
CEVA-XM4™

# Efficient Implementation of Neural Networks

**Jeff VanWashenova – Director of Automotive Market Segment**

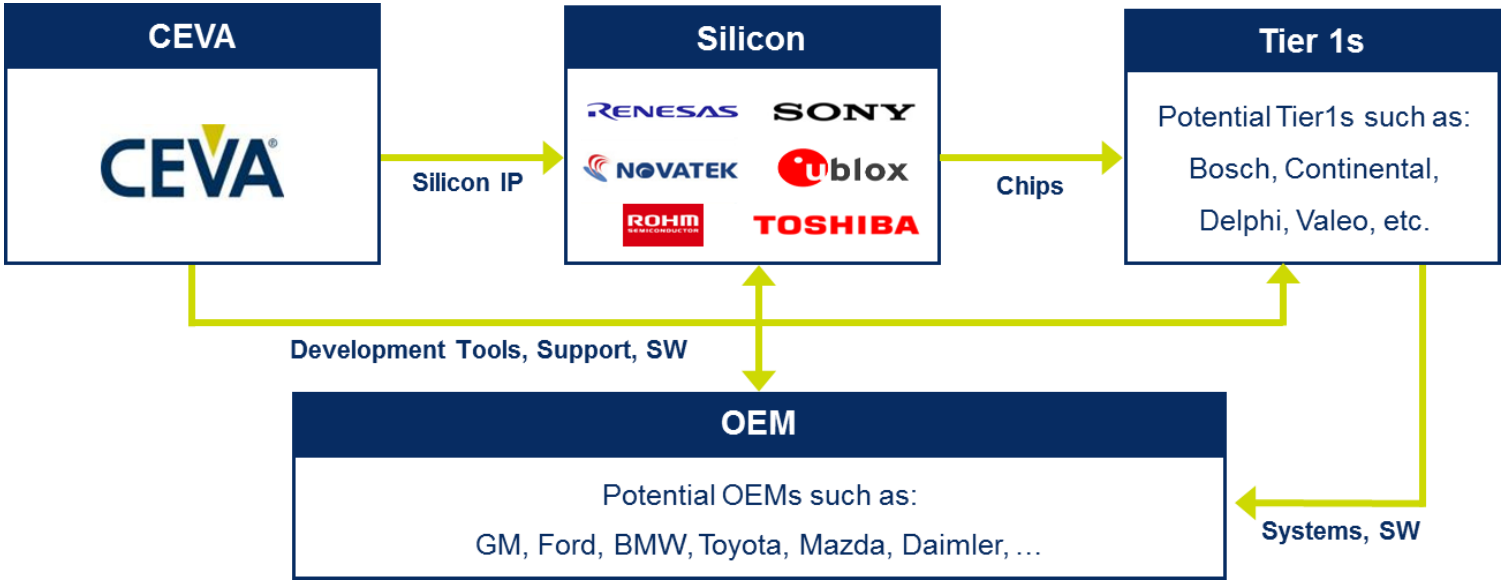
www.ceva-dsp.com

# Presentation Outline

1. CEVA
2. Automotive Trends Challenges in Deep Learning
3. Efficient implementation of Neural Networks
4. Alexnet Example

