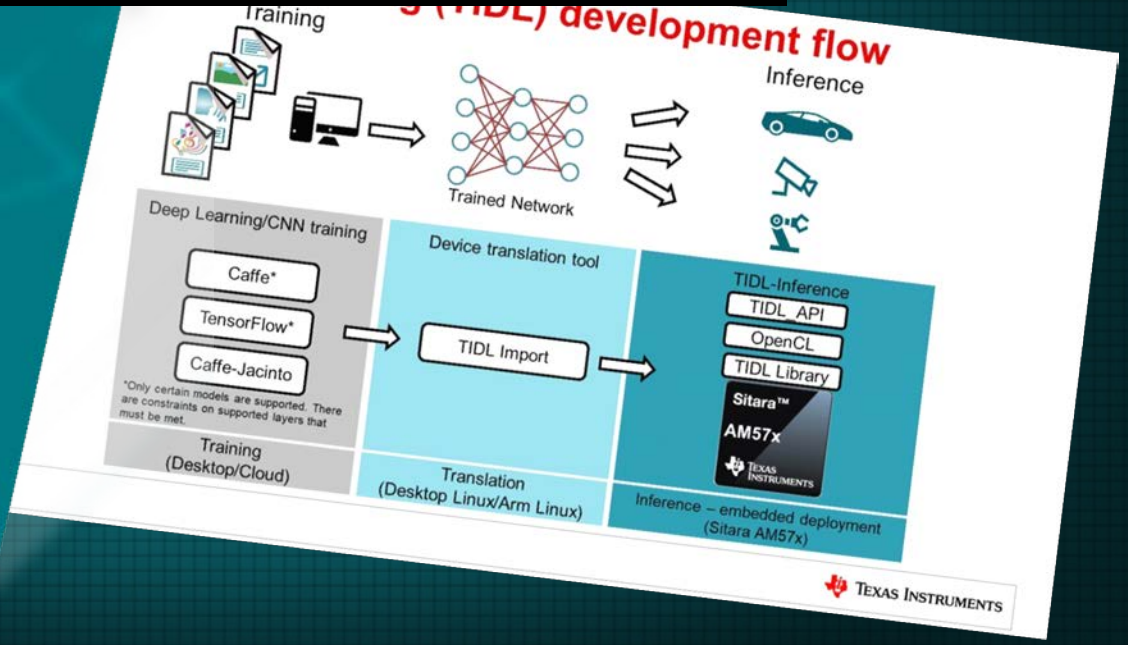
TEXAS INSTRUMENTS DEEP LEARNING (TIDL) FOR SITARA PROCESSORS

OVERVIEW



Texas Instruments Deep Learning (TIDL) for Sitara Processors Overview

Texas Instruments Deep Learning (TIDL) for Sitara Processors: Agenda

- AM57x SoC (System on Chip), AM5749 SoC, and Embedded Vision Engine (EVE) Subsystem
- TI Deep Learning (TIDL) development flow
- TIDL on AM57x SoC
- Validated framework and network models
- TIDL example use cases
- TI Design on TIDL: TIDEP-01004

AM57x Cortex-A15 Processor

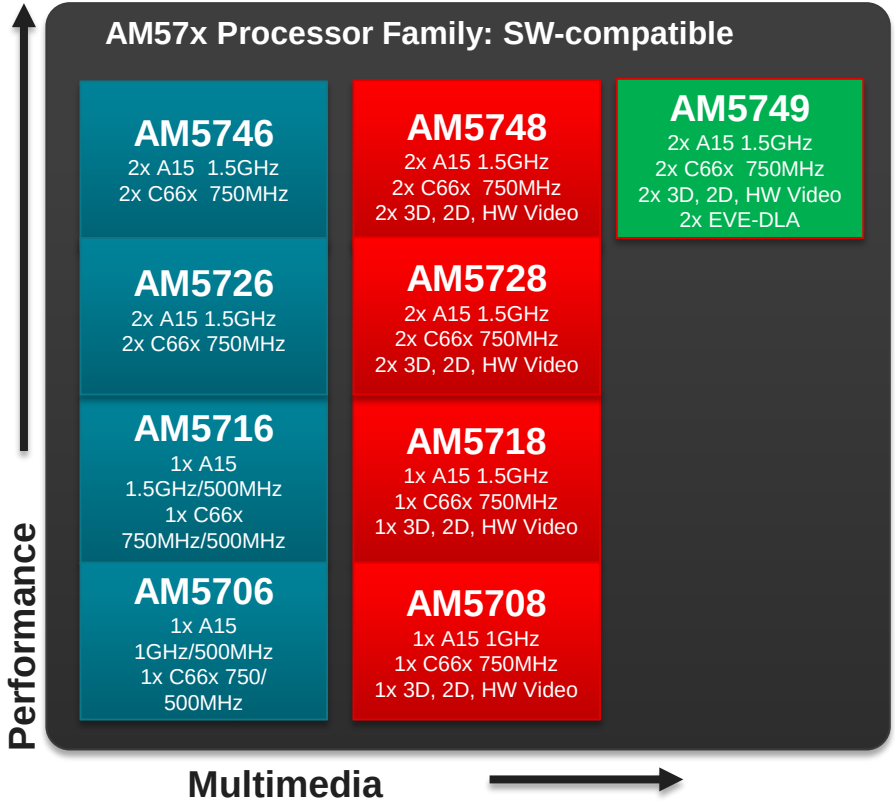
Highest integrated, generally available ARM Cortex-A15 in the market

