# 1.DeepStream介绍

高效、实时分析来自相机、传感器、IoT网关等的数据。

NVIDIA的DeepStream SDK提供了一整套数据流分析工具包，透过智能视频分析（IVA）和多传感器的数据处理来感知情景和意识。

DeepStream应用程序框架具有硬件加速构建块，可将深层神经网络和其他复杂处理任务带入流处理管道。开发者只需专注于构建核心深度学习网络和IP，而不是从头开始设计端到端解决方案。

可以节省计算机资源，一个服务器可能只能接10路视频，需要使用12G左右的显存，如果使用DeepStream，可能只需要2G的显存。

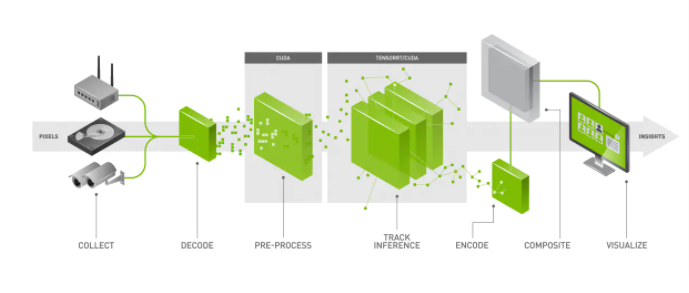
安装包：https://developer.nvidia.com/deepstream-getting-started

## 1.1工作流程

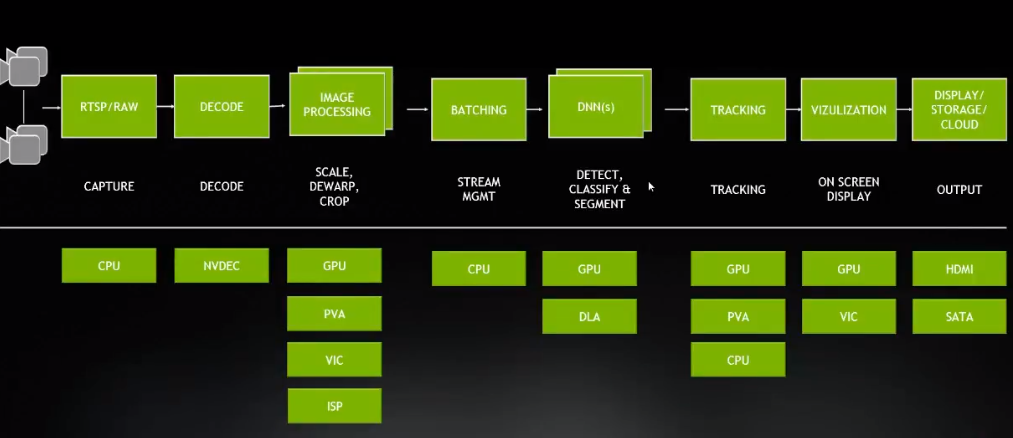
Deep Stream 实时视频流分析即结构化包含：实时视频解码和神经网络进行推理。

解码：由多个线程并行执行并将各种输入流输送到GPU HW硬件解码器;

