零样本图像分类指训练集和测试集在数据的类别上没有交集的情况下进行图像分类.

自提出此问题至今, 零样本 图像分类的研究已经大致有十年时间.

1. 零样本图像分类技术的研究意义及其应用价值,

2. 总结和归纳零样本图像分类的发展过程和研究现状

3. 接下来介绍常用的数据集和评价准则, 以及与零样本学习相关的技术的区别和联系

4. 分析有待深入研究的热点与难点问题, 并对未来的发展趋势进行了展望.

研究意义、应用价值

零样本分类 (zero-shot classification, ZSC), 也称作零样本学习 (zero-shot learning), 是指**依据一些可见类别 (seen classes) 的数据**, 辅以**相关常识信息或先验知识 (称为辅助信息, side information),** 用于训练某种学习模型, 实现对**未见类别 (unseen classes, 指数据标注完全缺失的类别) 的数据进行类别**预测和识别的一类技术.

已感知到的信息；辅助信息；🡪未感知到的信息。

将一些文字描述作为辅助信息（先验知识），然后学习一部分图像，结合文字进行推理，可用知道未见图像的类别。

现有的物体分类技术大多属于监督学习, 即需要大量标注数据训练相关模型 [7,8]. 据统计 [9], 人类可识别的物体种类大约有 3 万类，要对如此之多的类别进行数据标注是一件极其费力的工作。🡪所以如果有zero-short learning，我就可用只学会一部分类别，然后用文字告诉智能体其他类别和现有类别有什么区别、其他类别有什么特点，就可用在见到新的类别的时候给出更可靠的预测。

零样本分类技术很大程度上受到了人类学习过程的启发. 在学习机理上 与传统的分类技术的不同, 决定了零样本分类必须走一条继承和发展的研究道路。

假设一个小孩从未见过斑马 (斑马对小孩来说是未见类), 但他知道斑马是一种身上有黑白条纹的外形像马的动物 (马 对于小孩来说是可见类), 那么当他第一次看到斑马的时候, 一般都能认出这是斑马. 🡪 真的吗？真的不是大人的臆想吗？有实验证据支撑吗？

“触类旁通” 学习过程可以总结为 “已有学习模型 (认识马) + 常识或先验 (知道斑马是身上有黑白条纹的外形像马的动物) + 进行演绎推断 (识别斑马)” 的过程, 这正是零样本分类的思想核心, 即利用 常识或先验知识进行演绎推理. 🡪 所以零样本学习就是感知+推理的问题。

在 2008 年前后又提出了零样本学习技术 [1∼3], 目的是研究某些类别标注完全缺失的情 况. 零样本分类可看作视觉数据与文本等其他模态数据间的一种跨模态学习 [12∼14]. 这与人类的学习过程类似: 我们既不是从海量的图像或者视频数据集中学习相关类别概念, 也不是从海量语料库中学习语言 (语义), 而是从视觉和语言 (语义) 的交互中理解了世间万物. 🡪 从语义中推理，从视觉中感知。

从理论体系从属关系的角度来看, 零样本分类属于迁移学习 [26∼28] 领域的一个子任务. 🡪 我的感觉是对的，这就是迁移学习，从一个领域内学习文本表征，然后用到另一个领域。不如说迁移学习是零样本学习的前一个步骤。

迁移学习 的主要目的是将源域 (source domain) 和源任务 (source task) 中的知识迁移到目标域 (target domain) 中用于帮助目标任务 (target task) 进行学习.

从模型 实现的角度来看, 零样本分类的实现与另外两个研究领域密不可分, 其一是表征学习 (representation learning), 其二是度量学习 (metric learning).

表征学习是指通过对数据进行变换从而提取数据中的有效信息的一种学习方式, 涉及到人 工智能相关的诸多领域, 如语音识别 [29]、信号处理 [30]、目标识别 [7]、自然语言处理 [31], 以及迁移学习 [32] 等

嵌入学习：将多个模型的数据嵌入到同一语义空间中进行数据表征, 其本质可以视为一种以语义表示为核心的表征学习过程.

度量学习通常建立在表征学习的基础之上, 其本质是根据不同的任务, 根据特定空间中 的数据自主学习出针对某个特定任务的距离度量函数, 目前已被广泛应用于诸多计算机视觉相关的任 务, 如人脸识别 [33]、图像检索 [34]、目标跟踪 [35]、多模态匹配 [36] 等.

