的实用

性，该框架实现了简单，高效和稳健的性能。我们还表明，随着级别类数量的增加，所提出的方法也会很好地扩展。为了有效地理解关键点周围的补丁，训练的分类器使用数百个简单的二进制特征和模型类后验概率。此外，在任意特征集之间假定的独立性下，分类过程在计算上是便宜的。即使对于某些特定情况，此假设也可能无效。我们证明了有效分类器在图像数据集上的表现非常好，在照明环境和图像捕获视角方面有很大的变化。实验结果表明，在许多具有挑战性的数据集上可以实现一致的精度，同时为大分辨率图像提供交互速度。该方法展示了有希望的结果，其优于最先进的模式识别方法。

- 同时定位和映射（SLAM）是构建或更新未知环境的映射同时跟踪代理在其中的位置的问题。如何在移动设备甚至物联网等级设备上实现强大而持久的 SLAM，是当今业界面临的主要挑战。我们需要解决的主要问题是：1) 如何加速 SLAM 流水线以满足实时要求; 2.) 如何降低 SLAM 能耗以延长电池寿命。在深入研究问题后，我们发现特征提取确实是性能和能耗的瓶颈。因此，在本文中，我们设计，实现和评估硬件 ORB 特征提取器，并证明我们的设计在性能和能耗之间取得了很好的平衡，与 ARM Krait 和 Intel Core i5 相比。
- 本文旨在开发一种新的强大的特征表示方法。在自动编码器成功的推动下，我们首先理论总结了基于传统自动编码器的所有算法的一般属性：1) 输入的重建误差不能低于下限，可以看作是重

建输入的指导原则。另外，当输入被噪声破坏时，损坏的输入的重建误差也不能低于下限。 2) 实现其理想状态的隐藏表示的重建是重构输入以达到理想状态的必要条件。 3) 最小化隐藏表示的雅可比矩阵的弗罗贝尼乌斯范数具有不足并且可能导致更差的局部最优值。我们认为，最小化隐藏表示的重建误差比最小化隐藏表示的雅可比矩阵的 Frobenius 范数更稳健。基于上述分析，我们提出了一种称为双重去噪自动编码器（DDAEs）的新模型，该模型在输入和隐藏表示上使用损坏和重建。我们证明了所提出的模型具有高度的灵活性和可扩展性，并且具有更好的学习不变性和强大特征表示的能力。我们还表明，我们的模型比去噪自动编码器（DAE）更强大，可以处理噪音或不必要的功能。此外，我们分别通过以组合和分离的方式优化目标函数，详细说明如何使用两种不同的预训练方法训练 DDAE。对比实验表明，与最先进的模型相比，所提出的模型对于表示学习明显更好。

- 我们使用自动语音识别来评估英语口语学习者的发音，这是基于从支持向量机（SVM）分类器确定的学习者口语反应的真实可懂度或转录正确性的深度学习神经网络模型预测。使用 PocketSphinx 对齐模式生成的数字特征和许多识别通道搜索每个预期音素的替换和删除以及按顺序插入意外音素，SVM 模型与亚马逊机械土耳其群岛工作人员转录的准确率达到 82%，高于 75 多名独立研究人员报告的百分比将这些特征与 SVM 分类器概率预测模型一起使用可以帮助计算机辅助发音教学（CAPT）

系统提供可懂度修复。

- 传统的认证技术，如 RFID 标签和认证卡/徽章，存在不同的弱点，因此应该应用快速替代使用生物认证方法的认证。生物识别技术，例如指纹，语音和 ECG 信号，是可用于认证处理的独特人类角色。在这项工作中，我们提出了一个基于使用提取的 ECG 功能识别未知人员的物联网实时认证系统。离散余弦变换（DCT）用作 ECG 特征提取，其具有用于实时系统实现的更好特性。大量的研究具有高度的认证准确性，但是大多数研究忽略了认证个体的实时能力。在处理时间约 1.21 秒的准确率为 97.78% 的情况下，所提出的系统更适用于需要快速可靠的认证处理要求的许多应用中。
- 最近的声学事件分类研究集中在训练合适的滤波器以表示声学事件。但是，由于目标事件数据库的可用性有限以及传统滤波器的线性，因此仍有提高性能的空间。通过利用深度神经网络（DNN）的非线性建模及其在预训练环境之外学习的能力，本书提出了一种基于 DNN 的声学事件分类特征提取方案。使用室内监视环境数据库证明了所提方法的有效性和抗噪声性。
- 本文提出了一种新的集合框架，用于提取图像的高度辨别特征表示及其在野外群体级幸福感强度预测中的应用。为了产生足够多的决策，通过引导训练集并从中提取每个图像的  $n$  个特征来训练  $n$  个卷积神经网络。然后使用递归神经网络（RNN）来记住哪个网络提取更好的特征并为一个单独的图像生成最终的特征表示。

