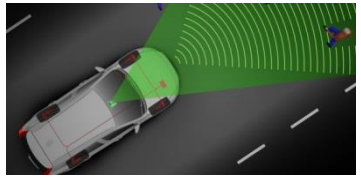
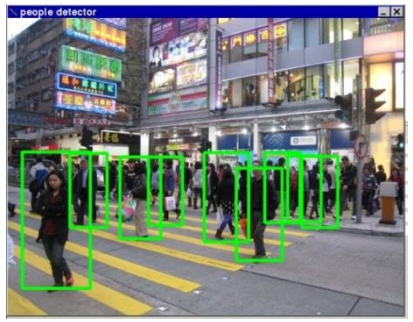
SYNOPSYS®
Silicon to Software™

Challenges and Solutions for Deep Learning Processors for Embedded Vision

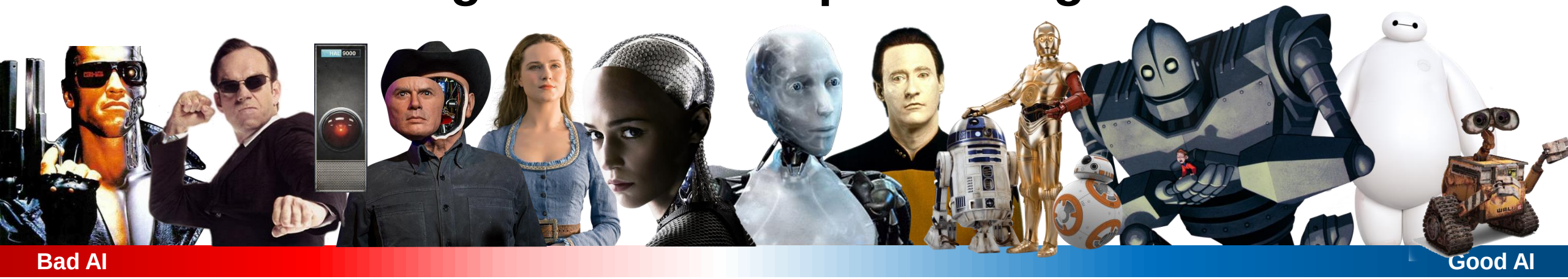
Pierre Paulin,
Director of R&D
Synopsys Inc.

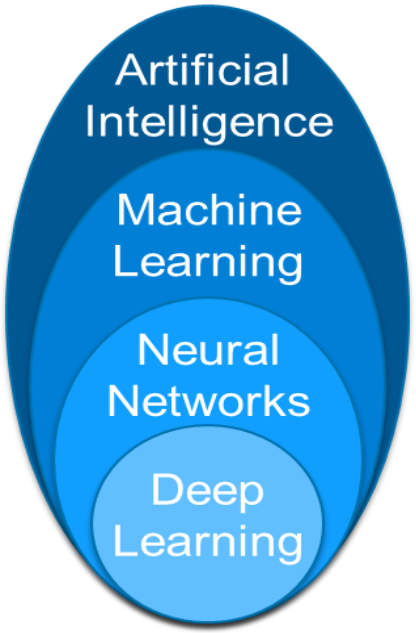




26 March 2018

Artificial Intelligence and Deep Learning





- Artificial Intelligence
  - Human levels of intelligence exhibited by machines
  - Narrow AI: Technology outperforming humans in a narrowly defined task
- Machine learning
  - An application of artificial intelligence that uses algorithms to analyze large amounts of data and then infers some information about the real world from the data
- Neural Networks
  - A class of machine learning algorithms – modeled after the human brain with a neuron representing the computational unit and the network describes how these units are connected to each other
- Deep Learning / Deep Neural Networks
  - A subset of machine learning using artificial neural networks with input, output and ‘hidden’ intermediate. Deep neural networks are capable of learning using large data sets