推理：由一个主线程通过调用TensorRT推理引擎来处理所有批量推理任务。其中，插件系统允许用户将更复杂的工作流添加到流水线中。

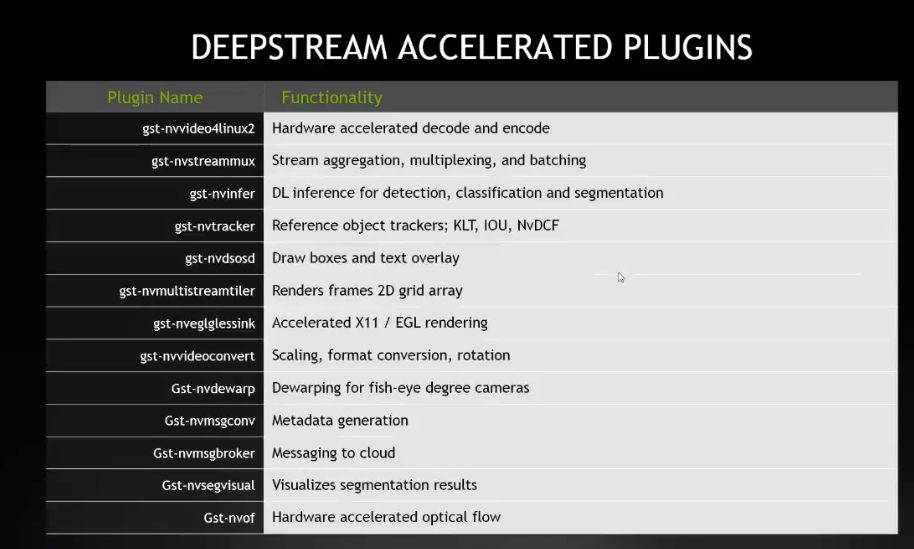


**Figure 1.Deepstream 工作流程**



**Figure 2.Deepstream 整体架构**

## 1.2 NV提供的组件



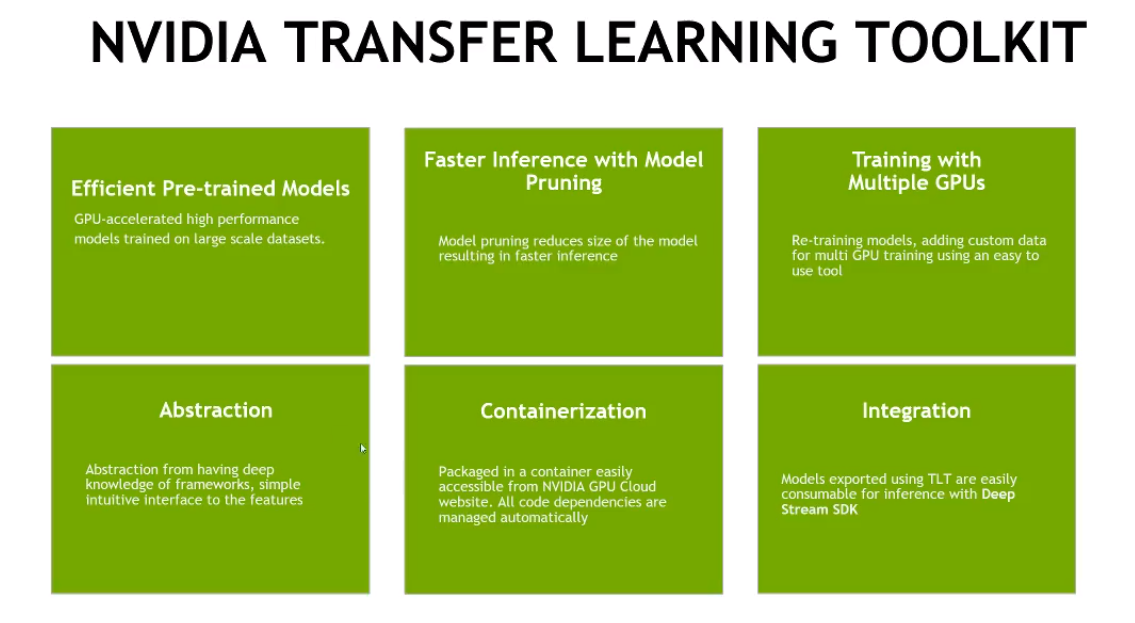
**Figure 3.NVIDIA提供的组件**

## 1.3 部署方式

如果pipeline跟整体APP比较类似，只需要更改config，替换model或者相关设置。位置：samples/configs/deepstream-app/

## 1.4 NVDIA TRANSFER LEARNING TOOLKIT (TLT)

一个Python工具包，涵盖了在视觉领域深度学习中开发的中的一个完整的线。传统开发过程：写脚本、训练、自定义一些内容。使用这个工具大大简化了工作流程。提供了大量的与预训练模型。可以直接生成一个tensert的推理引擎，生成的文件可以直接放到Deep Stream上利用。



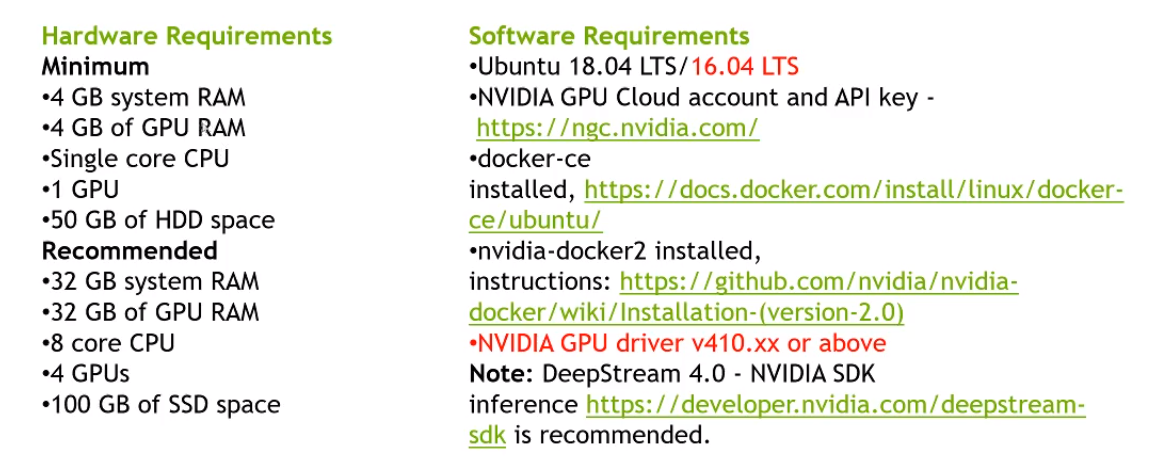
**Figure 4. TLT核心功能**

在指定的公共数据集上训练的图像分类和目标检测模型，可与Transfer Learning Toolkit一起使用。



**Figure 5.TLT支持的模型**

### 1.4.1 在服务器上部署TRANSFER LEARNING TOOLKIT



**Figure 6.部署TLT需要的环境**

步骤1：安装docker(https://www.docker.com/)