# CEVA – Ultra Low Power IP

- Image and Vision, Connectivity, Communication
- **+7 Billion Devices**

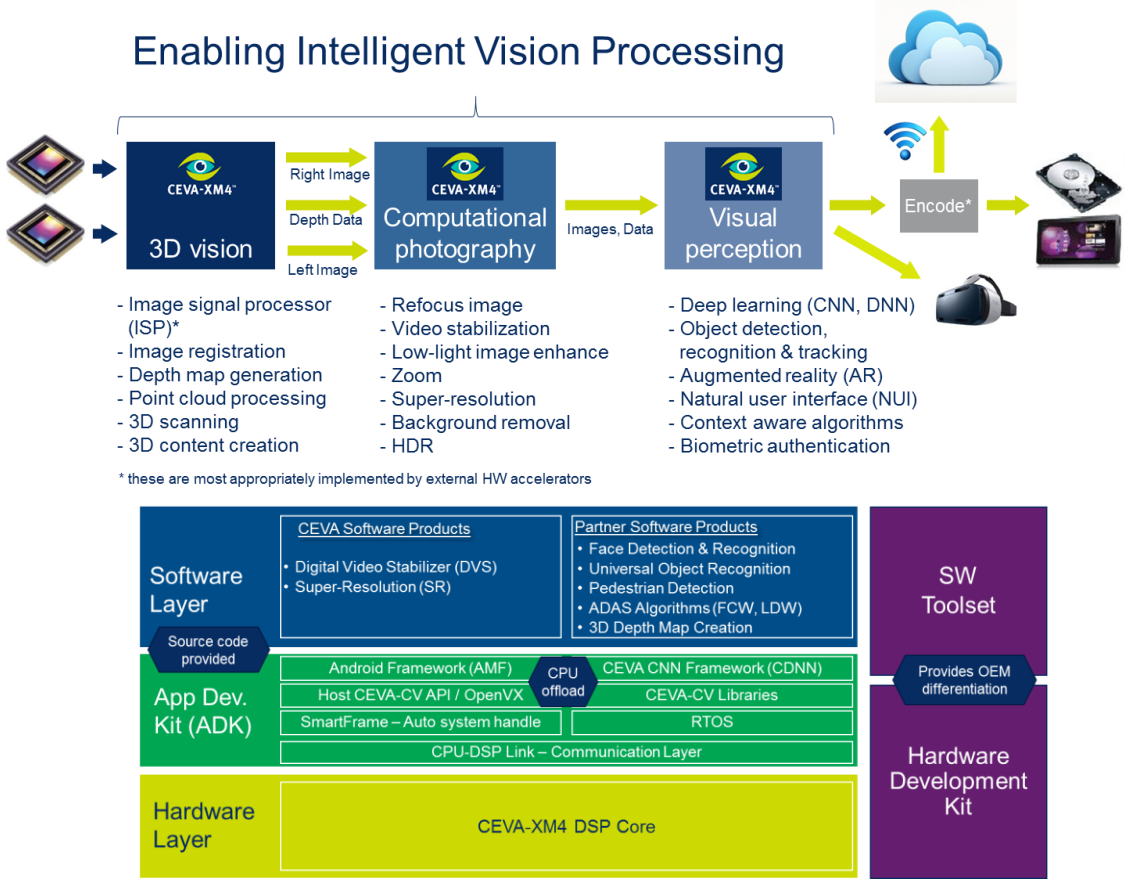



CEVA-XM4™

CEVA-XM4 vision processor wins THE LINLEY GROUP ANALYSTS' CHOICE 2015 BEST PROCESSOR IP

# CEVA Vision Platform

- Platform includes vision processor, libraries, tools and applications (CEVA and a variety of SW partners and service experts)
- **4th-generation imaging and vision processor IP**
  - Vector-type processor; combines fixed- and floating-point math; up to 4096-bit processing per cycle
- XM4:  10+ design wins, Silicon available in Q2/2016
- Platform enables efficient embedded vision applications



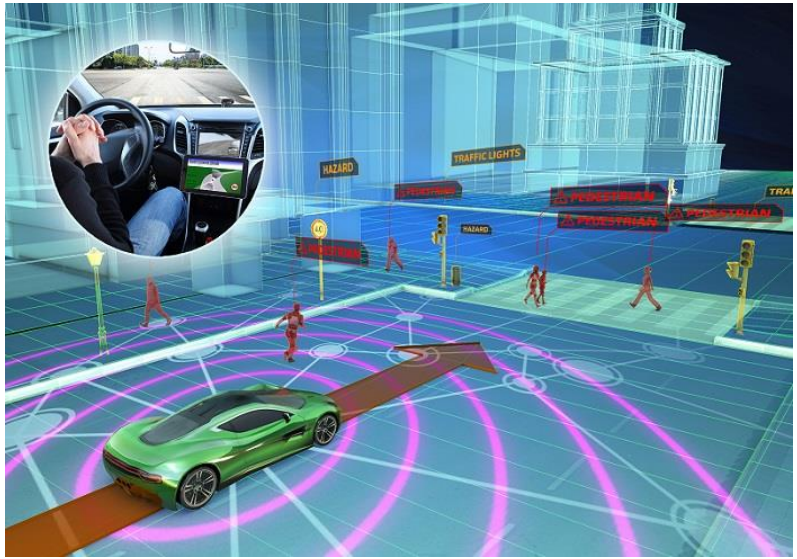
# Presentation Outline

1. CEVA
2. Automotive Trends Challenges
3. Efficient implementation of Neural Networks
4. Alexnet Example

# Deep Learning in Automotive

## High volume vision applications leading to autonomous driving

- Neural Networks
  1. Optimal for harsh automotive environment – Covered signs
  2. Re-trainable without code changes (implement once, use many times)
- Explosion of computational load; fixed power budget
- Challenges
  - **Cost** – Need for cost efficient embedded systems
  - **Power Efficiency** – Fixed power budget – 10 – 15W Centralized
  - **Time to Market** - Migration to production



