

袁 勇

湖北 | 男

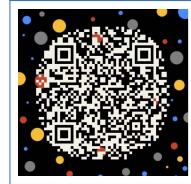
北京市海淀区

150-2955-xxxx

willard.yuan@gmail.com

yongyuan.name/blog

willard-yuan



教育背景

2013.9 - 2016.6 硕士学位, 中国科学院大学, 信号与信息专业, 保研.

2009.9 - 2013.6 学士学位, 西安电子科技大学, 电子信息科学与技术专业, 专业 top 3%.

工作经历

元石科技, 大模型初创公司. 2023.12 - now, 登山中...

快手, 搜索技术部-算法工程师, 职级 K4b, 2020-2023, 3 年绩效全部保持在 A 及以上

2020.12 - 2023.11 搜索精排 QIN (Query-dominant User Interest Network) 个性化大模型消费预估, 熟悉搜索推荐领域前沿行为序列建模方案.

- QIN 3.0 即时兴趣建模范式. QIN 个性化持续性优化的延续, QIN 3.0 弥补了 2.0 中侧重长期兴趣、短期兴趣而忽略了用户的即时兴趣。用户的兴趣通常存在中长期、即时兴趣, 在 QIN 3.0 中, 通过客户端下发的用户超实时行为序列 (最近 20 条), 通过 Target Attention 学习用户的即时兴趣, 构成对用户兴趣的长期、短期、即时表征。。实验收益: 设备均使用时长 +0.061%, 搜索播放时长 +1.838%, 搜索换 Query 率-0.521%, 内流总播放时长 +7.798%。成为全年推全实验中收益最显著的实验之一。
- QIN 个性化大模型样本实时化及流式训练. 提升训练与推理特征数据一致性, 避免样本特征发生穿越, 全局快速捕捉用户兴趣变化, 提高流量转化效率。为避免从离线拼接 (批式) 一步跨越到实时 (流式), 导致样本特征、label 存在较大 diff, 增加了一个中间过渡环节, 特征在线 dump, label 仍保持于离线 label 一致。在线取得搜索总播放时长 +1.151%, 综合搜索首页有点比 (自然结果)+0.227%, 搜索换 Query 率-0.431%。随机 GSB=0: 6: 89: 7: 0, 样本延迟: 2d 降低到 70 mins.
- 时长目标牵引, 在 QIN 中引入时长预估。针对快手搜索特点, 进行业务定制, 分别实验了 Youtube 的时长预估方案和采用物理时长分桶、播放时长再分桶进行时长削偏方式, 都取得了很好的收益。基于 Youtube 的 weighted logloss 方案, 在线取得搜索总播放时长 +1.318% 收益。采用时长削偏方式, 在线取得搜索总播放时长 +1.910% 收益。
- 创新性提出 QIN 2.0 双阶段检索个性化大模型. QIN 2.0 包含两个原创性工作, 分别为 RSU (Relevance Search Unit) 和 FAU (Fused Attention Unit) 模块。RSU 获取 target photo 的共享子序列和独有子序列, 先拿 query 的多模 embedding 检索最相似 N 条结果, 再拿 target 视频 embedding 检索最近 50 条。QIN 2.0 一阶段只计算一次, 扩展序列成本低, 筛选出与当前 query (显式 context) 相关性较好, 且对 target photo 预估有帮助的子序列。FAU 针对 ID 类特征与内容类特征表征特点, 设计出一种新颖的融合单元, ID 类特征与内容类特征分别计算 attention, 用户交互度反映用户对 photo 的兴趣, 通过用户交互度设计 Engagement Gate 具备良好解释性。线上累加收益: 搜索设备次日留存率 +0.087%, 搜索总播放时长 2.056%, 综合搜索首页有点比 (自然结果)+1.336%, 搜索换 Query 率-1.115%。同时该工作已被 CIKM 2023 接收。
- 原创性提出 QIN 1.0 个性化大模型. 在 query 与 photo 多塔匹配模型的基础上, 单独设计了个性化模块, 加入行为兴趣建模。为刻画用户在当前 query 下的兴趣, 同时兼顾搜索相关性, 构建了 query 和 photo 语义匹配多模 embedding, 从用户历史行为序列中通过 query embedding 检索出行为子序列, 作为 QIN 模型的序列特征。线上取得搜索设备次日留存率 +0.19%, 搜索总播放时长 +1.584%, 综合搜索首页有点比 (自然结果)+1.27%, 搜索主动换 Query 率-1.543% 收益, 随机 GSB=21 (2) -148 (12) -19 (1), 长尾 GSB=24 (5) -136 (23) -17 (2) .
- DNN 模型升级为 query 与 photo 多塔匹配模型. 按信号维度拆成相关、视觉感知、质量权威、时鲜维度构建多塔模型, 特征采用 beta 分布进行平滑处理, 模型加宽加深, 各个塔内部做柔性特征权重选择, 预估目标从单目标升级为 MMOE 多目标预估, 夯实了 query 与 photo 多塔匹配模型基础。
- 从 0 到 1 构建搜索精排消费分预估 DNN 模型. 搜索行为日志包含大量 query 和 photo 相关性信息, 用户 query 下的交互度刻画了用户对搜索结果满意程度, 利用用户海量行为以及 DNN 端到端建模, 对用户的消费偏好进行预估。在线替换了基于相关性与消费分档的树模型预估方案。

