

## Gennie v1.5

### 概述

Gennie 是格蠹科技推出的轻量级深度神经网络高速推理引擎，具有高速、异构、跨平台、运行时简单易用等优点。

Gennie 支持 Caffe、TensorFlow、Pytorch 等深度学习框架训练的模型，可以在 Windows、Linux、Android、MacOS 等操作系统下运行。

Gennie 能异构调度通用 CPU(x86、x64、ARM)、GPU (NVIDIA, AMD, INTEL, ARM), NPU, FTPA 等计算单元，充分利用硬件算力，提高推理速度。

### 适用场景

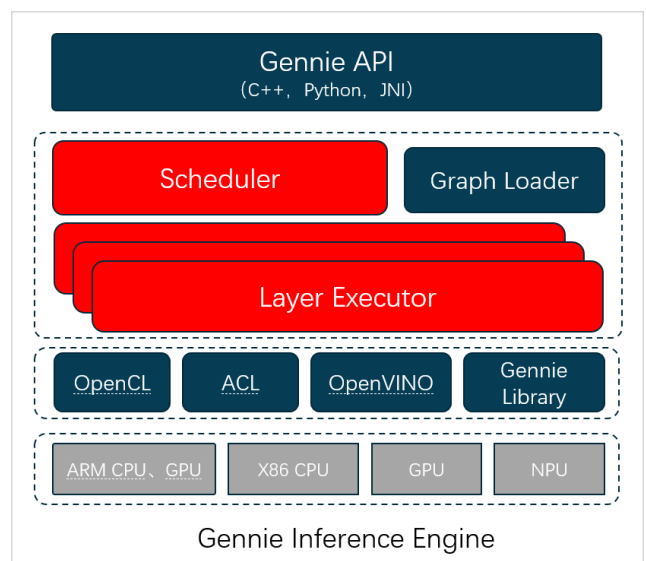
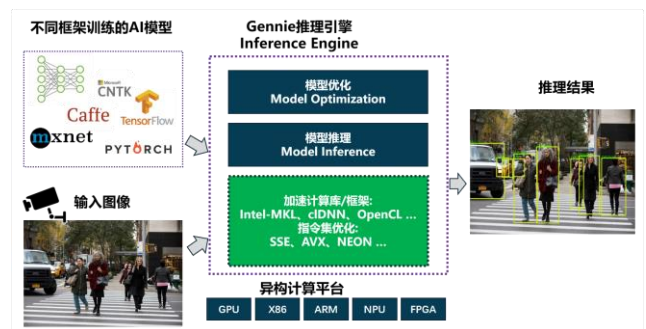
| 场景    | 场景说明                                     |
|-------|--|
| 安防监控  | 基于摄像头视频流的人脸识别，人脸认证，物体识别，物体检测，入侵检测等应用     |
| 移动应用  | 手机中的人脸解锁，AI 拍照，图像风格化等应用，以及智能穿戴设备中的 AI 应用 |
| 交通/汽车 | 自动驾驶和智慧交通中的 AI 应用，如车流和人流监测，目标检测，司机疲劳检测等  |
| 无人飞行器 | 无人机快递，无人机搜索和救援，目标跟踪和识别，故障检修等             |
| 游戏    | 游戏中的人脸识别，手势动作识别，人工智能游戏，AR，VR 游戏等         |
| 智能家居  | 基于人脸识别的智能门锁，智能音箱，陪伴机器人，智能摄像头等智能家居应用      |

### 系统架构

Gennie 由 Gennie Inference Engine、Gennie Model Transition Tool 构成。

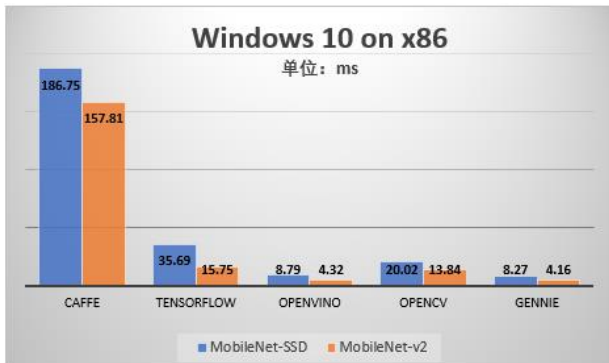
Gennie Inference Engine 即 Gennie 轻量级高速推理引擎。

Gennie 推理引擎引擎使用自有模型模型 (.gmn 后缀)。Gennie 模型具有体积小、安全性高的特点。Gennie Model Transition Tool 用于把 Tensor Flow, Caffe, PyTorch 等框架训练的模型转换成加密的 Gennie 模型。

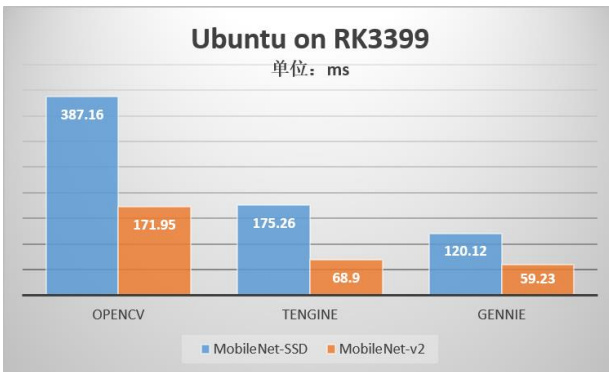


## 性能对比

### Win10 on X86



### Ubuntu on ARM



## 兼容性

### 支持的层

|                  |                 |            |
|------------------|-----------------|------------|
| BatchNorm        | Fully-Connected | Reorg      |
| Concat           | LRN             | ReLU       |
| ConstOp          | Input-op        | Reshape    |
| Convolution      | Normalize       | Resize     |
| Deconvolution    | Permute         | RoiPooling |
| Detection-Output | Pooling         | RPN        |
| Dropout          | Priorbox        | Scale      |
| Eltwise          | PReLU           | Slice      |
| Flattern         | Region          | Softmax    |

### 支持的网络

Gennie 支持几乎所有常见网络，如 AlexNet、ResNet、SqueezeNet、GoogleNet、MobileNet、MobileNet SSD、Inception、DenseNet、ShuffleNet 以及所有由 Gennie 支持的层构成的自定义网络。

### 支持的操作系统

| 支持操作系统                  | 版本                          |
|-------------------------|-----------------------------|
| Windows                 | Windows 7、8、10              |
| Linux (Ubuntu / CentOS) | CentOS 7+<br>Ubuntu 16.04 + |
| MacOS                   | 10.8 +                      |
| Android                 | Android 8.0+                |

## 产品优势

| 优势点         | 优势点描述   |
|-------------|---|
| 推理速度快       | 异构推理智能调度资源，提高资源利用率；<br>汇编级代码开发调优<br>可能是市场上最快的 AI 推理引擎     |
| 兼容性高        | 支持最常见的操作系统、硬件平台与推理框架训练的模型                                 |
| 接口简单        | 只需要调用 3 个函数即可以通过 Gennie 实现 AI 推理                          |
| 运行期占用资源少    | 运行期占用内存数不到 OpenCV 的 50%                                   |
| Footprint 小 | Footprint 只有 800K，不需要其他依赖即可完成 AI 推理                       |
| 领先的技术支持能力   | 由具有 5 年以上经验的技术专家提供专业支持。<br>除了推理引擎本身可以提供模型优化、代码调优等一系列技术服务。 |