REAL-TIME OBJECT DETECTION FOR DISASTER RESPONSE

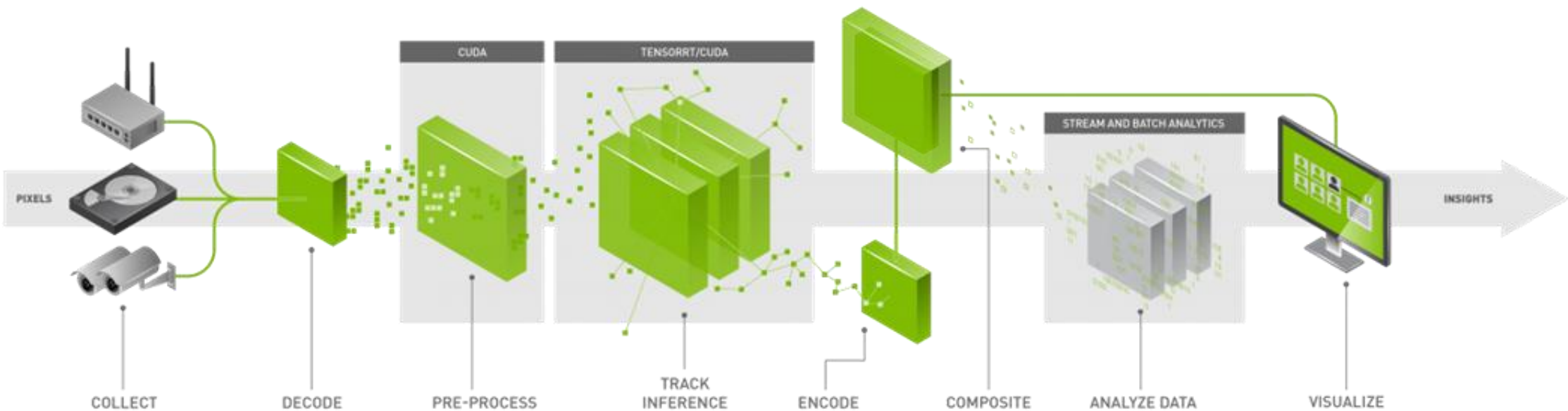
James Skinner, NVIDIA Solutions Architect

WEBINAR AGENDA

- Motivation: Finding people after a disaster
- Fast Intelligent Video Analytics (IVA) inference using **NVIDIA DeepStream 4.0**
- Introducing **Transfer Learning Toolkit (TLT) for Intelligent Video Analytics**
- Training an object detector using TLT
- Deploying to Jetson AGX Xavier using DeepStream

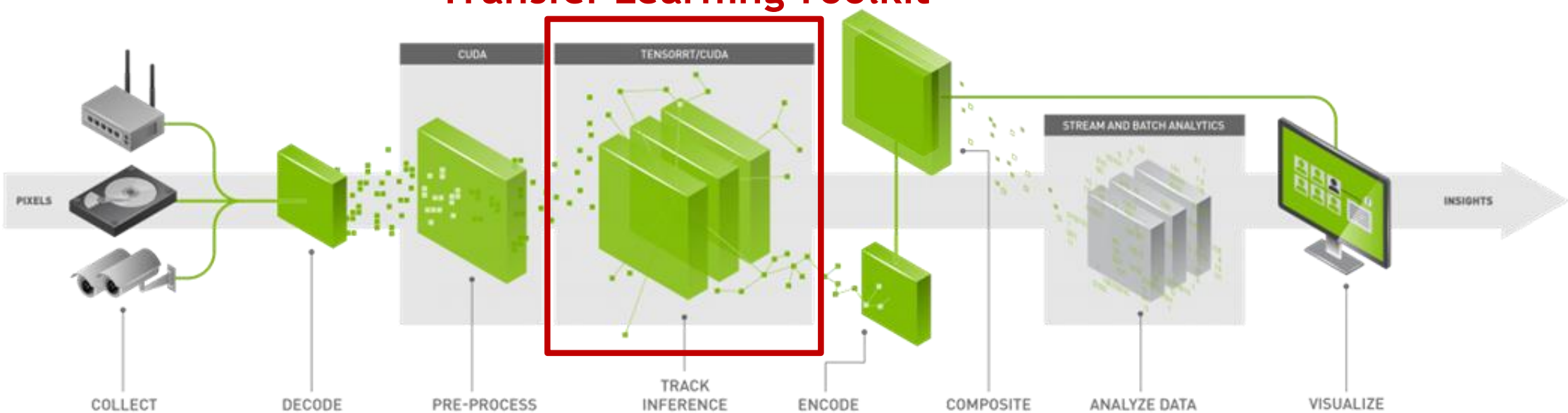
New Orleans (Aug. 30, 2005) U.S. Coast Guard Petty Officer 2nd Class Shawn Beaty looks for survivors in the wake of Hurricane Katrina
https://www.navy.mil/view_image.asp?id=27531