快手，多媒体内容理解部-视觉算法工程师，职级 K4a，2016-2020，4 年绩效全部保持在 A 及以上

短、长视频和图像检索、查重与拷贝检测，6 年视觉检索经历，掌握了若干工业视觉检索系统线上主流算法。

2018.12 - 2019.8 **智能配乐推荐，视频搜索**，基于视频内容推荐最相关的候选音乐、文本搜视频。

- 以视频检索为核心技术，为智能配乐增加相似（热歌）召回源以及原生召回源，增加精排序模型。相似（热歌）召回源 AB 实验召回效果优于或可比其他召回源；精排序模型 AB 实验相比 base 排序提升 2pp 左右，相似（热歌）召回源以及原生召回源线上已推全。
- 索引算法升级。以 DNN 及深度度量学习为基础，开发了一套基于 Spreading Vector 的特征索引算法，相比线上算法，离线评估额外增加召回率 9.8pp（填平与 Faiss 差 5pp 的距离，并在此基础上增加了 4.8pp），线上评估新增查重独立召回 2.7% 左右，最大时延由 190ms 降低到 110ms 左右。
- 长视频查重优化。从关键帧检测算法等方面，对长视频查重进行重点优化，优化后 Recall@10 相比优化前，在业务反馈的 badcase 上，召回率提升了 16% 左右。
- 负向与安全项目。从 0 到 1 新增政治海报查重自动检测算法。平均每日召回约 6 万左右，准确率在 90%+；负向直播截屏优化，第 1 期优化召回率提升 12.5%，第 2 期优化离线评估新增检测召回 3.1% 左右；logo 检测类别接近 1000 种，准确率维持在 99.9% 以上，热门页 logo 维持在万分位以下。
- 商品识别与检索。负责拍照采购项目，覆盖 28 个大类商品的检测与检索，商品识别准确率提升 XXX%。
- 中台业务需求。为业务提供主体检测算法、关键帧检测算法、短视频去水印算法、游戏识别算法、主色调提取算法等。