Benefits:

- Unmatched **performance** in the class
- **Scalable** family with pin-compatible single and dual core devices with single Processor SDK for all Sitara devices

Software & development tools:

- Processor SDK with LTS Linux
- TIDL supported by the SDK
- TI GP EVM and Industrial Development Kit (IDK) available.

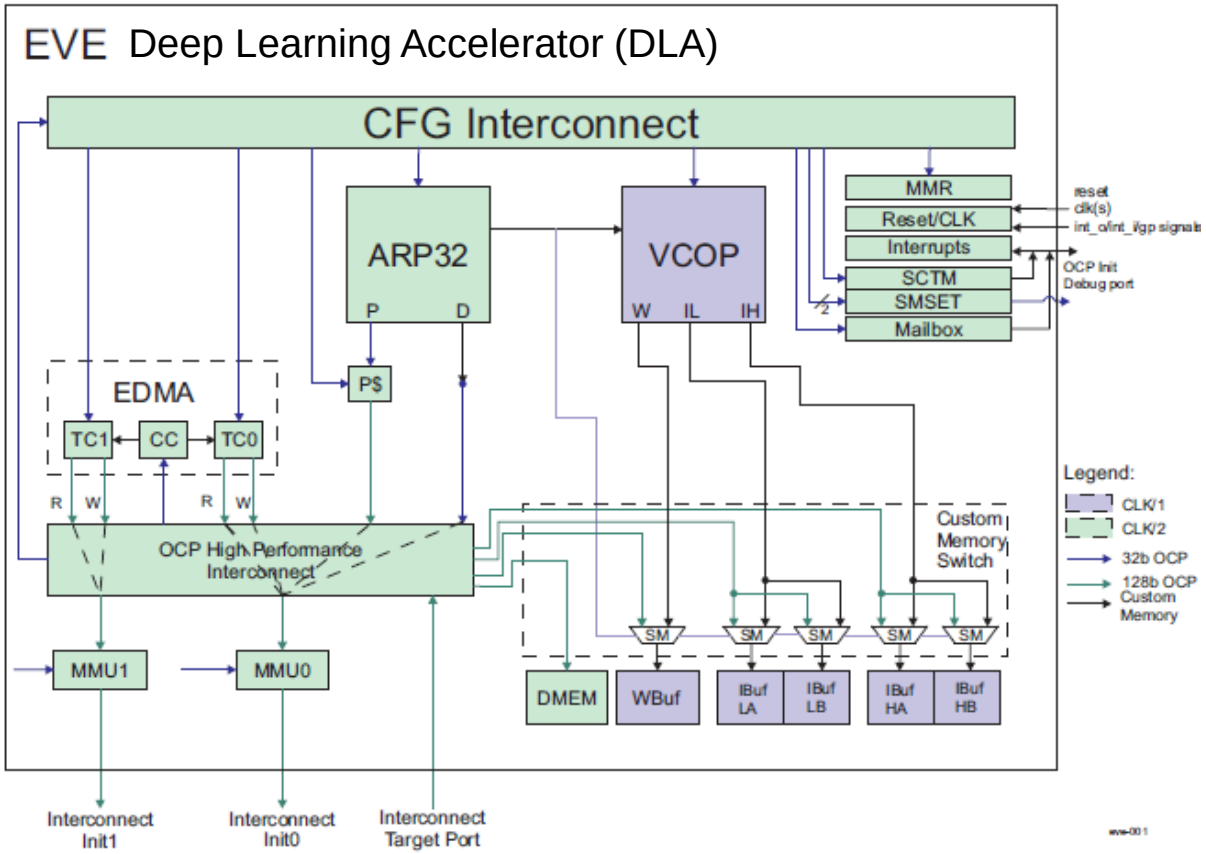


Embedded Vision Engine (EVE)