# Presentation Outline

1. CEVA
2. Automotive Trends Challenges
3. Efficient implementation of Neural Networks
4. Alexnet Example

# Efficient Implementation of Neural Networks

## Cost

- **Specialization**: IP that is optimized for vision and deep learning
- **Die Size:** Specialized core for vision and neural networks are smaller than GPU for best performance/mm$^2$

## Efficiency

- **Utilization** – Dedicated Vision Processor can achieve higher utilization of core
  - > 95% utilization improves **Performance/Watt/**
  - **Floating point to Fixed Point Operations**
  - Fixed point sufficient for CNN

## Time to Market

- **Flexibility:** Quickly adapt to algorithm and network changes
- **Qualifications:** Select designs with need safety packages (ISO26262)
- **Software Libraries –** Quick development and deployment of common vision/CNN functions
- **Software Tools** - SW Tools to migrate from R&D to Production quickly

**17x**

Smaller Die Size*

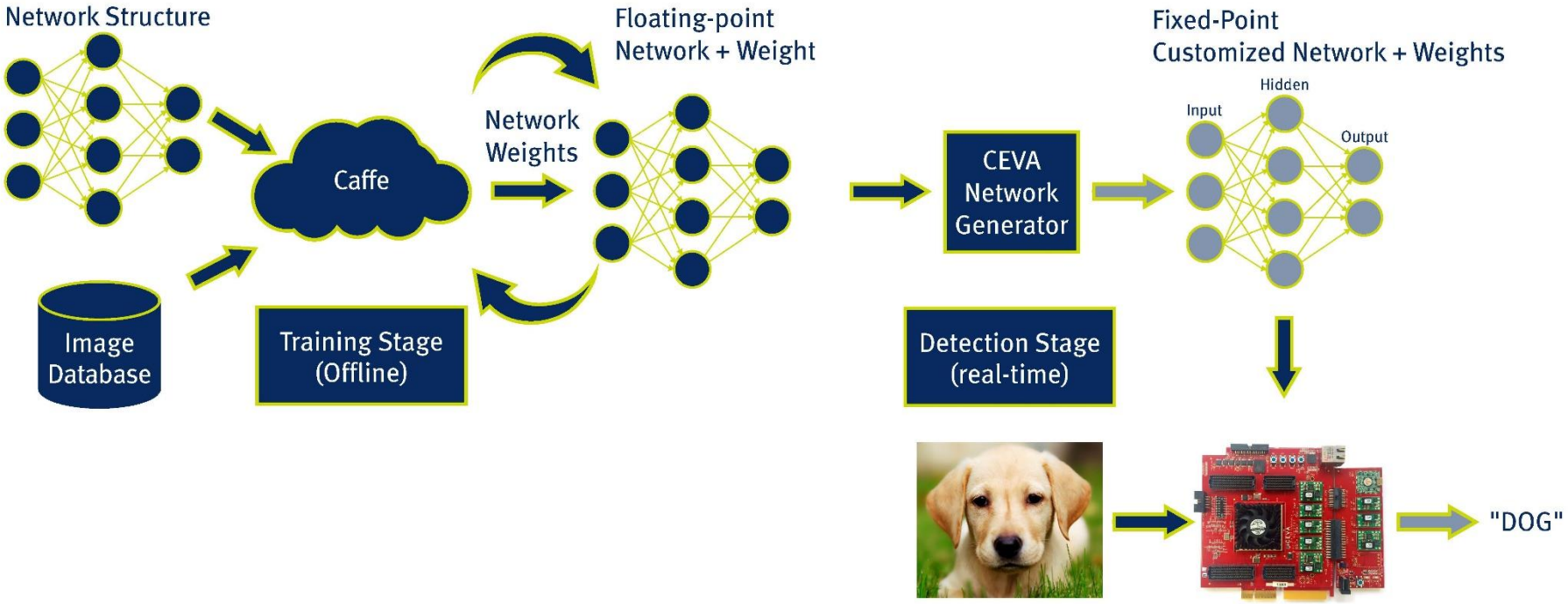
**9x**

Lower Power**

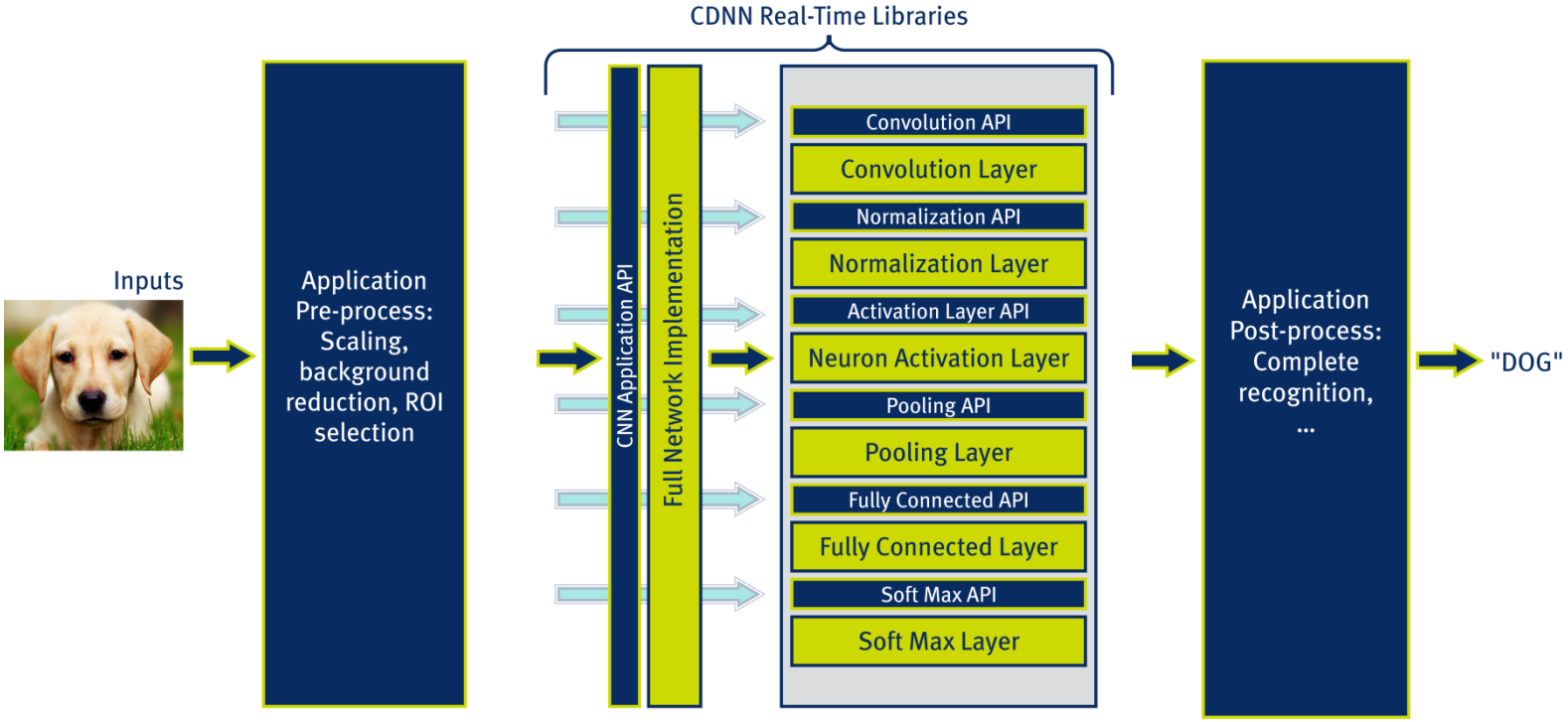
**30%**

Faster Processing*

# CNN Usage Flow with Caffe & CDNN



# Real-Time CDNN Application Flow



# Presentation Outline

1. CEVA
2. Automotive Trends Challenges
3. Efficient implementation of Neural Networks
4. Alexnet Example

# Example CNN – AlexNet

Example based on Caffe open source implementation for CNN

| Object | AlexNet PC Probability (floating point) | AlexNet on XM4 Probability (fixed point) |
|---|---|---|
| Labrador retriever | 90.44% | **91.01%** |
| Golden retriever | 4.45% | 3.98% |
| Beagle | 0.21% | **0.18%** |
| Kuvasz | 0.12% | **0.10%** |
| Classification Probabilities | | |