几组情绪模型（GEM）用于聚合一组中的面部特征，并使用参数优化的支持向量回归器（SVR）来获得最终结果。通过大量实验，所提出的复发随机深度集合（RRDE）的有效性在结构和决策方面得到了证明。最佳结果在 HAPPEI 数据集的验证集上产生 0.55 均方根误差（RMSE），明显优于基线 0.78。

- 许多现实世界的系统需要在异构信息网络上运行，该网络由许多不同类型的交互组件组成。例子包括对生物信息网络进行数据分析的系统;社交网络;和信息提取系统处理非结构化数据以将原始文本转换为知识图。许多以前的工作描述了在这种网络上执行特定类型的分析，挖掘和学习的专门方法。在这项工作中，我们提出了一个统一的框架，包括一个数据模型 – 一个带有一阶模式的图形以及一个声明性语言，用于以促进关系和结构化机器学习的方式构建，查询和操纵这些网络。特别是，我们为关系和图遍历查询语言提供了一个初始原型，其中查询直接用作结构化机器学习模型的关系特征。通过进行声明性图遍历查询来执行特征提取。学习和推理模型可以直接对这种关系表示进行操作，并使用新的数据和知识对其进行扩充，然后将新的数据和知识无缝集成到关系结构中以支持新的预测。我们通过展示自然语言处理和计算生物学领域的任务来展示该系统的功能。
- 为了识别他们的偏好然后在这些模型的基础上个性化服务而对用户进行建模是一项复杂的任务，主要是因为需要考虑各种显式和隐式信号，缺失或不确定的信息，上下文方面等等。在这项研究



中，提出并评估了一种用于从用户数据中发现潜在偏好模式的新颖通用方法。该方法依赖于使用图表表示数据，然后系统地提取基于图形的特征，并使用它们来丰富原始用户模型。提取的功能封装了用户，项目和元数据之间的复杂关系。然后，增强的用户模型可以用作任何推荐算法的输入。所提出的方法是与领域无关的（在电影，音乐和商业推荐系统的数据上证明），并且使用几种最先进的机器学习方法，在不同的推荐任务上以及使用不同的评估度量来评估。结果显示了跨任务和域的推荐准确性的一致改进。此外，评估还提供了有关特殊情景下方法绩效的更深入分析，包括高稀疏性和评级的可变性。

- 细粒度分类可以从基于部件的特征中受益，这些特征揭示了对象类别之间的细微视觉差异。手工制作的功能已被广泛用于零件检测和分类。尽管最近的趋势试图使用诸如卷积神经网络（CNN）之类的强大的深度学习模型自动地学习这些特征，但是他们的训练以及可能还需要测试需要手动提供的注释，这些注释成本很高。为了放宽这些要求，我们在本研究中假设一般问题设置，其中原始图像仅提供用于模型训练的对象级类标签，而不需要其他辅助信息。具体地，通过提取和解释由 CNN 学习的分层隐藏层特征，我们提出了用于细粒度分类的精细的基于 CNN 的系统。当在 Caltech-UCSD Birds-200-2011, FGVC-Aircraft, Cars 和 Stanford 狗数据集上进行评估时，只有对象级别的标签用于培训，而且没有其他注释可用于培训和测试，我们的方法实现了与现有