2018.01 - 2018.12 **第三代视频原创保护系统**，对百亿级短视频（千亿级视频帧）进行第三代算法升级，同时拓展至长视频原创保护，个人绩效 S。

- 特征升级为 simNet 和 dupNet。带领同事设计并研发了一种统一视觉近似重复和语义相似的特征学习框架。在此基础上分别获得了专门面向视频查重的 dupNet 特征和既可以面向视频查重又可以面向语义相似的 simNet 特征，7000 万离线数据集上评估，在暴力搜索的条件下，simNet 比线上第二代特征 top@10 召回率提升了 14.5%，dupNet 在 100 万离线数据集上评估，相比 simNet 在 top@10 召回率高 8% 左右，上线后平均每日新增召回 25000 左右，dupNet 上线应用于 A 站长视频原创保护。
- 索引升级。带领同事设计并研发了一种结合图 ANN 和 OPQ 的索引算法 hnsWOPQ。保持了 OPQ 特征高压缩比、索引动态增删灵活的特征，同时引入先进的图索引技术，保证获得比 OPQ 更高的召回率，在 1 亿 128 维业务数据集上测试，以暴力搜索结果作为基准，top@50 召回率提升由 61% 左右提升至 82% 左右。
- 局部特征检索系统。从 0 到 1 构建了一套局部特征检索系统，在公开数据集 ROxford 数据集上 MAP 为 XXX，应用于安全、政治违禁等查重检测上。
- 基于 LBS 视觉检索同城热点事件感知与发现。以视频检索技术为基础，结合 LBS 信息构建 Graph 聚合相似视频。平均每日向同城输出 150 个可用的热点，上线后，漫游页的 dau 从原来的 40w 50w 提升至目前的 240w 左右；为天象平台开发地方热点、城市热点、站外魔表发现、灾害监测自动发现算法。

2016.12 - 2017.12 **第二代短视频原创保护系统**，对从 0 到 1 的短视频原创保护系统进行算法升级，以适应百亿级短视频原创保护，个人绩效 S。

- 特征索引。设计并研发了一种基于 OPQ 特征量化编码的视频多帧索引及排序方法，全量索引平台 xx 亿视频量，与视频多帧倒排结构相比，检索召回提升 4 倍以上，全量上相似查重独立覆盖由 2% 提升至 5.3% 左右，线上每周收到的 basecase 显著减少。
- 校验算法。设计并研发了一种基于 CNN 特征和局部特征匹配相结合的视频查重多帧校验方法，提高了视频查重校验对错帧、镜像等编辑的鲁棒性；开发了一种基于 GIST 全局特征更严格的图像校验方法，弥补局部特征校验的召回以及适配游戏查重校验等高度相似的场景。
- 重排算法。设计并开发了一种视频（多帧）相似重排方法，上线并显著改善了相似查重的视频查重效果。
- 特征表达。设计基于 CNN 和传统特征融合的视觉检索系统，解决实例检索问题，Oxford Building 数据集 mAP 取得 80%。

视频物体检测，视频 Logo 检测游戏识别等。

- 设计并开发了两种 Logo 识别算法，分别为：基于 OCR 文字识别的 Logo 识别算法，基于 Faster RCNN 结合 BoVW 模型以及局部特征校验的 Logo 识别算法，升级了并维护基于局部特征匹配 Logo 识别系统。检测准确率：99.5%+，Logo 检测类别达 1000+ 多种，Logo 每日召回占视频上传总量 xxx%。
- 设计并开发了尺寸极小（目标尺寸 \leq xxx）且目标极不明显的 Logo 检测方法；开发了王者荣耀英雄、视频截屏、直播游戏（xxx 种）等识别算法，检测准确率 100%。
- 为嗨声、珍珠港、kwaigo 等提供自动去水印需求。

美团, 算法工程师, 2016

2016.7 - 2016.12 美团, 外卖风控部, 算法开发.

- 设计开发了新的商户抓取策略, 维护、优化已有的商户抓取逻辑, 对新美大外卖业务的数据做日常的分析、监控及报表.

中科院西安光学精密机械研究所, 学生创业团队, 算法工程师

2015.10 - 2016.6 ETRACK 眼控技术, 学生创业团队, 算法设计与开发.

- 负责设计瞳孔检测与瞳孔中心检测算法并用 C++ 实现, 采用 OpenMP 实现多线程实时处理.
- 负责设计瞳孔中心到屏幕坐标的映射方法并用 C++ 实现, 通过安卓 NDK 供 JAVA 调用.
- 负责桌面版本的开发, 使用 QT 框架构建图形界面, 人眼检测算法核心部分: [视频演示](#).