FAST IVA INFERENCE USING DEEPSTREAM



FAST IVA INFERENCE USING DEEPSTREAM

We create this model in Transfer Learning Toolkit



NVIDIA TRANSFER LEARNING TOOLKIT FEATURES

Efficient Pre-Trained Models

GPU-accelerated high performance models trained on large scale datasets.

Faster Inference with Model Pruning

Model pruning reduces model size, accelerating inference

Training with Multiple GPUs

Re-training models using multi-GPU training using an easy to use tool

Abstraction

Abstraction of deep learning, using a simple intuitive interface.

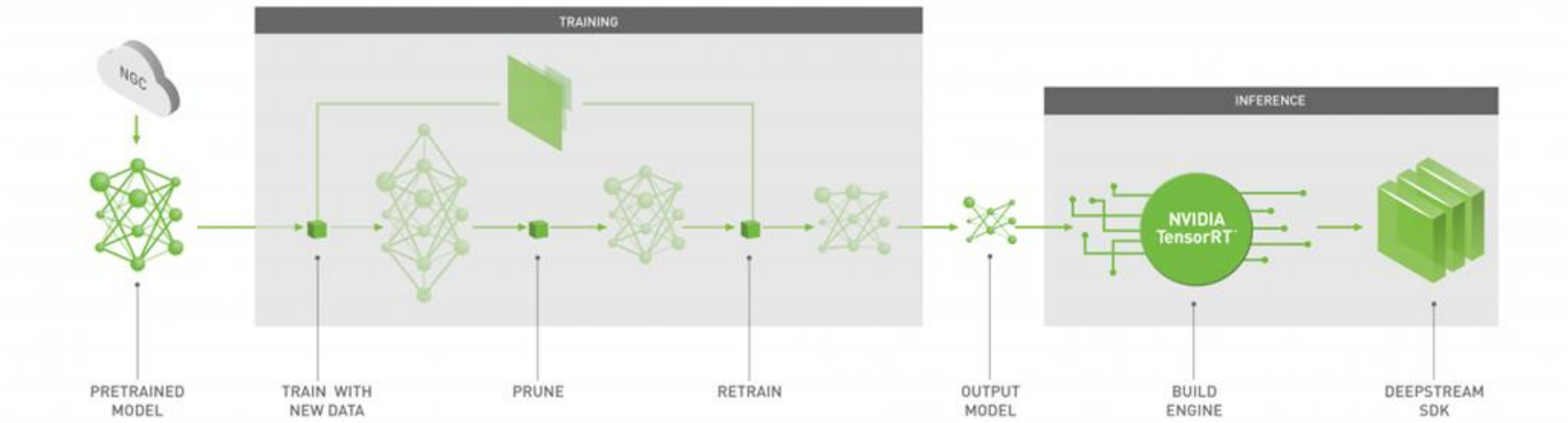
Containerization

Packaged in a container on the NVIDIA GPU Cloud.