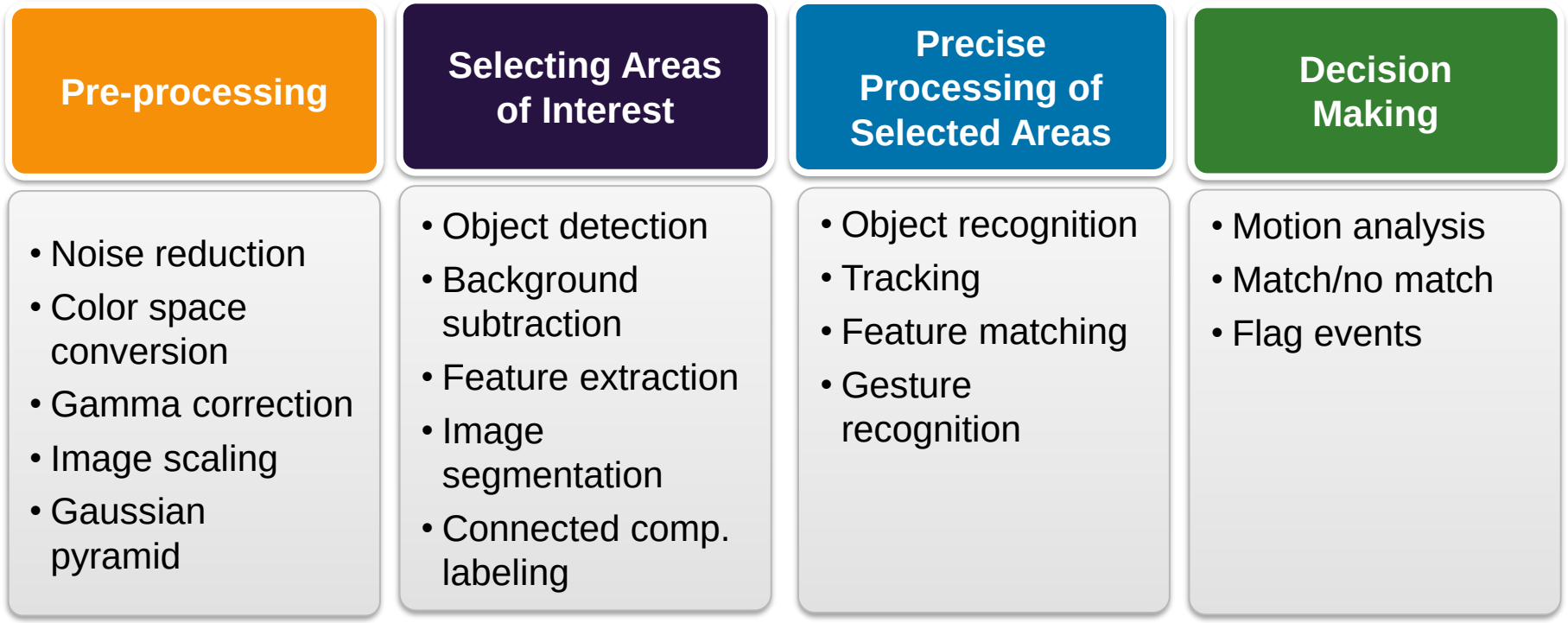
SYNOPSYS®

Embedded Vision Applications

- Surveillance cameras
- Automotive
  - ADAS
  - Autonomous driving
- Consumer
  - Drones
  - Smart Home / IoT
- Mobile
- Mixed/augmented reality glasses
- Image Super Resolution
  - Digital TV
  - Digital Still Camera

