



AI 的嵌入式技术实现

CONTENTS

1. 嵌入式AI发展现状

2. 嵌入式AI典型产品

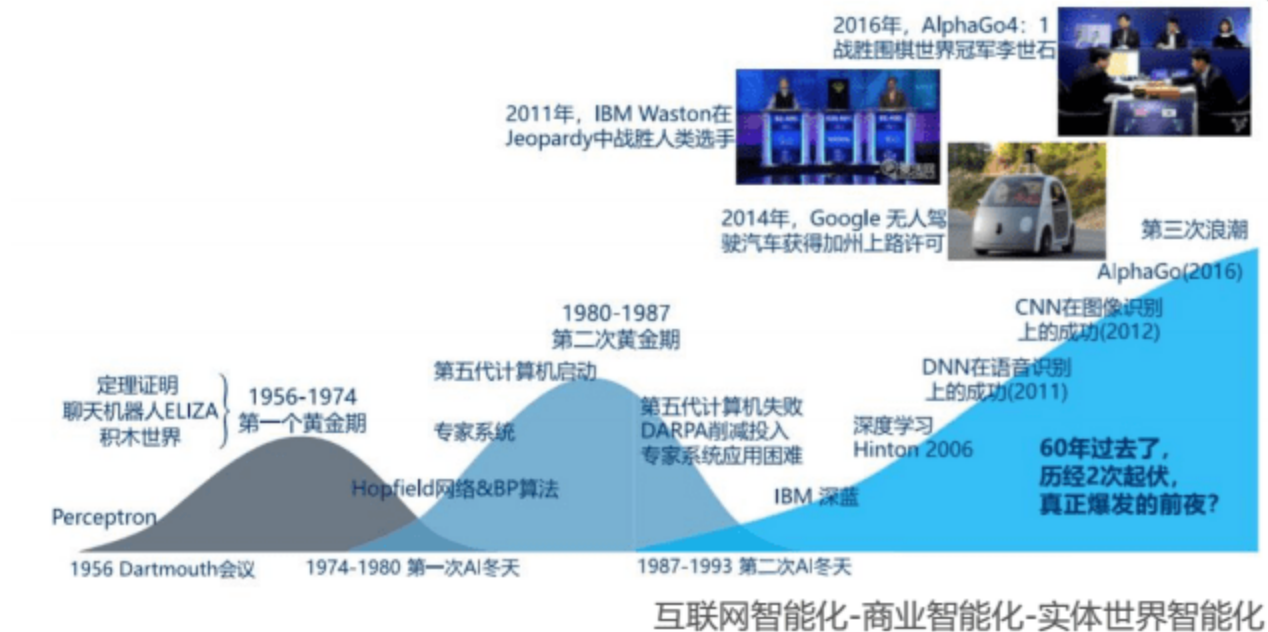
3. 嵌入式AI技术实现

原创力文档

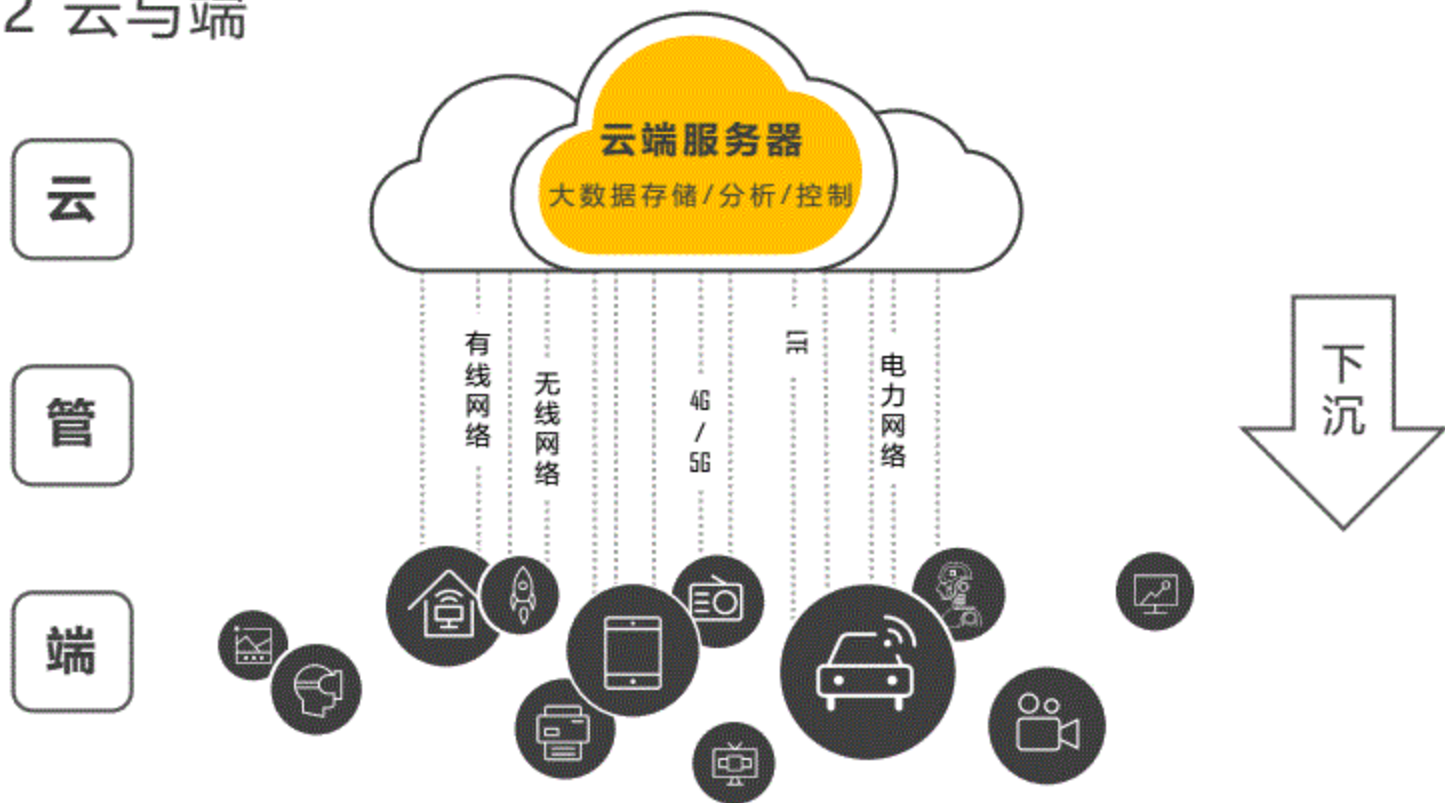
max.book118.com

预览与源文档一致, 下载高清无水印

1.1 AI第三次浪潮



1.2 云与端



1.3 嵌入式AI驱动力

数据：大爆炸

1. 物联网设备量
2013年：100亿台
2020年：500亿台
增长5倍
2. 物理网数据量
2013年：0.1ZB
2020年：44ZB
增长近500倍

算法：普惠化

1. 识别率：AI超越人类
(语音/图像识别准确率超过人类平均水平)
2. 开源：(GitHub: AI换脸)
3. 自动化：(Google: AutoML)

算力：供应丰富，性价比提升

1. 效率升级：CPU-DSP-NPU
2. 供应丰富：
 - AI公司 (寒武纪/地平线)
 - 传统IP Vendor
(ARM/IMG/Synopsys/Cadence/Verisilcon)
 - 系统厂商&互联网公司 (海思、BAT)