Integration

Integration with **DeepStream SDK** simplifies the process of creating IVA applications.

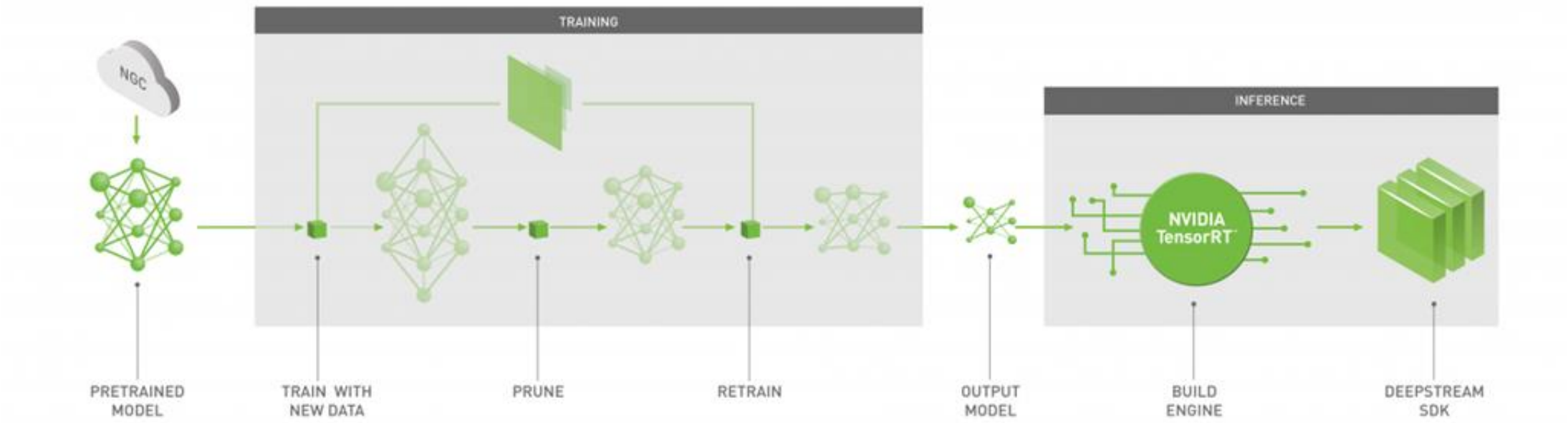
TRANSFER LEARNING TOOLKIT



On Tesla GPU (eg DGX-1, cloud provider)

1. **Download pre-trained model**
2. **Convert data to TFRecords**
3. **Train model on your data**

