

专业阅读与写作第二次作业

58119304 朱启鹏

A. 选择方向

我选择的子方向是计算机图形学与多媒体。

认知总结：计算机图形学（英语：computer graphics，缩写为 CG）是研究计算机在硬件和软件的帮助下创建计算机图形的科学学科，是计算机科学的一个分支领域，主要关注数字合成与操作视觉的图形内容。虽然这个词通常被认为是指三维图形，事实上同时包括了二维图形以及影像处理。多媒体（Multimedia），在电脑应用系统中，组合两种或两种以上媒体的一种人机交互式资讯交流和传播媒体。使用的媒体包括文字、图片、照片、声音（包含音乐、语音旁白、特殊音效）、动画和影片，以及程序所提供的互动功能。即计算机图形学与多媒体主要关注的是如何用计算机解决图像、文本、声音的问题。

理由：计算机图形学与多媒体目前应用广泛，在游戏行业、音视频行业等都有很大的应用。并且随着 5G 时代的来临，网络传播的速度越来越快、越来越稳定，短视频行业与自动驾驶行业开始蓬勃发展。而这两个行业都与该领域息息相关，在未来，如何快速、高效处理以不同载体为媒介的信息将至关重要。而我本身就对多媒体有关的领域十分感兴趣，所以此次作业选择深入理解这个方向。

B. 高水平论文

我选择的文章是 [1]Y. R. Choi and R. M. Kil, "Face Video Retrieval Based on the Deep CNN With RBF Loss," in IEEE Transactions on Image Processing, vol. 30, pp. 1015-1029, 2021, doi: 10.1109/TIP.2020.3040847.

理由：目前人脸识别已经在世界有了广泛的应用。生活中我们时常能见到在利用人脸识别技术验证你的身份，只有当我们验证通过我们才有可能能够通过一些例如动车站或者其它站点的关卡。除此之外，人脸识别技术还在追击嫌疑犯上发挥了巨大的作用，而中国准确度最高的人脸识别技术被应用于天眼系统。

这篇文章提出了一个新的框架来提取人脸视频的高度紧凑和判别特征使用深度卷积神经网络的检索任务（美国有线电视新闻网）。人脸视频检索任务是找到视频包含来自数据库的特定人脸与查询相同的人的面部图像或面部视频。关键挑战是用小存储提取判别特征来自具有较大类内变化的面部视频的空间通过不同的角度、光照和面部表情。在最近多年来，基于 CNN 的二进制哈希和度量学习方法在图像/视频检索任务方面取得了显著进展。然而，现有的基于 CNN 的二进制哈希和度量学习在不可避免的信息方面有局限性分别是损失和存储效率低下。为了应对这些问题，提出的框架由两部分组成：第一，使用径向基函数核（RBF）的新型损失函数损失）被引入训练神经网络以生成紧凑和判别高级特次，优化使用逻辑函数（Logistic Quantization）的量化是建议将实值特征转换为 1 字节整数最小的信息损失。通过人脸视频检索在具有挑战性的电视连续剧数据集（ICT-TV）上进行的实验，它是表明提议

的框架优于现有最先进的特征提取方法。此外，RBF 损失的有效性也通过 CIFAR-10 上的图像分类和检索实验和使用 LeNet-5 的 Fashion-MNIST 数据集。索引词——人脸视频检索，卷积神经网络、损失函数、径向基函数、度量学习。

C. 高水平会议

名称：IEEE Visualization Conference

网址：<http://dblp.uni-trier.de/db/conf/visualization/index.html>

理由：IEEE Visualization Conference (VIS) 是计算机图形学与多媒体的一个重要会议。根据 2016 年 Google Scholar 的 h-index 指标排名，VIS 是视觉研究会议评分最高的会议，也是计算机图形学评分第二高的会议。它在澳大利亚 ICT 会议排名中被评为“A”级。以及巴西教育部的“A”评级。该会议具有高度选择性，所有论文的接受率通常小于 25%。这是一场在该领域认可度极高的会议，所以我选择这场会议。

D. 高水平期刊

名称：IEEE Transactions on Image Processing

网址：<http://dblp.uni-trier.de/db/journals/tip/>

理由：IEEE Transactions on Image Processing 主要发表关于图像处理和计算机视觉领域的学术论文。IEEE Transactions on Image Processing 的发布是经过经验丰富的科学同行严格审查后的结果，所以具有其发布的论文具有非常高的价值，且在业界认可度较高。该杂志鼓励各地研究所投稿，他们主要关注投稿论文中的创新性以及实际意义。所以，这是一份十分有价值的期刊。