技术相比具有优越性或可比性的令人印象深刻的性能。此外，它揭示了 CNN 学习的等级特征的巧妙运用，它具有广泛的适用性，远远超出目前的细粒度分类任务。

- 与传统的 SIFT 和 SURF 相比，最近提出的开源 KAZE 图像特征检测和描述算法提供了前所未有的性能，因为它依赖于非线性尺度空间而不是高斯线性尺度空间。然而，改进的性能带来了显著的计算成本，限制了其在许多应用中的使用。我们报告了 KAZE 算法的 GPGPU 实现，而没有使用二进制描述符来获得加速。对于  $1920 \times 1200$  大小的图像，我们的基于计算统一设备架构 (CUDA) C 的 GPU 版本在 NVIDIA GeForce GTX Titan X (Maxwell Architecture–GM200) 卡上花费了大约 300 毫秒，而多线程 CPU 版本则为近 2400 毫秒 (16 线程 Intel (R) Xeon (R) CPU E5-2650 处理器)。通过 GPU 和 CPU 实现之间的细粒度比较，详细描述了基于 CUDA 的并行实现。通过实现近 8 倍的加速而不会降低性能，我们的工作扩展了 KAZE 算法的适用性。另外，这里描述的策略可以证明对于其他基于非线性尺度空间的方法的 GPU 实现是有用的。
- 尖峰神经网络 (SNN) 由于其稀疏的基于尖峰的编码方案而实现了功率高效的实现。本文开发了一种生物启发的 SNN，它使用无监督学习从语音信号中提取判别特征，随后可用于分类器。该体系结构包括尖峰卷积/池化层，后面是完全连接的尖峰层，用于特征发现。泄漏，积分和激发 (LIF) 神经元的卷积层代表主要

声学特征。完全连接的层配备有概率性尖峰定时依赖的可塑性学习规则。该层通过概率 LIF 神经元表示辨别特征。为了评估学习特征的辨别力，它们被用在隐藏的马尔可夫模型（HMM）中用于口头数字识别。实验结果表明，性能高于 96%，与流行的统计特征提取方法相比具有优势。我们的结果提供了 SNN 中无监督特征获取的新颖演示。

- 近年来，许多用于记录和观察听诊器声音的创新解决方案已经可用。然而，为了充分利用这种装置，需要一种用于检测异常肺部声音的自动方法，这比通常使用小的和非多样化的数据集开发和评估的现有方法更好。我们提出了一种基于机器学习的方法，用于在大型健康调查中使用听诊器检测肺部声音中的噼啪声。我们的方法经过训练和评估，使用 209 个文件，由专业听众分类。我们的分析管道基于从音频文件中的小窗口中提取的功能。我们评估了几种特征提取方法和分类器。我们使用 175 个裂纹窗口和 208 个普通窗口的训练集来评估管道。我们进行了 100 次交叉验证循环，我们在循环之间改组了训练集。对于所有培训和评估之间的划分是 70%–30%。我们发现并评估了具有来自时域的四个特征和来自谱域的一个 5 维矢量。我们评估了几个分类器，发现带有径向基函数核的 SVM 表现最佳。我们的方法具有 86% 的精确度，并且在窗口中对噼啪声进行分类时回忆率为 84%，这比在卫生人员的研究中发现的更准确。低维特征向量使 SVM 非常快。该模型可以在 1.44 秒内在普通计算机上进行训练，并且可以

在 1.08 秒内对 319 个爆裂进行分类。我们的方法可以检测并显示录制的音频文件中的各个噼啪声它准确，快速，资源要求低。它可用于培训医务人员或作为蓝牙听诊器的智能手机应用程序的一部分。

- 实时可视化分析任务（如跟踪和识别）需要快速执行计算密集型算法。视觉传感器网络可以通过用处理节点增加传感器网络并以摄像机争用处理节点同时试图最小化其任务完成时间的方式分配计算负担来执行这样的任务。在本文中，我们制定了将所有相机传感器的完成时间最小化作为优化问题的问题。我们提出了完全分布式优化的算法，分析均衡分配的存在，评估网络拓扑和视频特性的影响，以及中央协调的好处。我们的结果表明，只要有足够的信息，分布式优化就可以提供较低的完成时间，而且可以通过额外的稀疏中心协调来实现可预测和稳定的性能。
- 组织病理学评估，包括手术切除和核心针穿刺活检，是前列腺癌诊断的标准程序。目前对组织病理学图像的解释包括肿瘤面积的确定，Gleason 分级以及某些预后关键特征的鉴定。这样的过程不仅乏味，而且修剪内部/相互观察的变化。最近，FDA 批准了第一个用于数字病理学的整个载玻片成像系统的营销。这开启了基于数字组织病理学图像的计算机辅助前列腺图像分析和特征提取的新时代。在这项工作中，我们提出了一个分析管道，包括癌症区域的定位，分级，不同格里森等级的面积比，以及细胞学/建筑特征提取。该算法结合了人工工程特征提取以及深度神经网络