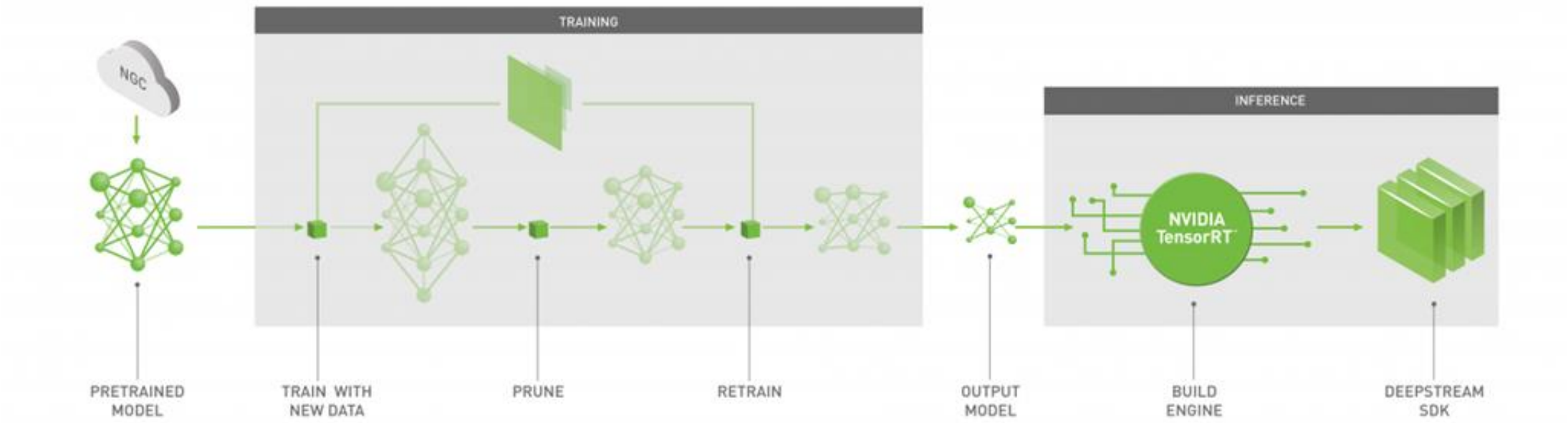
TRANSFER LEARNING TOOLKIT



On Tesla GPU (eg DGX-1, cloud provider)

1. Download pre-trained model
2. Convert data to TFRecords
3. Train model on your data
4. (Prune trained model)
5. (Retrain)
6. Export model

TRANSFER LEARNING TOOLKIT



On Tesla GPU (eg DGX-1, cloud provider)

1. Download pre-trained model
2. Convert data to TFRecords
3. Train model on your data
4. (Prune trained model)
5. (Retrain)
6. Export model

On edge device (eg Jetson AGX Xavier)

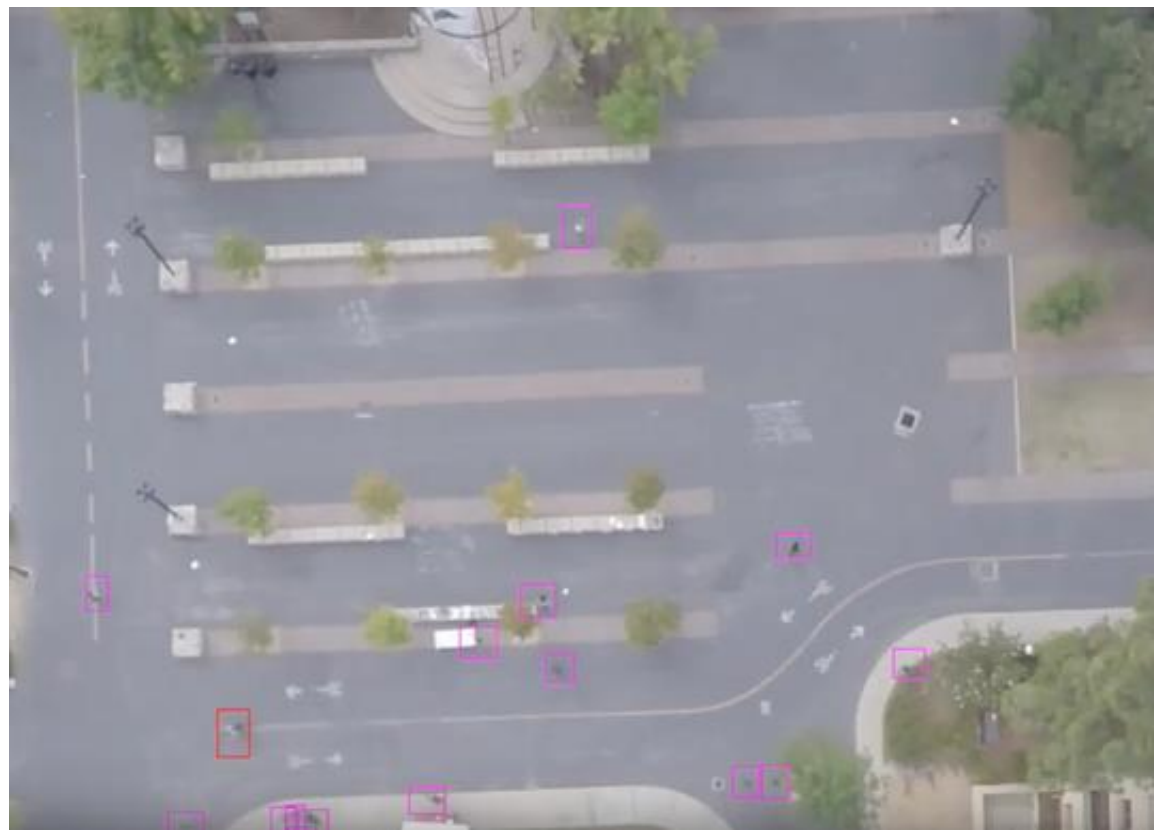
7. (Build TensorRT engine)
8. Deploy in DeepStream