E. 相关领域专家

名字：李飞飞

主页地址：<https://profiles.stanford.edu/fei-fei-li>

理由：李飞飞博士是斯坦福大学计算机科学系的首任红杉教授。她于 2013 年至 2018 年担任斯坦福大学 AI 实验室主任。2017 年 1 月至 2018 年 9 月在斯坦福大学休假期间，她担任谷歌副总裁，并担任谷歌云 AI/ML 首席科学家。李飞飞博士于 1999 年获得普林斯顿大学物理学学士学位，2005 年获得加州理工学院（Caltech）电气工程博士学位。2009 年加入斯坦福大学担任助理教授。在此之前，她曾在普林斯顿大学（2007-2009）和伊利诺伊大学香槟分校（2005-2006）任教。

李飞飞博士目前的研究兴趣包括受认知启发的人工智能、机器学习、深度学习、计算机视觉和人工智能 + 医疗保健，尤其是用于医疗保健服务的环境智能系统。过去，她还致力于认知和计算神经科学。李博士在 Nature、PNAS、Journal of Neuroscience、CVPR、ICCV、NIPS、ECCV、ICRA、IROS、RSS、IJCV、IEEE-PAMI、New England Journal 等顶级期刊和会议上发表学术论文 200 余篇医

学、自然数字医学等。李博士是 ImageNet 和 ImageNet Challenge 的发明者, ImageNet Challenge 是一个重要的大规模数据集和基准测试工作, 为深度学习和 AI 的最新发展做出了贡献。除了她的技术贡献外, 她还是倡导 STEM 和 AI 多样性的全国领先声音。

她还是 ACM 研究员、对外关系委员会 (CFR) 成员、2019 年 IEEE PAMI Longuet-Higgins 奖、2019 年国家地理学会进一步奖、2017 年雅典娜学术领导奖、IAPR 2016 JK Aggarwal 获得者奖, 2016 IEEE PAMI Mark Everingham 奖, 2016 nVidia Pioneer in AI 奖, 2014 IBM Faculty Fellow Award, 2011 Alfred Sloan Faculty Award, 2012 Yahoo Labs FREP 奖, 2009 NSF CAREER 奖, 2006 Microsoft Research New in Faculty Fellowship 其他。李博士是许多学术或有影响力的会议的主讲人, 包括世界经济论坛 (达沃斯)、Grace Hopper Conference 2017 和 TED2015 主要会议。李博士实验室的工作已在各种杂志和报纸上发表, 包括《纽约时报》、《华尔街日报》、《财富》杂志、《科学》、《连线杂志》、《麻省理工科技评论》、《金融时报》等。她被 ELLE 杂志选为 2017 年科技女性, 被 Good Housekeeping 选为 2017 年杰出女性奖, 被外交政策选为 2015 年全球思想家, 2016 年被卡内基杂志选为“伟大的移民: 美国的骄傲”之一 Foundation, 历届获奖者包括阿尔伯特·爱因斯坦、马友友、谢尔盖·布林等。

她虽然是一名华裔, 但是却在计算机视觉领域做出了很大的贡献。她通过锲而不舍的收集图片大数据, 最终制作出 ImageNet 这样优秀的数据集, 为计算机视觉的蓬勃发展奠定了坚实的基础。并且他在 Ted 上的演讲也另外印象深刻, 她曾说过, “现在回头看, 用大数据来训练计算算法的做法或许显而易见。但是在 2007 年, 这并不显然。在这段旅途中, 我很长一段时间都非常孤独, 有些同事友情建议我做些对获得终身教职更有用的事。我们在研究经费方面也一直碰到麻烦, 我可能需要重开我的干洗店来为 ImageNet 筹资。”可见 FeiFei Li 为 ImageNet 数据集的付出是显而易见的。这不正是作为一名计算机科研人员要学习的榜样吗?

REFERENCES

- [1] Young Rok Choi and Rhee Man Kil. “Face Video Retrieval Based on the Deep CNN With RBF Loss”. In: *IEEE Transactions on Image Processing* 30 (2020), pp. 1015–1029.