络学习的特征提取。此外，实现整个管道以直接操作由数字扫描仪产生的整个幻灯片图像，因此可能很容易转化为临床实践。该算法在来自 TCGA 数据集的 368 个整个幻灯片图像上进行测试，并且在用 4 + 3 幻灯片区分格里森 3 + 4 时实现 75% 的总体准确度。

- 本文旨在通过神经网络体系结构促进有关文本特征提取技术的讨论。本文讨论的研究问题集中在最先进的神经网络技术上，这些技术已被证明是语言处理，语言生成，文本分类和其他计算语言学任务的有用工具。
- 这项工作提出了一项研究，从三种类型的眼球运动事件中提取和分析一组 101 类眼动特征：固定，扫视和后扫视振荡。在阅读任务期间记录眼睛运动。对于记录中具有多个实例的要素类别，我们通过计算这些实例的分布的描述性统计来提取相应的要素子类型。提供了详细描述和数学公式的统一框架，用于提取特征集。使用来自 298 名受试者的规范人群的大型眼动记录数据库来执行特征值的分析。我们展示了实验人群中特征值的集中趋势和整体变异性，更重要的是，我们量化了每个单独特征的重测信度（可重复性）。所描述的方法和分析可以在探索眼睛运动的领域中提供有价值的工具，例如行为研究，注意力和认知研究，医学研究，生物识别识别和人机交互。
- 与仅使用运动测量的标准多目标跟踪（MTT）算法相比，特征辅助跟踪通常可以产生改进的跟踪性能。然而，在许多应用中，目

标的特征信号由稀疏傅立叶域信号组成。它在时域中快速且非线性地变化，并且特征测量被错过的检测和错误关联破坏。这两个因素使得难以提取要在 MTT 中使用的特征信息。在本文中，我们开发了一个特征辅助最近邻联合概率数据关联滤波器

(NN-JPDAF)，用于在密集目标环境中进行联合 MTT 和特征提取。为了估计来自不完全和损坏的测量的快速变化的特征信号，我们使用原子范数约束来表示特征信号的稀疏性，并使用  $\ell_1$  范数来表示由错误关联引起的损坏的稀疏性。基于稀疏表示，通过求解凸的半定规划 (SDP) 来估计特征信号。我们还提供了一种迭代方法，通过乘法器的交替方向方法 (ADMM) 来解决这个 SDP，其中每次迭代都涉及闭式计算。利用估计的特征信号，执行重新滤波以估计目标的运动状态，其中该关联利用运动学和特征信息。仿真结果表明了该算法在雷达应用中的性能。

- 计算机视觉是当今信息技术中最活跃的研究领域之一。让机器和机器人能够以视线的速度观察和理解周围的世界，创造无限的潜在应用和机会。特征检测和描述算法确实可以被认为是这种机器和机器人眼睛的视网膜。然而，这些算法通常是计算密集型的，这使得它们无法实现视觉实时性能的速度。此外，它们的能力不同，有些人可能会因为与其他人相比特定类型的投入而更青睐和更好地工作。因此，必须紧凑地报告其利弊，以及他们的表现和最近的进展。本文致力于全面概述特征检测和描述算法的最新进展和最新进展。具体而言，它首先概述了基本概念。然后，它会