步骤2：安装英伟达GPU驱动V410以上

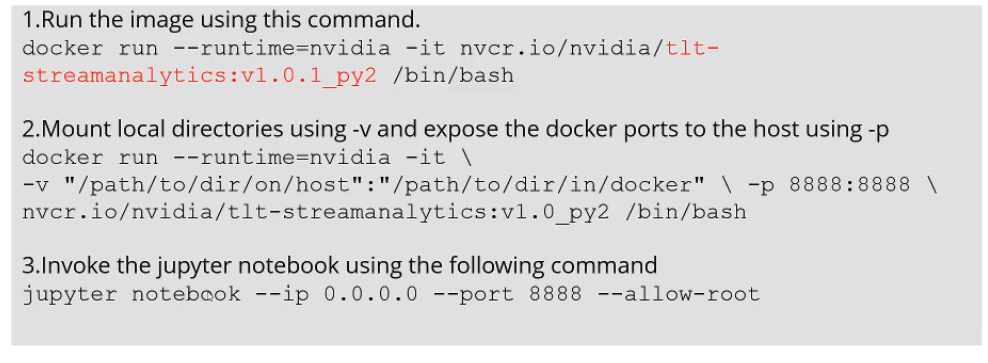
(https://www.nvidia.com/Download/index.aspx?lang=en-us)

步骤3：安装docker2(https://github.com/NVIDIA/nvidia-docker)

步骤4：得到NGC API key(https://ngc.nvidia.com/)：Sigin in -> setup -> Get API key ->

Gennerate API key

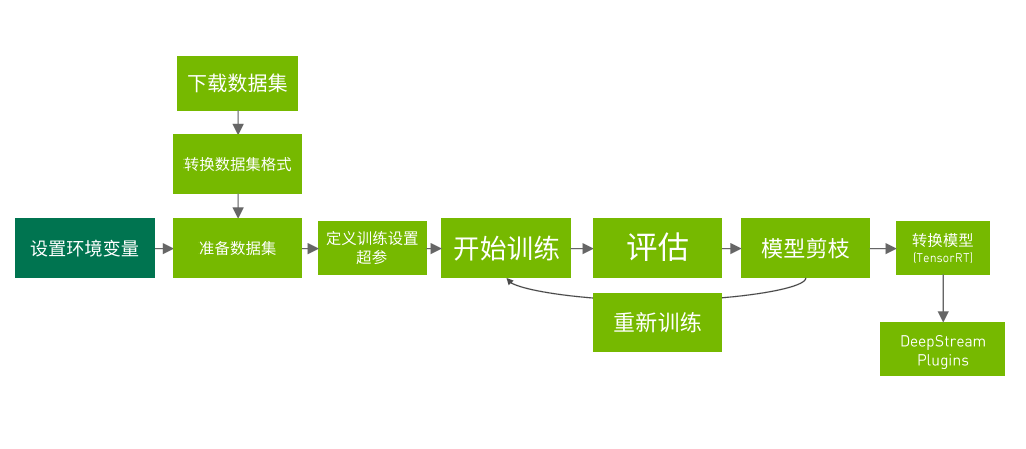
步骤5：在服务器上运行TLT镜像



### 1.4.2 使用NVDIA TRANSFER LEARNING TOOLKIT

**步骤：**

1. 从NGC（NVIDIA GPU CLOUD）上下载整个工具(docker)。
2. 下载模型，训练新的数据集、更改模型类别。
3. 对模型进行剪枝、再训练。
4. 输出模型
5. 将模型放入Inference工具，包括TensorRT、DeepStream



**Figure 7.TLT工作流程**

# 2相关网站

DEEPSTREAM DOCKER CONTAINER FOR GPU

NGC Container：<https://ngc.nvidia.com/catalog/containers/nvidia:deepstream>

NVIDIA Container Runtime：<https://github.com/NVIDIA/nvidia-docker>

DEEPSTREAM DOCKER CONTAINER FOR JETSON

NGC Container：<https://ngc.nvidia.com/catalog/containers/nvidia:deepstream-l4t>

NVIDIA Container Runtime：<https://github.com/NVIDIA/nvidia-docker/wiki/NVIDIA-Container-Runtime-on-Jetson>

Download：<https://developer.nvidia.com/deepstream-sdk>

Forums：<https://devtalk.nvidia.com/default/board/209/>

<https://forums.developer.nvidia.com/t/deepstream-sdk-faq/80236>

Github samples：<https://github.com/NVIDIA-AI-IOT/deepstream_4.x_apps>

<https://github.com/NVIDIA/TensorRT>

TRANSFER LEARNING TOOLKIT：<https://developer.nvidia.com/transfer-learning-toolkit>

<https://devtalk.nvidia.com/default/board/417/transfer-learning-toolkit/>