出版物

论文与书刊

- 2023.8 **Query-dominant User Interest Network for Large-Scale Search Ranking**, Tong Guo, Xuanping Li, Haitao Yang, Xiao Liang, **Yong Yuan**, Jingyou Hou, Bingqing Ke, Chao Zhang, Junlin He, Shunyu Zhang, Enyun Yu and Wenwu Ou. CIKM' 23 .
- 2016.1 Latent Semantic Minimal Hashing for Image Retrieval. IEEE TIP, 2016 (MINOR REVISION) , [project](#), Xuelong Li, **Yong Yuan**, and Xiaoqiang Lu.
- 2014.4 Learning Hash Functions Using Sparse Reconstruction. ACM ICIMCS, pp. 14-18, 2014 , **Yong Yuan**, Xiaoqiang Lu, and Xuelong Li, (**Best Paper Runner-up Award**).
- 2021.2 [Python 计算机视觉编程](#) (译作), 图灵出版社, 2014 , 朱文涛, **袁勇**.

授权专利

- 2023.9 [多媒体文件的检索方法、装置、设备及存储介质](#) (CN112015922A), 张存义, **袁勇**, 刘岸, 丁冬, 张金帅.
- 2023.6 [一种视频展示的方法、装置、电子设备及存储介质](#) (CN112752158A), 颜秉武, 韩少鹏, 白丽娟, **袁勇**.
- 2023.4 [Method and Device for Verifying Image and Video](#) (US20210118119) , **Yong Yuan**, Yan Li, An Liu, Jie Yang, Bingqing KE, Yanjun Ma, Cunyi Zhang, and Dong Ding.
- 2022.9 [特征提取网络训练方法、图像处理方法及装置](#) (CN113505256A), 张磊, 李攀登, 李岩, **袁勇**, 谢洪涛, 张勇东.
- 2022.6 [Method and Electronic Device for Displaying Video](#) (US20220210511) , Bingwu Yan, Shaopeng Han, Lijuan Bai, **Yong Yuan**.
- 2022.5 [图像检索的方法、装置、服务器及存储介质](#) (CN110275970A), 张存义, **袁勇**.
- 2021.9 [图像检索方法、装置及存储介质](#) (CN110134804A), 杨杰, **袁勇**, 李岩, 刘岸.
- 2021.2 [视频展示方法、装置、电子设备、服务器及存储介质](#) (CN112367559A), 周静, 王慧, 刘付家, **袁勇**, 李新.
- 2022.8 [一种视频查重方法、装置及电子设备](#) (CN112399236A), **袁勇**, 刘岸, 杨杰, 李岩.
- 2016.10 [一种基于潜在语义最小哈希的图像检索方法](#) (CN106033426A), 李学龙, 卢孝强, **袁勇**.

科研经历

中科院西安光学精密机械研究所, 2013-2016

2013.3 - 2016.6 基于内容的图像检索 (CBIR), 课题研究方向.

- 熟练掌握 BoW、VLAD、FV 等特征编码方法, 精通 CBIR 及大规模索引技术.
- 掌握并积累机器学习中常用的降维、聚类、分类以及图像物体识别技术.
- 提出并发表一种基于稀疏表达的哈希编码方法, 详见[HABIR 工具包主页](#).
- 对同款物体的检索有较多的积累经验 (衣服、鞋子等大型图像库 30 万); 有对特定类图像诸如皮革、纺织图像等进行检索的经历; 在 13 万量级的图库上做过广告 logo 的搜索.
- 对人脸数据检索与识别、医学影像检索有相应的经历, 并对深度学习 (CNN 卷积神经网络) 具备一定的理解.