SYNOPSYS®

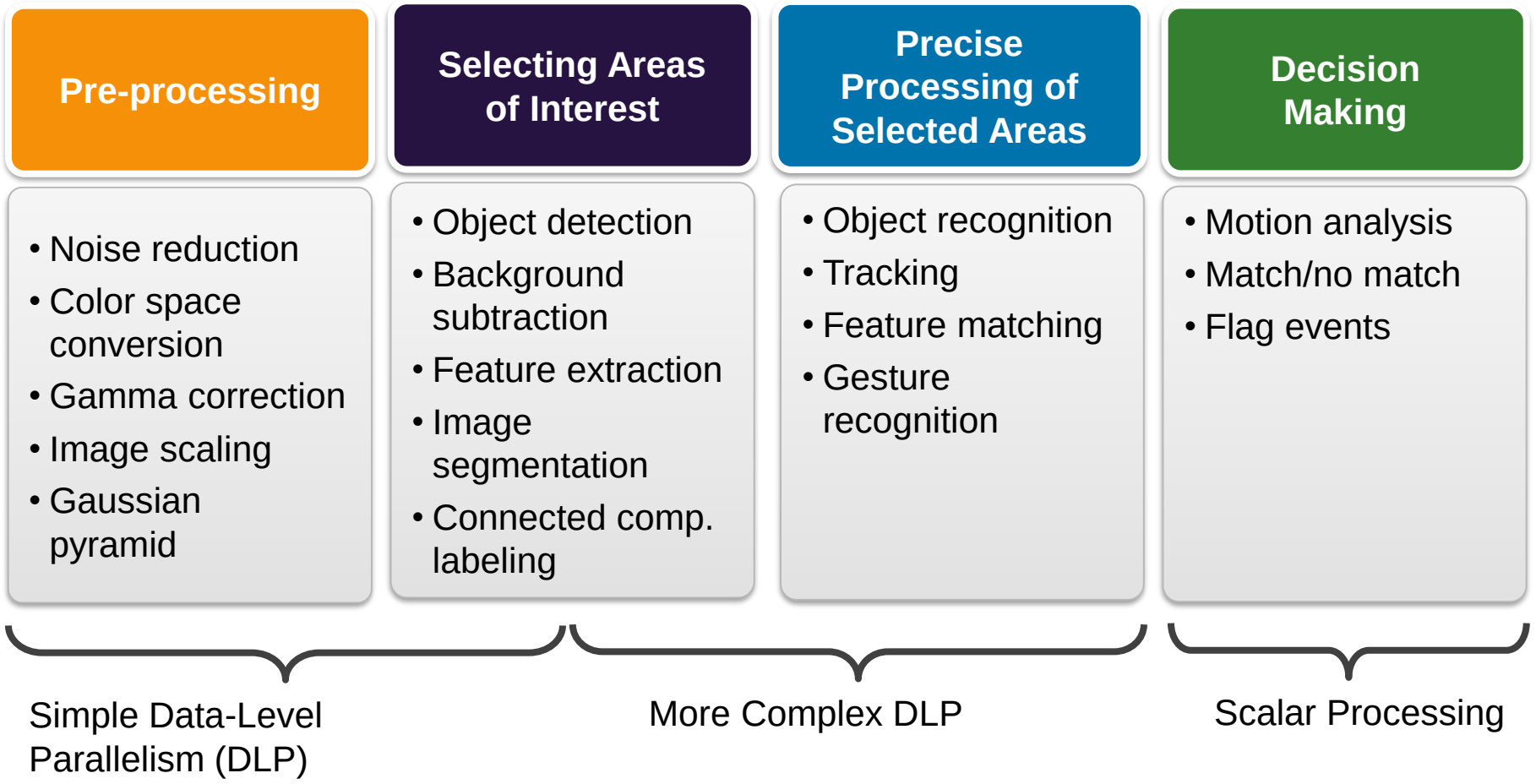
Traditional Embedded Vision Pipeline



- Mathematical approach to image analysis and interpretation
  - Extracting simple features: lines, edges
  - Integrating simple features into higher level features
  - Segmentation by simple analysis of localized pixels
- Shallow machine learning for classification