嵌入式AI的春天，才刚刚开始

CONTENTS

1. 嵌入式AI发展现状

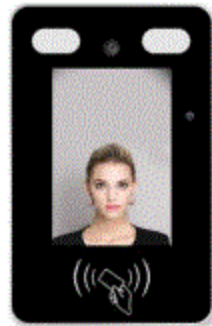
2. 嵌入式AI典型产品

3. 嵌入式AI技术实现

2.1 嵌入式AI典型产品

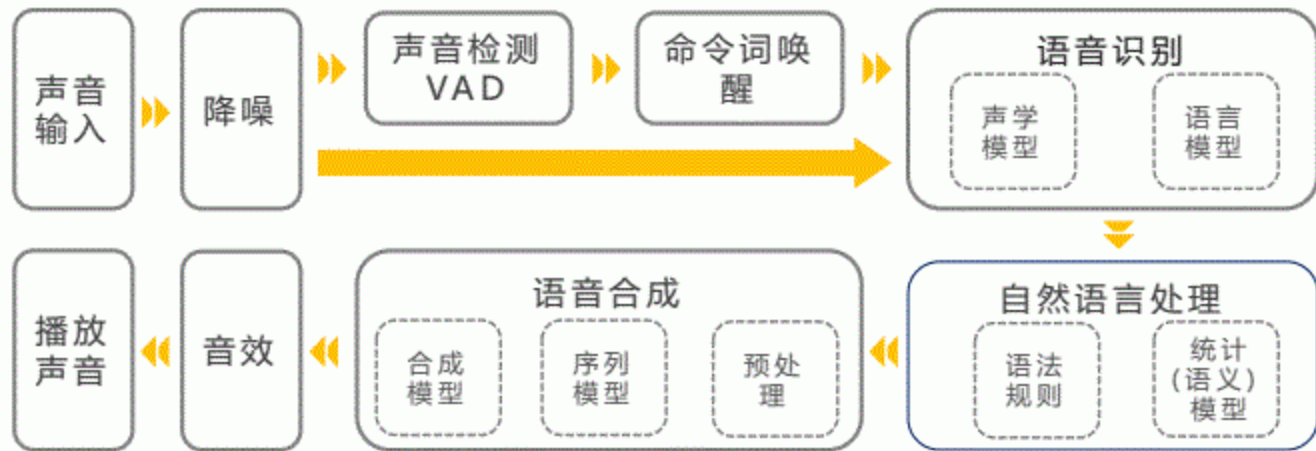


语音



视觉
(人脸/人型/物体)