对于零样本分类任务, 在获取 到合适的数据表征空间之后, 则需要对跨模态样本间的距离度量进行学习, 目的是保证嵌入到语义空 间后样本间的语义相似度关系得以保持. 因此, 零样本分类又可以看作是在不同域中的度量学习过程.

**零样本学习可以看作是在进行表征学习和度量学习的基础上, 通过借助辅助信息 (属性或 文本) 实现跨模态知识的迁移, 从而完成可见类信息到未见类信息推断的迁移学习过程.**

关于零样本分类的综述文献较 少, 其中影响力较大的分别是 Xian 等 [37], Fu 等 [38], 以及 Wang 等 [39] 的工作. 其中, Xian 等 [37] 对目 前零样本分类的部分经典分类模型做出了一些概括性总结,

Xian Y Q, Schiele B, Akata Z. Zero-shot learning-the good, the bad and the ugly. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, 2017. 3077–3086

Fu Y W, Xiang T, Jiang Y G, et al. Recent advances in zero-shot recognition: toward data-efficient understanding of visual content. IEEE Signal Process Mag, 2018, 35: 112–125

Wang W, Zheng V W, Yu H, et al. A survey of zero-shot learning. ACM Trans Intell Syst Technol, 2019, 10: 1–37

Zero-shot learning 在NLP中的应用

|  |  |
| --- | --- |
| Natural language processing. In the area of natural language processing (NLP), the application of zero-shot learning has been emerging in recent years. In bilingual dictionary induction [102, 105] and bilingual phrase table induction [28, 102] problems, from some seed bilingual word or phrase pairs, zero-shot learning helps to construct bilingual dictionaries or bilingual phrase tables. It is especially useful for scarce language pairs, in which word or phrase pairs are rare. In machine translation problems, zero-shot learning is used for zero-shot translation (also called zero resource translation) for language pairs that do not have parallel corpora [37,69,106,177]. Besides, zero-shot learning has also been used for spoken language understanding [36, 168] (i.e., spoken language classification in new domains) and semantic utterance classification [22] (i.e., classification of new semantic classes). Finally, it has been used in other problems related to NLP, such as entity extraction from web pages [117], fine-grained named entity typing [99], cross-lingual document retrieval [47], and relation extraction [85]. | 自然语言处理。在自然语言处理(NLP)领域，近年来出现了零样本学习的应用。在双语词典归纳法[102,105]和双语短语表归纳法[28,102]问题中，从一些种子双语单词或短语对中，零样本学习有助于构建双语词典或双语短语表。它对稀有的语言对特别有用，在这些语言对中，单词或短语对是罕见的。在机器翻译问题中，对于没有平行语料库的语言对，零样本学习用于零样本翻译(也称为零资源翻译)[37,69,106,177]。此外，零样本学习也被用于口语理解[36,168](即零样本学习)。(1)新领域口语分类;(2)语义话语分类。，新语义类的分类)。最后，它还被用于与NLP相关的其他问题，如从web页面中提取实体[117]、细粒度命名实体类型化[99]、跨语言文档检索[47]和关系提取[85]。 |

Applications of Zero-Shot Learning

<https://towardsdatascience.com/applications-of-zero-shot-learning-f65bb232963f>

zero-short learning in natural language processing

<https://www.google.com/search?sxsrf=ALeKk00rDsqnU2DnXr2vRCBs1ZXpqiMATQ%3A1588596271978&source=hp&ei=Lw6wXpONOdPG-gSbqICgCQ&q=zero-short+learning+in+natural+language+processing&oq=zero-short+learning+in+natural+language+pro&gs_lcp=CgZwc3ktYWIQAxgAMgUIIRCgATIFCCEQoAEyBQghEKABOgIIADoFCAAQywE6BAgAEB46BggAEAoQHjoGCAAQCBAeOggIABAIEAwQHjoICAAQBRAKEB46CAgAEAwQBRAeOggIABAIEAoQHjoHCCEQChCgAVCqJ1i1e2CMggFoCHAAeAKAAYAFiAGAe5IBCzItMzkuMTAuMS4ymAEAoAEBqgEHZ3dzLXdpeg&sclient=psy-ab>