Huawei HiAi介绍

1.概述

HiAI是面向移动终端的AI计算平台。

构建三层AI生态：服务能力开放、应用能力开放和芯片能力开放。

服务能力开放：根据用户所需，适时适地的推送服务；

应用能力开放：轻松将多种AI能力与App集成，让App更加智慧强大；

芯片能力开放：快速转化和迁移已有模型，借助异构调度和NPU加速获得最佳性能

HiAI API为对第三方开发者开放的HiAI资源包。其中HiAI API是移动计算平台中的人工智能计算库，该计算库面向人工智能应用程序开发人员，让开发者便捷高效地编写在移动设备上运行的人工智能应用程序。

HiAI API将作为统一的二进制文件发布。这组API主要作用是通过HiAI异构计算平台来加速神经网络的计算，当前仅支持在Kirin SoC上运行。

使用HiAI API，开发人员可以专注于开发新颖的AI应用程序，而不用关注针对计算的性能调优。

HiAI API集成在使用Kirin SoC芯片的android系统上，开发者可以在集成环境中运行神经网络模型，调用HiAI API进行加速计算。HiAI API无需安装，使用移动设备默认镜像即可支持相关集成、开发和验证。

2. 功能介绍

HiAI API为人工智能应用开发人员提供了两个主要功能：

* 提供常用的人工智能业务功能API，可在移动设备上高效运行。
* 提供了一个与处理器硬件无关的加速API，应用厂商和开发者可以在HiAI异构加速系统上加速模型计算、算子计算。

基本功能：

* + 支持模型编译、模型加载、模型运行、模型销毁等AI模型管理接口。
  + 支持基本的算子计算接口，包括卷积、池化、全链接等接口。

3.技术介绍

* + HiAI支持针对神经网络模型运算的专用AI指令集，可以用最少的时钟周期高效并行执行更多的神经网络算子。
  + HiAI通过工具可以把神经网络各种算子，比如卷积、池化、激活、全链接等离线编译成NPU的专用AI指令序列，同时将数据和权重重新摆放，指令与数据融合在一起生成离线执行模型。在离线编译的时候，可以把前后层（卷积、Relu、Pooling）能够融合的算子进行层间融合，这种方式可以减少DDR的读写带宽，提升性能。
  + HiAI能够把神经网络模型中的相关数据以最高效的方式重新摆放，特别是Feature Map的通道数据，在做卷积运算的时候，对于通道相关联的计算效率会得到大幅提升。
  + HiAI支持稀疏化模型加速。在不损失计算精度的前提下，对权重进行置零稀疏优化，NPU可以跳过系数为零的乘加运算，极大地提高了计算效率，同时也降低了带宽。

HiAI支持多种智能平台框架，包括Caffe、Tensorflow等，使用不同的智能平台框架，第三方需要在接口中指出本次计算需要使用的具体的智能平台框架，其他接口和参数无需修改。

HiAI支持绝大部分模型和神经网络算子，并持续优化与完善。

4.应用领域

人脸对比，人脸检测，五官特征检测，人脸属性，人脸朝向识别，美学评分，图片分类标签，图片超分辨率，场景检测，文档检测矫正，人像分割，图像语义分割，文字图像超分辨率，分词，词性标注等等。

5.合作案例

**快手**：快手通过接入HUAWEI HiAI生态，应用强大的HiAI端侧计算机视觉能力，能够实现在同等运算时间下动作识别更精准。借助HiAI芯片开放能力(Kiri 970)，将已有的深度学习算法模型快速转化迁移到华为NPU（进行加速，以高于普通CPU 50倍的能效优势和25倍的性能优势，大幅提高动作识别速度和稳定性，有效解决吞帧、卡顿、不连贯的问题。通过华为NPU对深度学习算法进行加速，也将显著改善功耗表现，在获得极高性能的同时，有效降低手机耗电量，抑制机身发热发烫。同等运算时间下，动作识别错误率降低42%

**Prisma：**将拍摄的普通照片，模仿著名艺术家画作的风格进行处理，实现几十种艺术风格转换。通过接入HiAI开放能力，基于华为Kirin970开发基于端侧的学习模型，实现图像风格渲染的端侧处理，摆脱对网络的依赖。将已有的深度学习算法模型快速转化迁移到华为NPU进行加速，能够轻松处理卷积计算等复杂计算，可以快速实现场景渲染，运用HiAI场景检测能力，轻松准确辨别动物、绿植、食物、建筑、汽车等多类场景，同时适合这类场景的风格滤镜将被标记推荐，风格滤镜选择更智能。

**搜狐新闻客户端：**通过接入HiAI Engine图像超分辨率能力，搜狐新闻能够将图片分辨率放大9倍（高宽各放大3倍），展现更加清晰的图片细节，并有效抑制原图中JPEG压缩噪声，改善小图片在移动终端中放大浏览时的体验。实现图像的端侧处理，摆脱对网络的依赖。基于深度学习在计算机视觉中的广泛应用，将图片进行智能放大，将有效规避传统放大弊端，获得比传统图像处理更加清晰、锐利、干净的照片。

**苏宁易购：**通过接入HiAI Engine图像超分辨率能力，将商品图片分辨率智能放大9倍（高宽各放大3倍），得到更加清晰优质的画面，显著改善用户图片阅览体验。HiAI Engine图像超分辨率能力，还可运用于节省存储和流量；通过在图片存储、传输时加大图片压缩率，缩小分辨率，用户浏览时再用3x超分智能放大，实现同等清晰度下，大幅节省存储空间和数据流量，解决加载缓慢问题。

**携程旅行：**通过接入HiAI Engine美学评分能力，实现自动综合图像的失焦、抖动等技术因素与歪斜、色彩、构图等主观美感，对图像评价打分，让用户可以通过评分高低快速了解照片质量，解决心中疑虑，并据此调整，拍出最美风景。运用HiAI Engine图像超分辨率能力，能够将照片分辨率放大9倍（高宽各放大3倍），画面更加清晰、干净、锐利，轻松解决照片模糊的问题，为用户呈现更优质的旅行照片。

**WPS：**引入HiAI Engine文档检测校正能力后，WPS可以精准识别感知文档所在区域，并自动裁剪虑除屏幕、墙面、桌椅等其他景物。能够自动将拍摄视角调整到正对文档的角度上，最大校正角度达45度。HiAI将对包含文字内容的图像进行9倍分辨率放大（高宽各放大3倍），从而显著提升画质清晰度、增强文字可辨识度。通过接入通用OCR，可自动识别并提取图片中文字信息，实现自由修改、剪切、复制、删除PPT图片中的文字内容。

**汽车之家：**运用HiAI Engine美学评分能力，结合失焦、抖动等技术因素与歪斜、色彩等主观美感，对视频各帧画面进行综合打分，并自动挑选出得分最高的画面作为视频封面，轻松解决第一帧画面欠佳、用户自己挑选耗时费力的问题。

**Microsoft Translator：**在端侧即可完成复杂的神经网络参数运算，从而实现离线翻译，摆脱对网络的依赖。

**StorySign：是一款听障儿童阅读识字应用，**自动校正手机检测画面角度，能够有效识别并校正手机所检测到的书页画面。即使手机和书页的相对角度不佳，也能使所检测到的书页画面在手机上变得规整而易于查看，让后续的文本识别更容易、更准确。HiAI的通用文字识别能力支持StorySign进行多国语言识别，令其可快速准确提取书页上的各国文字内容并将之翻译成手语。同时，它也支持多国手语翻译，用户还可以选择自己偏好，选择英国、法国、意大利等国家的手语种类。