2.2 产品 workflow-语音

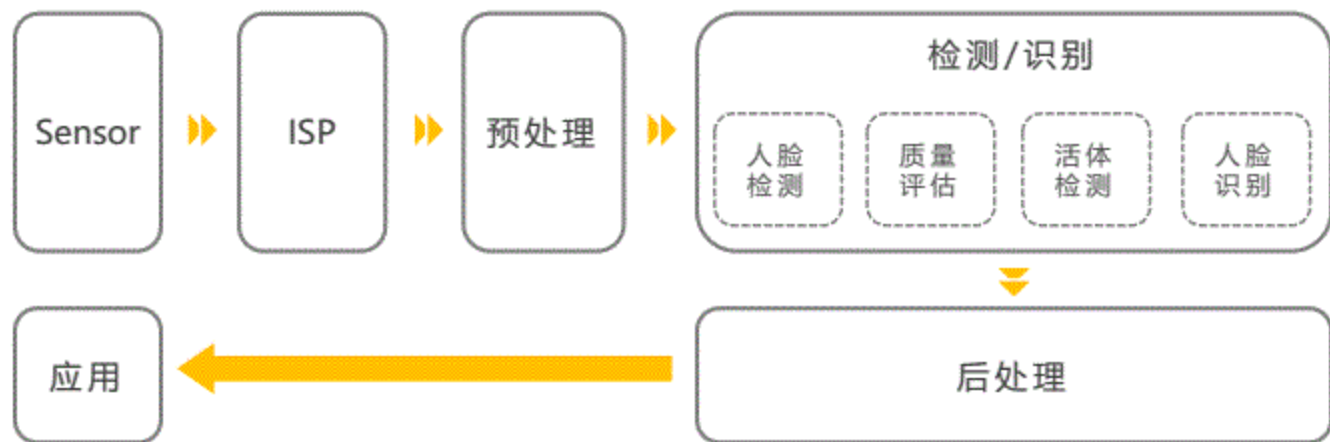


原创力文档

max.book118.com

预览与源文档一致, 下载高清无水印

2.2 产品 workflow-视觉



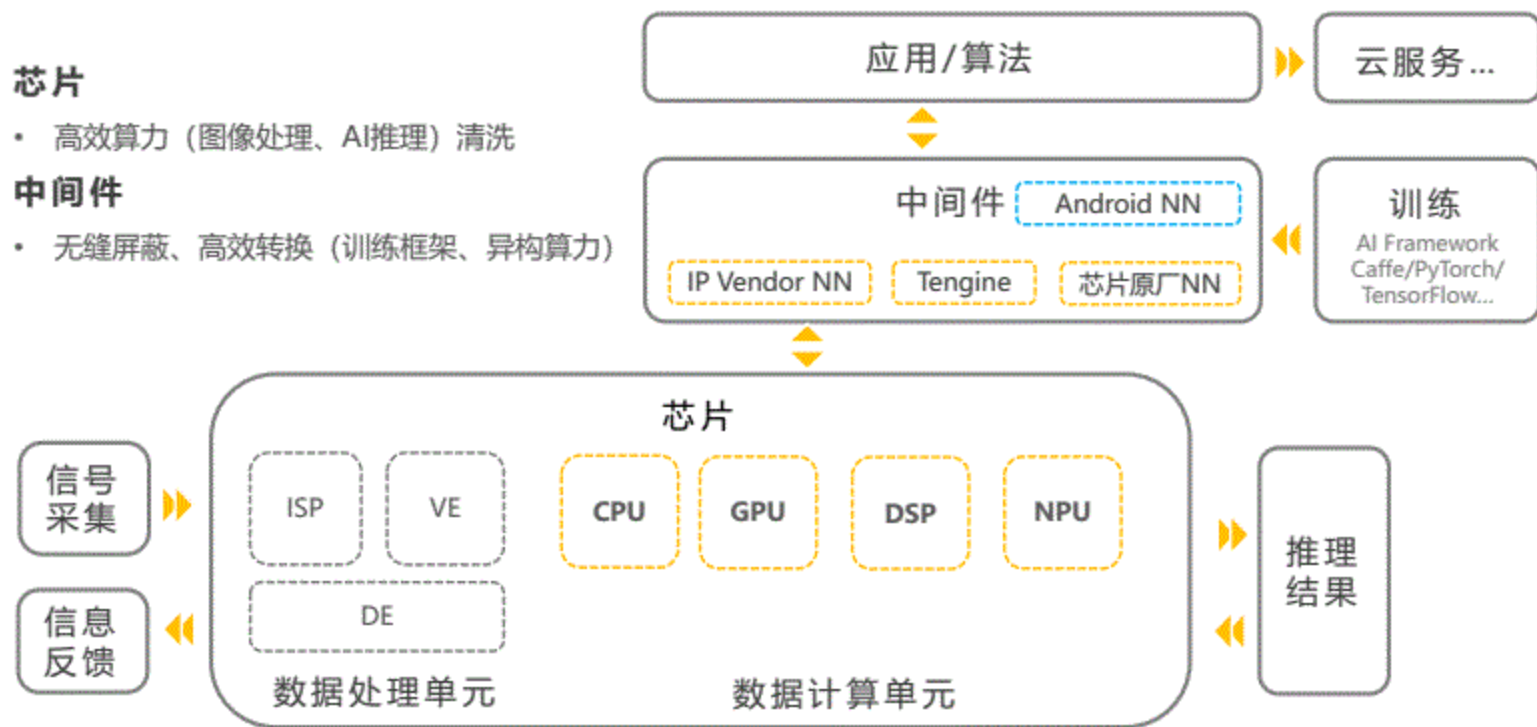
2.3 嵌入式AI技术架构（视觉）

芯片

- 高效算力（图像处理、AI推理）清洗

中间件

- 无缝屏蔽、高效转换（训练框架、异构算力）



CONTENTS

1. 嵌入式AI发展现状
2. 嵌入式AI典型产品
- 3. 嵌入式AI技术实现**

原创力文档

max.book118.com