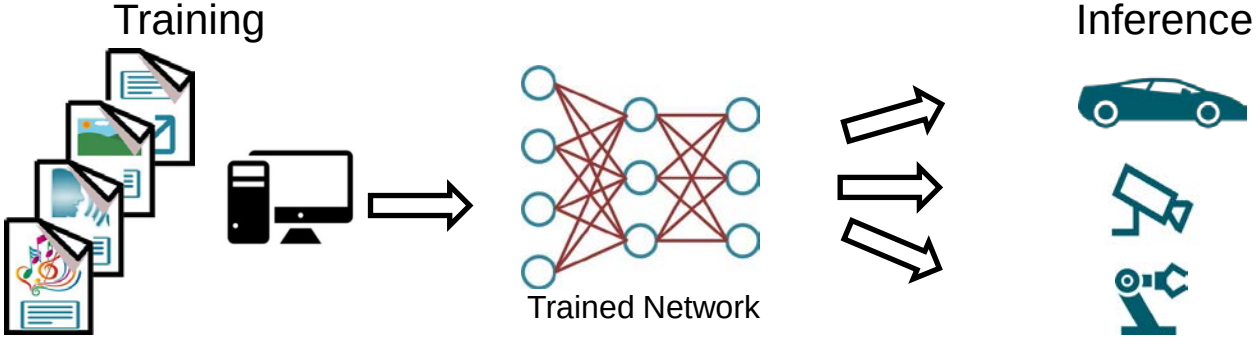
- 32-bit RISC processor (ARP32)
- 512-bit Vector CO-Processor (VCOP)
  - 16x 16-bit MAC/cycle
  - **10.4 GMAC/sec (or 20.8 GOP/sec) per EVE subsystem at 650 MHz**

AM5749
2x A15 1.5GHz
2x C66x 750MHz
2x 3D, 2D, HW Video
2x EVE-DLA

- **AM5749 EVE supports 20.8 GMAC/sec at 650 MHz**



TI Deep Learning (TIDL) development flow



Deep Learning/CNN training	Device translation tool	TIDL-Inference
Caffe* TensorFlow* Caffe-Jacinto *Only certain models are supported. There are constraints on supported layers that must be met.	TIDL Import	TIDL_API OpenCL TIDL Library Sitara™ AM57x Texas Instruments
Training (Desktop/Cloud)	Translation (Desktop Linux/Arm Linux)	Inference – embedded deployment (Sitara AM57x)

TI Deep Learning (TIDL) on AM57x SoC

- TIDL is set of software components that enables a Deep Learning inference at the edge.
- TIDL is available for free as part of Processor SDK Linux.
- TIDL runs on all AM57x devices:
  - Runs on EVE subsystem and/or C66x DSP cores
  - AM5749 is the highest performance TIDL device (2x EVE) currently available from TI.

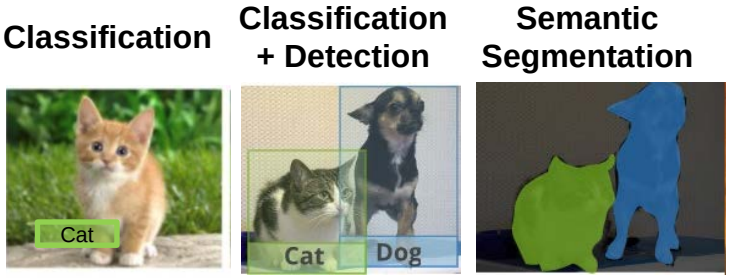
  NOTE: TIDL runs 1.5x-4x faster on EVE subsystems and consumes less power compared to C66x

  Fully loaded 2x EVE consumes 220 mW compared to 2x C66x cores which consumes 520 mW.
- The initial TIDL release:
  - Supports Convolution Neural Networks (CNN) for spatial input data (video, images, ToF, Radar, etc.)
  - Does not support Recurrent Neural Networks (RNN), Long-Short Term Memory (LSTM) and Gated Recurrent Unit (GRU) model used for time-variant data (speech, machine data, etc.)

  NOTE: Support for RNN/LSTM/GRU is planned for future releases
- TIDL APIs are used for programming of EVE and DSP cores.

TI Deep Learning (TIDL) network models

- TIDL provides several TI network models:
  - Object classification
  - Object detection
  - Pixel-level semantic segmentation



- TI's network models and TIDL leverage the following for embedded performance optimization:
  - Efficient CNN configurations/network pruning
  - Sparsity
  - Fixed-point dynamic quantization
- These tools result in major computational and bandwidth reductions with minor decrease in accuracy.