比较，报告和讨论它们的性能和功能。选择最大稳定极值区域算法和尺度不变特征变换算法，它们是两种类型中最好的算法，以报告它们最近的算法导数。

- 深度神经网络是表示学习技术。在训练期间，深网能够在机器学习中生成前所未有的大小和细节的描述性语言。提取在训练有素的 CNN 模型中编码的描述性语言（在图像数据的情况下），并将其重新用于其他目的是感兴趣的领域，因为它提供对在处理数百万图像之后 CNN 先前学习的视觉描述符的访问，无需昂贵的培训阶段。到目前为止，对该领域的贡献（通常称为特征表示转移或转移学习）纯粹是经验性的，从靠近输出的单个层提取所有 CNN 特征，并通过将它们馈送到分类器来测试它们的性能。这种方法提供了一致的结果，尽管其相关性仅限于分类任务。在一种完全不同的方法中，在本文中，我们统计地测量深 CNN 中发现的每个特征的判别能力，当用于表征每个类别的 11 个数据集时。我们试图提供有关 CNN 特征行为的新见解，特别是来自卷积层的行为，因为这可能与它们在知识表示和推理中的应用相关。我们的结果证实，低级和中级特征可能与高级特征的行为不同，但仅在某些条件下。我们发现所有 CNN 特征都可以用于知识表示目的，无论是通过它们的存在还是通过它们的缺席，将单个 CNN 特征可以提供的信息加倍。我们还研究了这些特征可能包含多少噪声，并提出了一种阈值方法来丢弃其中的大部分。所有这些见解都直接应用于 CNN 嵌入空间的生成。

- 特征提取是许多应用数据科学工作流程的关键组成部分。近年来，人工智能和机器学习的快速发展导致了特征提取工具和服务的激增，使数据科学家能够以各种各样的维度廉价有效地注释他们的数据 – 从检测图像中的面部到分析连贯文本中表达的情绪。不幸的是，强大的特征提取服务的激增已经通过对特征提取服务的不同接口的数量的相应扩展来反映。在几乎每个新服务都有自己的API，文档和/或客户端库的世界中，需要组合从多个源获得的各种功能的数据科学家经常被迫编写和维护更精细的特征提取管道。为了应对这一挑战，我们引入了一个全新的多模式特征提取开源框架。Pliers 是一个开源 Python 包，支持各种数据类型（视频，图像，音频和文本）的标准化注释，并且明确地考虑了易用性和可扩展性。用户可以在几行 Python 代码中将各种预先存在的特征提取工具应用于他们的数据，并且还可以通过编写模块化类轻松地添加他们自己的自定义提取器。基于图形的 API 可以快速开发复杂的特征提取管道，从而以单一的标准化格式输出结果。我们描述了包的体系结构，详细介绍了它与以前的特征提取工具箱相比的主要优势，并将示例应用程序用于大型功能性 MRI 数据集，以说明钳子如何能够显著减少构建复杂特征提取工作流程所需的时间和精力，同时提高代码清晰度和可维护性。
- 本文提出了一种基于遗传算法的频域特征搜索（GAFDS）方法用于癫痫的脑电图（EEG）分析。在该方法中，首先搜索频域特征，然后将其与非线性特征组合。随后，选择并优化这些特征以

对 EEG 信号进行分类。通过实验分析提取的特征。 GAFDS 提取的特征显示出显著的独立性，并且它们在类间距离和类内距离的比率方面优于非线性特征。此外，所提出的特征搜索方法还可以在希尔伯特变换之后搜索信号中的瞬时频率的特征。使用这些特征获得的分类结果是合理的，因此，GAFDS 表现出良好的可扩展性。多个经典分类器（即 k-最近邻，线性判别分析，决策树，AdaBoost，多层感知器和朴素贝叶斯）通过使用 GAFDS 方法生成的特征和优化的选择来获得良好的结果。具体而言，两个分类和三个分类问题的准确度可分别达到 99%和 97%。几个交叉验证实验的结果表明，GAFDS 在脑电图分类的特征提取中是有效的。因此，所提出的特征选择和优化模型可以提高分类准确性

