|  |  |  |
| --- | --- | --- |
| **文章信息** | **摘要及贡献** | **方法概述及原理简述** |
| **阅读程度**：  仅浏览  **文章标题：**  Representation Learning with Contrastive Predictive Coding  **中文标题：**  使用对比预测编码进行表征学习  **发表于：**  Arxiv2019  **作者：**  Aaron van den Oord  **单位：**  Deepmind | 虽然有监督学习在许多应用中取得了巨大的进展，但无监督学习还没有得到如此广泛的应用，这仍然是人工智能的一项重要和具有挑战性的努力。  在这项工作中，我们提出了一种通用的无监督学习方法来从高维数据中提取有用的表示，我们称之为对比预测编码。  我们的模型的关键灵感是在潜在空间中通过使用强大的自回归模型预测未来，来学习某种表征。  我们使用一个概率对比损失，它诱导潜在空间来捕获信息，这对预测未来的样本是最大有用的。  它还通过使用负抽样使模型易于处理。  虽然之前的大多数工作都集中在特定方式上来评价学到的特征，但我们证明了我们的方法能够学习有用的表征，在四个不同的领域上取得了强大的表现：语音、图像、文本和3D环境中的强化学习。 | 基本编码器（enconder）使用卷积神经网络即可记为genc，将原始数据x映射到潜在空间中z  自回归编码器记为gar能够提取时序信息，将潜在空间中的一段时序序列编码生成  c，c即为通过gar获得的上下文信息  公式1：  公式2：    语音数据：以预测zt+3为例，W3为一组权重，乘以ct后即为模型的预测值，X中包含zt+3，及N-1个负样本，负样本可以来自本语音样本的其他时段，可以来自其他语音样本。    图片数据：图片为256\*256，进行随机数据增强后，使用一个64\*64的滑动窗口，以32\*32的步长对图片进行取样，最终获得7\*7=49个patch，每个patch的大小为64\*64，先随机裁剪为60\*60，再拉伸为64\*64，这是对每个patch进行的数据增强，每个patch经过CNN进行特征提取后变成长度为1024的特征向量，即49个patch经过特征提取网络genc后得到一个7\*7\*1024的特征向量张量（即z，其中zij表示第i行第j列这个patch对应的特征向量）。每个zij及其上方的所有特征向量（z1jz2j...zij）构建为一个序列送入gar，获得cij然后用cij进行zij下方的预测。  原理：通过c能够良好地预测下面的x，这意味着模型最大化了c与x之间的互信息。互信息是指已知c后，x的不确定性的减少量，互信息越大意味着已知c时，x的不确定越小，意味着c这个特征能够较好地预测x，这也意味着经过genc对原始样本进行编码后获得地编码信息蕴含着样本中的上下文信息。 |