预览与源文档一致, 下载高清无水印

3.1 算法

算法

算法架构

- 应用场景（室内vs户外、配合式vs非配合式）
- 算法通路（检测、质量评估、活体检测、识别）
- 模型选择（速度、精度、资源的平衡）

数据准备

- 采集
- 清洗
- 标注
- 预处理

模型训练

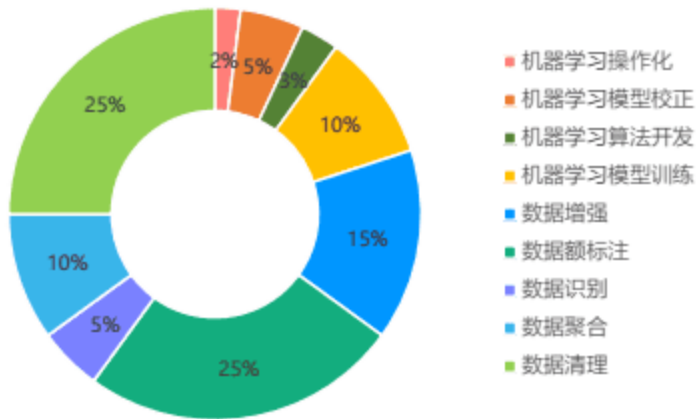
- Benchmark 设定
- 训练框架选择：TF/PyTorch/Caffe/MXNet
- 预训练与调优
- 测试（光线、遮挡、角度、姿态）

芯片

方案

3.1 算法

机器学习中不同任务的分配



AI算法任务 80%的时间花在数据准备



数据的均衡性很重要:

有色人种被错误匹配的比例显著过高。近40%的错误匹配来自有色人种，尽管他们只占国会议员的20%。

3.2 芯片

算法

芯片选择

- AI计算能力：CPU、DSP、NPU
- 图像处理能力：ISP、图像加速器
- 系统资源：带宽、内存、Storage

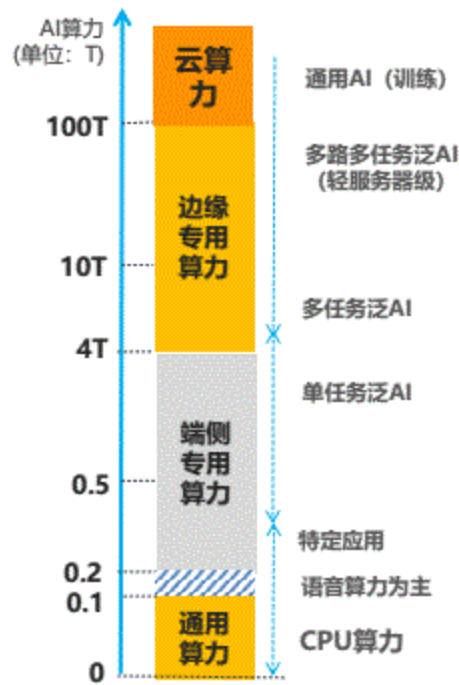
芯片

软件优化

- 异构计算：指令加速(加速lib: NCNN/MNN/Tengine)、多核多线程 (OPENMP) 、异构算力分配
- 图像处理：裁剪、旋转、畸变处理等数据流优化
- 系统工具：模型转换支持 (TensorFlow、PyTorch、Caffe) 、权重量化压缩 (32bit-8bit)

方案

3.2 芯片-AI算力



AI算力分析

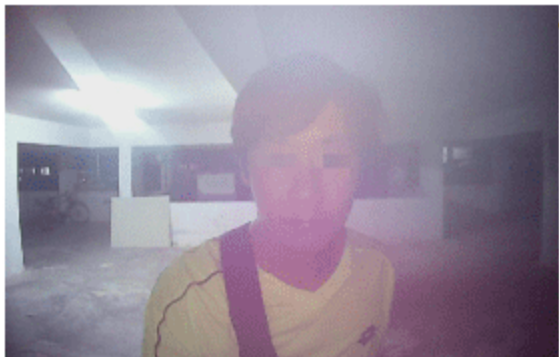
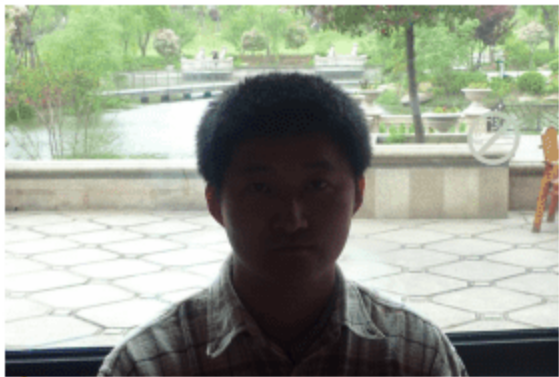
算力 (TOPS)	计算单元	典型场景问题
>100	NPU	模型训练, 视频结构化, 城市大脑 (大数据分析)
10~100	NPU	多路多目标视频结构化, 多路视频语义理解, 大数据分析
4~10	NPU	单路多目标视频结构化, 单路视频语义理解, 视频像素分割
0.5~4	NPU	实时万人级人脸识别, 精准图像物体分割, 实时 (>25fps) 目标检测/分类
0.2~0.5	NPU	准实时千人级人脸识别, 低帧率 (<10fps) 目标检测/分类
0.1~0.2	CPU/DSP/NPU	配合式秒级人脸识别, 语音识别, 语音合成, 特定场景语义理解 (5万词, 百毫秒级) 说明: 算力交叠区
0~0.1	CPU/DSP	配合式秒级人脸识别, 语音识别, 语音合成, 特定场景语义理解 (5万词, 接近秒级)

3.2 芯片-图像



理想

3.2 芯片-图像



现实

3.2 芯片-图像



Before ISP



After ISP

另：图像数据流通路软硬件优化（配合场景、算法裁剪数据流；使用专用图像处理加速器）

3.3 方案

算法

芯片

方案

方案

1. Sensor：选型（视角/WDR/成本）

- 大视角-距离近，多人脸；（小视角-距离远，单人脸）
- 全光线-大靶面和宽动态sensor，保证暗光和逆光等；（室内-普通sensor）

2. 补光设计：环形补光灯、回避光污染

3. 安装：室外光线干扰等



小 结

1. 数据驱动AI
2. 嵌入式AI产品和技术核心主线：数据流
3. 算法 关键技术：数据准备度
4. 芯片&方案 关键技术：数据质量/数据处理/数据计算
 - 数据质量：业务设计、Sensor、ISP
 - 数据处理：图像处理通路（硬件加速，软件优化）
 - 数据计算：CPU/DSP/NPU异构计算

谢谢