USING TLT TO TRAIN A
REAL-TIME DETECTOR

STANFORD DRONES DATASET

http://cvgl.stanford.edu/projects/uav_data/

- 68 videos
- Randomly take frames from each video & crop each to 768 x 768.
- Save data in KITTI format.
- Six classes, including pedestrian.
- You could also fine-tune on a domain specific dataset.



TLT PROCESS IN ONE SLIDE

`ngc registry model download-version`	Download a pre-trained model
`tlt-dataset-convert`	Convert KITTI dataset into TFRecords
`tlt-train`	Train model using your data
`tlt-evaluate / tlt-infer`	Evaluate on val data / infer new images
`tlt-prune`	Prune the model, to reduce no. of params.
`tlt-train`	Retrain, to recover accuracy.
`tlt-export --data_type fp16`	Export your model to `.etlt` format
(Jetson) `tlt-converter – fp16`	Convert to TensorRT engine.


TRAINING ON DGX-1

DEPLOYING ON JETSON AGX XAVIER

JETSON AGX XAVIER DEVELOPER KIT

https://developer.nvidia.com/embedded/jetson-agx-xavier-developer-kit

- Install the latest JetPack (currently 4.2.2)
- Follow the instructions in the TLT Getting Started Guide
  - Download tlt-converter from the NVIDIA dev zone.
  - Install Open SSL:
    `sudo apt-get install libssl-dev`
- Copy your `.etlt` and `.bin` files to your Jetson device.

INFERENCE PERFORMANCE

My model: **DetectNet v2** with **ResNet 50** backbone; 768 x 768 pixel video frames.

Number of streams	Precision	Total FPS
1	FP32	11
1	FP16	40
1	INT8	60
4 (interval + tracker)	INT8	240 (60/stream)
8 (interval + tracker)	INT8	251 (31/stream)

START USING TRANSFER LEARNING TOOLKIT

https://developer.nvidia.com/transfer-learning-toolkit

- Sign up for a free NVIDIA GPU Cloud (NGC) account
- Download the TLT for IVA Docker container
  - `docker pull nvcr.io/nvidia/tlt-streamanalytics:v1.0_py2`
- Train, prune, re-train, export & deploy!
- Let me know which problems you have quickly and accurately solved with the NVIDIA IVA tools!
  - jskinner@nvidia.com

NVIDIA

GTC GPU TECHNOLOGY CONFERENCE

November 4 - 6, 2019 | Washington, D.C.

CONNECT

Connect with experts from NVIDIA, GE Healthcare, NSF Carnegie Mellon, Google, and other leading organizations

LEARN

Gain insight and valuable hands-on training through over 100 sessions

DISCOVER

See how GPU technologies are creating amazing breakthroughs in important fields such as deep learning

INNOVATE

Explore disruptive innovations that can transform your work

Join us at GTC DC | Use VIP code NVJSKINNER for 25% off | Govt attends free

Don’t miss the premier AI conference.

nvidia.com/en-us/gtc-dc/


NVIDIA