- 特征选择的所有相关问题是识别所有强相关和弱相关属性。对于诸如预测性维护或生产线优化的工业应用中的时间序列分类和回归，该问题尤其难以解决，其中每个标签或回归目标同时与若干时间序列和元信息相关联。在这里，我们提出了一种用于时间序列的高效，可扩展的特征提取算法，该算法在机器学习管道的早期阶段过滤可用特征，以及它们对分类或回归任务的重要性，同时控制所选择的预期百分比。不相关的功能。所提出的算法将已建立的特征提取方法与特征重要性过滤器相结合它具有较低的计算复杂度，允许仅利用有限的领域知识来开始问题，可以简单地并行化，具有高度可扩展性并且基于充分研究的非参数假设检验。我们将所提出的算法与 UCR 时间序列分类存档的所有二元分类

问题以及来自生产线优化项目的时间序列和具有潜在的动态定性变化的模拟随机过程进行对比。

- 在本文中，我们比较了基于事件的衰减和基于时间的衰减记忆表面，用于使用基于事件的摄像机进行高速事件跟踪，特征提取和对象分类。高速识别任务涉及检测和分类在相机镜头附近自由下落的模型飞机，以便产生表现出目标速度的显著变化的具有挑战性的数据集。与基于时间的衰减存储器表面相比，这种差异促使研究基于事件的衰减存储器表面以捕获基于事件的数据的时间方面。然后，这些表面用于执行无监督的特征提取，跟踪和识别。为了生成内存表面，事件分箱，线性衰减内核和指数衰减内核被研究，发现指数衰减的内核表现最佳。发现基于事件的衰减记忆表面在识别中优于基于时间的衰减记忆表面，特别是当需要对目标速度的不变性时。研究了一系列网络和感受野大小。该系统在飞机进入视野的 156 毫秒内实现 98.75% 的识别精度，仅使用 25 个基于事件的特征提取神经元与线性分类器串联。通过将线性分类器结果与 ELM 分类器进行比较，我们发现少数基于事件的特征提取器可以有效地将数据集的复杂时空事件模式投影到特征空间中几乎线性可分离的表示。
- 深度卷积神经网络已经在许多实际的机器学习任务中取得了突破性成果，例如 ImageNet 数据集中的图像分类，用于玩 Atari 游戏或棋盘游戏 Go 的控制策略学习以及图像字幕。这些应用中的许多应用首先执行特征提取，然后将其结果馈送到可训练的分类器

中。用于特征提取的深度卷积神经网络的数学分析由 Mallat, 2012 开始。具体而言, Mallat 考虑所谓的基于小波变换的散射网络, 随后是每个网络层中的模非线性, 并证明了平移不变性 (渐近地) 在小波尺度参数中) 和相应特征提取器的变形稳定性。本文通过开发一种包含一般卷积变换的理论, 或者更多技术用语, 一般半离散帧 (包括 Weyl–Heisenberg 滤波器, 曲线, 剪切, 脊波, 小波和学习滤波器), 一般 Lipschitz 连续, 补充了 Mallat 的结果。非线性 (例如, 整流线性单位, 移位逻辑 sigmoids, 双曲正切和模数函数), 以及模拟例如子采样和平均的一般 Lipschitz 连续汇集算子。此外, 所有这些元素在不同的网络层中可以是不同的。对于得到的特征提取器, 我们证明了垂直性质的平移不变性结果, 在特征意义上随着网络深度的增加逐渐变得更加平移, 并且我们建立了适用于信号类的变形灵敏度边界, 例如, 带限制函数, 卡通函数和 Lipschitz 函数。

- 通常我们在实践中遇到的数据是矩阵的集合而不是单个矩阵。这些多块数据是自然链接的, 因此通常具有一些共同的特征, 同时由于它们被测量和收集的背景, 它们具有各自的特征。在这项研究中, 我们提出了一种新的通用和个性化特征分析 (CIFA) 方案, 该方案以链接的方式处理多块数据, 旨在发现和分离它们的共同和个体特征。根据是否给出共同特征的数量, 提出了两种有效的算法来提取所有数据共享的共同基础。然后通过结合诸如降维和盲源分离的技术, 分别对公共空间和个体空间进行特征提取。我

们还讨论了提议的 CIFA 如何通过分别利用样本的共同和个体特征来显着提高分类和聚类任务的性能。我们的实验结果表明，与合成和实际数据的最新方法相比，所提出的方法具有一些令人鼓舞的特征。