| ▲ | <1%

# CEVA-XM4 CDNN Development Platform

i.MX6

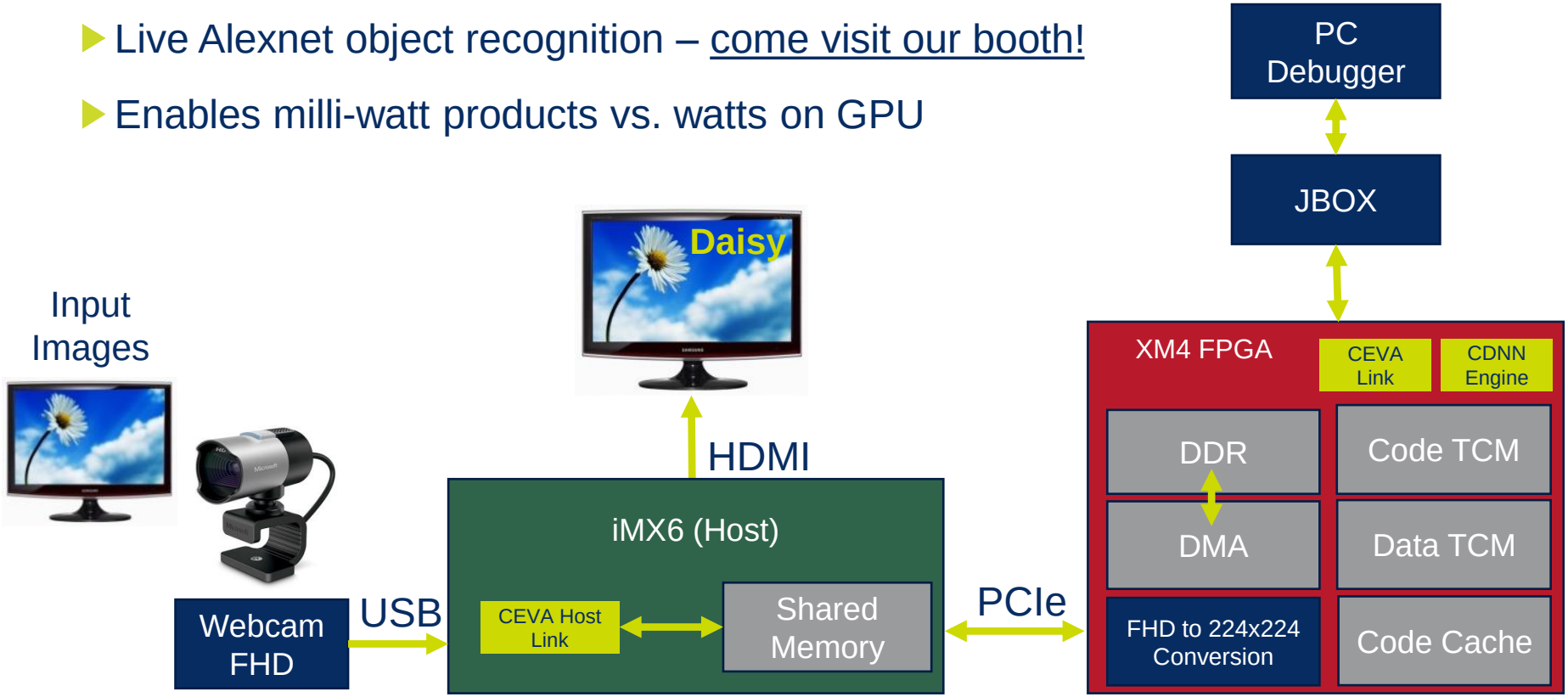
Host running Linux applications

PCIe

XM4 FPGA

# CEVA-XM4 CDNN Demo

- Live Alexnet object recognition – come visit our booth!
- Enables milli-watt products vs. watts on GPU



# CEVA – Efficient System Approach

## XM4 – Vision Processor

- Flexible:  Fully programmable
- Efficient:  Near full utilization of core processing
- Fixed point and floating point operations
- Optimized for image and vision applications
- Small Die size to enable cost effective designs

## Offline Network Generator

- Auto convert fully trained networks
- Floating to fixed point conversion – Embedded Target
- Network optimized for embedded vision IP core – CEVA XM4

## Real Time NN Libraries

- RT Libraries available for quick algorithms development and deployment
- Optimized for CEVA-XM4 vision DSP
- Common vision and  neural network libraries

**CEVA Vision Platform approach for efficient implementation**

CEVA®

# Thank You

CDNN

www.ceva-dsp.com