- 2015.1 - 2015.4 **基于卷积神经网络的 CBIR 演示原型系统 PicSearch**, 兴趣驱动型项目, 协作开发.
- PicSearch 是一个在线图像检索原型系统, 使用了 CNN 卷积网络模型.
- 线下完成图像特征的提取, 并做了一定的降维处理, 后台在线特征匹配与排序用 python 实现, 服务器采用了 python 轻量级 web 开发框架 CherryPy, 采用 Bootstrap 框架优化前端交互界面.
 - 图库为包含 29780 张图片的 Caltech-256 公开数据集, 采用特征常驻内存的方式进行了代码的优化, 使其能及时地响应用户的查询请求 (毫秒级), 在线演示地址 PicSearch: search.yongyuan.name(已下线), 本地演示效果: [演示视频](#).
- 2015.3 - 2015.7 **基于词袋模型的物体检索原型 DupSearch 系统**, 兴趣驱动型项目, 独立开发.
- DupSearch 是一个针对 Object Retrieval 或 Duplicate Search 而写的图像检索原型系统.
- 在 Oxford Building 公开数据库上平均检索精度达到 83.35%, 对于光照、旋转、视角等具有较好的适应性, 在线匹配在服务器上能较快的响应查询, 并且在不复杂化现有模型情况下仍有改进提高 MAP 的空间.
 - 图像库测试规模 30 万, 取得了很不错的检索效果, 算法原型系统已售予某公司, 15 万衣服库检索示例详见[GitHub](#), 此外, 对于广告 logo 的搜索也能取得很高的检索精度.
- 2014.7 - 2015.5 **复杂低空飞行的自主避险理论与方法研究 (973)**, 项目参与者.
- 多源协同感知周围环境, 对复杂低空环境中可能的危险障碍物进行实时检测, 并完成飞行器的自主避险.
- 负责可见光传感器数据与激光雷达传感器点云数据的融合, 消除高压线检测时的误检.
 - 负责桥梁、高压线塔、作为异常目标入侵的滑翔机等危险障碍物的实时检测.
 - 采用 opencv、dlib 等计算机视觉开源库, 非电力线类障碍物检测采用 HOG+SVM 物体检测方法.

开源项目

- 2017.7 - 2020 构建 CBIR 领域传统特征与深度学习方法做图像检索的对比框架, 详见[GitHub](#).
- 2016.8 - 2016.9 以 SeetaFaceEngine 为基础, 使用 LSH 索引技术构建了一个人脸检索系统, 详见[GitHub](#).
- 2015.4 - 2016.4 基于 MatConvNet 以及 VGGNet 卷积神经网络模型构建的一个用于图像检索的实验工具包, 详见[GitHub](#).
- 2013.2 - 2016.6 实现并整理了一些流行的哈希算法及多种指标评价, 目前该 Matlab 工具包已更新至 V2.0, 详见[GitHub](#).
- 2013.12 - 2014.6 翻译《Programming Computer Vision with Python》时, 为使读者更易于理解书中的内容, 重新对书上的代码做了整理, 并放在 github 上, 详见[项目主页](#).
- 2014.2 - 2014.5 基于稀疏重构的哈希编码方法的 Matlab 代码及检索指标评价, 详见[GitHub](#).


奖项

- 快手优秀个人奖 (2022)
- 中科院三好学生 (2016.4)
- 优秀学生巡回报告团成员 (2012.12)
- 校内一等奖学金 (2011.11)
- Best Paper Runner-up Award(2014.7)
- 国家奖学金 (2012.11)
- 国家励志奖学金 (2010.11)

个人技能栈

- 编程语言 会 C++/C、OpenCV 以及 QT, 熟练 Python, Matlab, SQL, 熟悉 HTML, CSS, Spark, Lua
- 算法技能 精通搜索推荐行为序列建模, 3 年搜索经验, 5 年图像视频检索经验, 熟练掌握深度学习与传统机器学习算法, 对 LLM 大模型有较好的理解
- 英语 CET-6 和 CET-4, 具备专业英文文献阅读、写作及翻译能力, 平时会保持对 CVPR 等论文的阅读.
- 常用工具 OS X、Linux、TensorFlow、Pytorch、OpenCV、VSCode、Ipython Notebook、Git
- GitHub github.com/willard-yuan

其他

-  喜欢编码, 热爱开源, 有写[博客](#)进行总结的习惯; 具备一定的管理能力.