Validated framework and network models

Framework	Network	Comment
BVLC-Caffe *	SqueezeNet 1.1	Supports floating-point and on-the-fly 8-bit quantized model
TensorFlow *	InceptionNet V1 (GoogleNet)	Supports floating-point and on-the-fly 8-bit quantized model
	Mobilenet 1.0	
Caffe-Jacinto	**JacintoNet11**	**Object classification network** 
	JDetNet	**SSD-based object detection** 
	JSegNet21	**Pixel-wise semantic segmentation** 

* Only certain neural network models are supported with BVLC-Caffe and TensorFlow frameworks.
Certain constraints on layer parameters must be met in order for TI device translator tool to work on it.
Details: http://software-dl.ti.com/processor-sdk-linux/esd/docs/latest/linux/Foundational_Components_TIDL.html

TI Deep Learning (TIDL) example use cases

TIDL Library 5.0 and later
Spatial 2D data (RGB sensor, ToF, radar/mmWave, etc.)

Classification
+ Localization


Object detection


Semantic
Segmentation

Example Applications	Role of deep learning
Vision computers, optical inspection, semiconductor test and manufacturing	Classify products as good or defective. Deep learning can improve accuracy and flexibility.
Building automation	Can track/identify/count people and objects
Automated Sorting Equipment, industrial robots	Ability to recognize objects and guide movement/placement
ATMs, currency counters	Determine valid/counterfeit currency
Vacuum robots, robotic lawn mowers	Improve ability to recognize obstacles and things like power cords/wires.
Smart appliances	Recognize objects in refrigerator, determine food & auto cook

TIDL TI Design: TIDEP-01004

Features

- Embedded deep learning inference on AM57x SoC
- Performance scalable TI deep learning library (TIDL library) on AM57x using C66x DSP only, EVE only, or DSP + EVE.
- Performance optimized reference CNN models for object classification, detection and pixel-level semantic segmentation.
- Walk-through of TIDL development flow: training, import and deployment.
- Benchmarks of several popular deep learning networks on AM5749
- This reference design is tested and includes TIDL library on DSP and EVE, reference CNN models and Getting Started Guide.

Applications

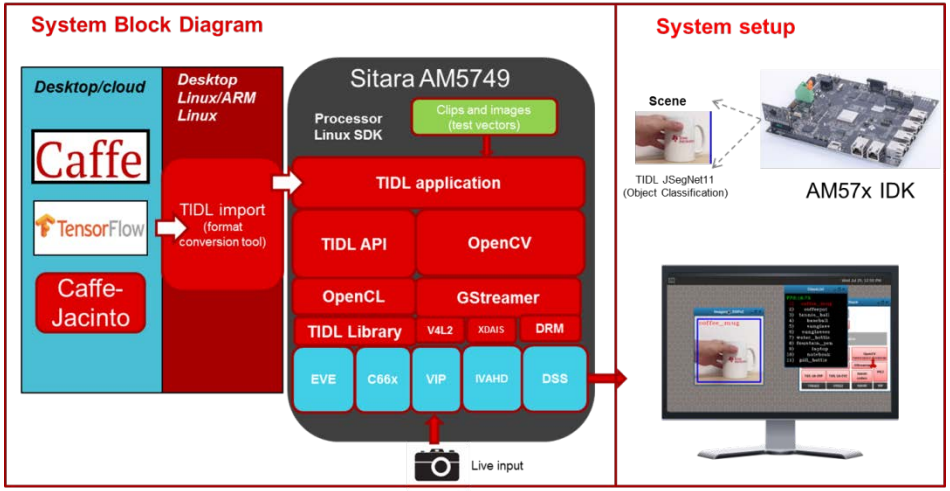
- Machine Vision: Vision Computer, Code Readers
- Automated Machinery: Automated Sorting Equipment, Optical Inspection
- EPOS: ATMs, Currency Counter
- Logistics Robots: Logistics Robots CPU board
- Many others…

Tools & Resources

- **Design Guide**
- **Design Files:** Schematics, BOM, Gerbers, Software, etc.
- **Device Datasheets:**
  - AM5749 IDK (not yet public)
  - AM5749 SoC (not yet public)

Benefits

- Guide through dense and sparse model training, importing and deploying the model for inference run on AM57x SoC.
- Highly integrated solution bringing deep learning inference to the edge.
- Example trained networks ready to run
- TIDL API to run multiple (same or different) networks in parallel on different EVE subsystem and DSP cores.



For more information

- Sitara™ AM57x Processors: http://www.ti.com/am57x
- TI Deep Learning on Processor SDK Linux: http://software-dl.ti.com/processor-sdk-linux/esd/docs/latest/linux/Foundational_Components_TIDL.html
- TIDL API: http://downloads.ti.com/mctools/esd/docs/tidl-api/intro.html
- Caffe-Jacinto training framework: https://github.com/tidsp/caffe-jacinto
- Caffe Models trained by TI: https://github.com/tidsp/caffe-jacinto-models
- For questions regarding topics covered in this training, visit the support forums at the TI E2E Community website.