

智眼·图像视频识别理解

产品简介

拓尔思信息技术股份有限公司

“智眼” 图像视频识别理解服务平台

产品背景

近年来，随着大数据、机器学习、计算机、脑科学等领域研究的不断进步，人工智能技术发展迅速；同时伴随着短视频以及社交媒体应用的大规模兴起，图像及视频数据量急速增多。在此背景下，为把握人工智能产业机遇，适应行业技术迅速发展需要，拓尔思推出了“智眼”图像视频识别理解服务平台，为客户提供更好的产品及解决方案，提升公司的盈利能力和市场竞争能力。

方案概述

该平台集成了主流的深度学习框架及深度学习模型。在图像及视频检索和分类、图像及视频特征提取、重复图像及视频检测、图像及视频目标检测识别、图像及视频 OCR 文字提取和场景分类等关键技术的基础上，通过深度学习模型理解图像及视频，提取图像及视频内容特征，建立图像及视频检索系统。用户通过输入图像及视频，可快速实现图像及视频的智能识别和处理；以及在图像视频库中快速检索到与输入视频相似的图像及视频片段。

平台可以实时地从海量图像及短视频数据中识别并筛选出涉黄涉暴、政治敏感等非法数据和广告推送、恶趣低俗等违规数据，保障互联网的图像及视频内容安全和品质服务，可广泛用于版权鉴定及内容审核等场景。

产品功能



◎ 图像分类

本功能基于定制的深度卷积神经网络模型实现。支持 ImageNet 的 1000 类分类以及定制模型分类功能。

在线接口：<http://ai.trs.cn/DL-VRS/docs> 中图像分类部分

◎ 以图搜图

此功能基于定制的深度卷积神经网络模型提取的深度特征实现。深度特征集成图像的高层语义特征，利用 CNN 和 RNN 等深度学习算法得到的深度特征不仅保持了一定的不变性，而且还包含了更多的高层语义信息，可以有效地缩小底层特征与高层语义之间的鸿沟。同时，相邻位置的特征往往是高度相关的，因此可以检测到图像的局部特征，可以有效解决相似图像检索问题。

在线接口：<http://ai.trs.cn/DL-VRS/docs> 中以图搜图部分

图像检索 Demo：http://ai.trs.cn/image_demo/image/image_search

舰船检索 Demo：<http://ai.trs.cn/graphdemo/>

◎ 图像敏感识别

此功能采用定制的深度卷积神经网络模型实现。支持 NSFW 识别及暴恐识别。

在线接口：<http://ai.trs.cn/DL-VRS/docs> 中图像敏感识别部分

◎ 人脸识别

此功能采用目标检测和人脸特征比对相结合的方法。第一步先把图片中的人脸位置检测出来；第二步提取人脸部分图片的特征；第三步和库中的人脸进行相似度比对，得到结果。支持自建库。帮助用户识别图像、视频中的重要人物及敏感人物。

在线接口：<http://ai.trs.cn/DL-VRS/docs> 中人脸识别部分

人脸检索在线 Demo：http://ai.trs.cn/image_demo/image/face_search

◎ OCR识别

此功能基于 CRNN+CTC 模型框架，在保证识别效果的同时实现模型蒸馏，使模型尺寸缩小，加快推理速度。目前支持简体中文、英文、数字、常用标点的识别。帮助用户识别图像、视频的字幕；帮助用户识别图像、PDF 中的表格及文字信息。

在线接口：<http://ai.trs.cn/DL-VRS/docs> 中 OCR 识别部分

在线 Demo：http://ai.trs.cn/table_image_demo/static/index.html

◎ 目标检测

此功能基于经典深度学习模型架构，实现敏感旗帜、商标 Logo、电视台标及道路行人车

辆检测等功能。模型可定制。帮助用户识别图像、视频中感兴趣目标的位置。

在线接口：<http://ai.trs.cn/DL-VRS/docs> 中目标检测部分

旗帜检测 Demo：<http://ai.trs.cn/flag-detection/>

商标 Logo 检测 Demo：<http://ai.trs.cn/logo-detection/>

电视台标检测 Demo：<http://ai.trs.cn/tv-detection/>

车辆行人检测 Demo：<http://ai.trs.cn/street-detection/>

◎ 音视频处理

此功能主要基于视频处理 SDK，实现音视频基础信息提取、视频抽帧、音视频分离、音视频编码转换以及音视频格式转换功能。帮助用户编辑音视频数据，更方便地对接智能识别业务。

在线接口：<http://ai.trs.cn/DL-VRS/docs> 中音视频处理部分

◎ 智能语音处理

此功能基于深度学习模型，实现语音识别、语音生成及语音分类功能。支持语音中英文识别及生成。帮助用户更方便地将音频数据打标以及将音频转换为文字，更全面地理解视频中的信息。

在线接口：<http://ai.trs.cn/DL-VRS/docs> 中智能语音处理部分

◎ 视频检索

基于深度学习模型提取出的图像特征、文字及人脸信息，实现综合的视频检索功能。

在线 Demo：http://ai.trs.cn/image_demo/image/video_search

产品优势

◎ 接口规范

统一采用 HTTP RESTful 接口实现，并配有 Swagger UI，方便调试。

◎ 处理效率高

支持图像/视频的多种接口传输类型，支持多线程调用以及 GPU 加速计算。

◎ 全要素提取

采用大数据、语义智能等技术，实现对图像/视频内容进行全要素智能提取，包括图像特征、文字信息、语音信息、位置信息等，进行细颗粒度打标，极大地提高处理效率。

◎ 绿色软件

不修改注册表，不修改系统目录。

◎ 模型定制

可根据智拓提供的服务对模型进行训练和验证。

◎ 免费试用

可根据该文档中提供的在线地址或 TRS 人工智能开放平台提供的服务进行免费试用。

◎ 全场景扩展

采用微服务架构，实现图像及视频检索、图像及视频场景分类、视频编辑处理、图像视频审核及版权比对等全场景应用进行可插拔式的扩展。

应用场景

◎ 图像及视频检索

面向大规模图像及视频库，以及多样的文件格式，帮助用户实现更加高效、准确、便捷的检索服务。

◎ 图像及视频场景分类

面向大规模图像及视频库，提供全方位的特征及标签信息。帮助用户实现更加高效的数据打标及数据分类功能。

◎ 视频编辑处理

面向大规模视频库及视频编辑场景，帮助用户更加高效、统一地编辑和转换视频库。

◎ 图像视频审核/版权比对

面向大规模图像及视频库，帮助用户更加精准、准确及高效地进行内容审核工作。

用户价值

◎ 提升数据处理效率

◎ 助力数据检索精准直达

◎ 规范数据存储及应用更加合法和有序

【服务声明】

本文档意在向您介绍拓尔思部分产品、服务的当时相关概况，部分产品、服务的内容可能不时会有所调整。您所购买的拓尔思产品、服务的种类、服务标准等应由您与拓尔思之间的商业合同约定，除非双方另有约定，否则，拓尔思对本文档内容不做任何明示或默示的承诺或保证。

【联系我们】

我们致力于为您提供个性化的售前购买咨询服务，及相应的技术售后服务，任何问题请联系 4006 300229。