SYNOPSYS®

Traditional Embedded Vision Pipeline

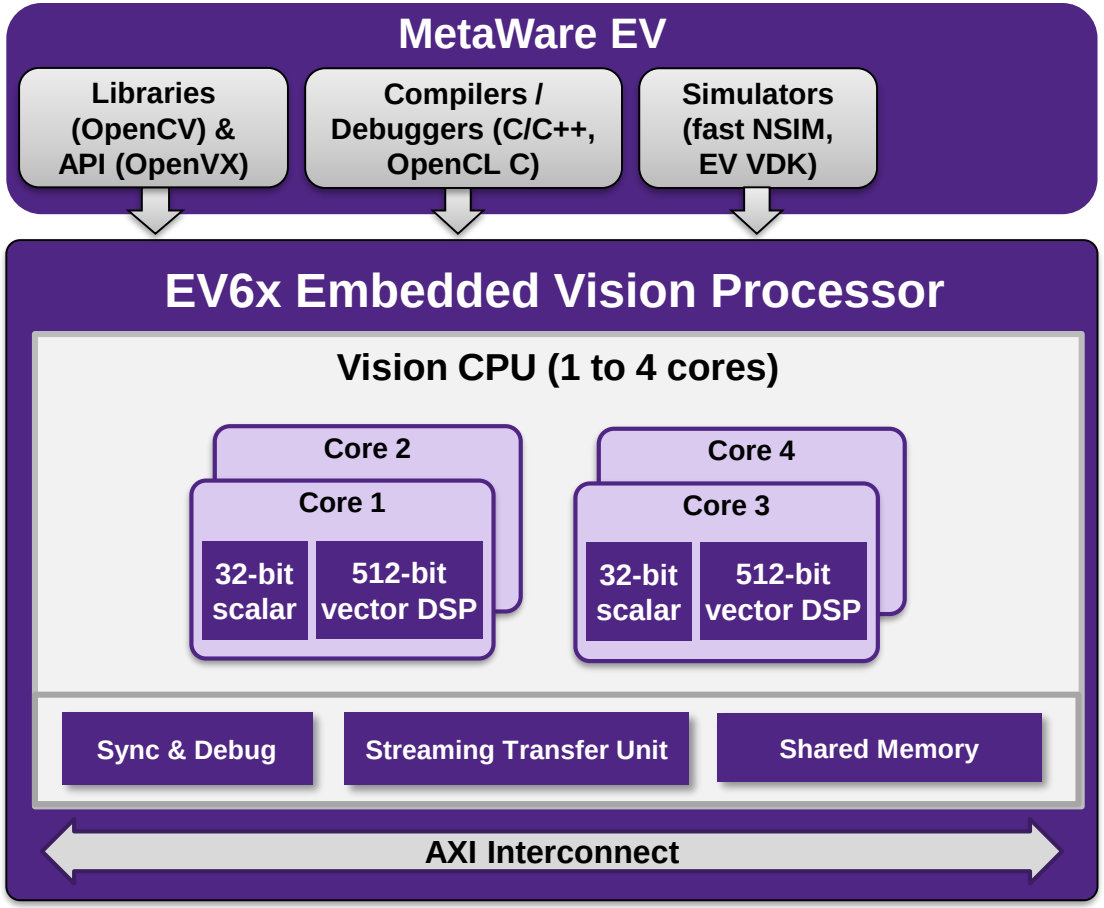


- Complex, heterogeneous processing
- High performance required → multi-core vector processing

SYNOPSYS®

DesignWare® EV6x Embedded Vision Processor IP

Scalable Hardware-Software Solution for High Accuracy Vision Processing



- Wide vector DSP processing
  - Supports full range of vision algorithms
  - Scalable from 1 to 4 vision CPU cores
    - EV64 delivers up to 990 GOP/s*
    - EV64 delivers up to 328 GMAC/s*
- High productivity standards-based toolset
  - OpenCV libraries, OpenVX framework, OpenCL C compiler, C/C++ compiler
- Supports requirements for traditional embedded vision processing
  - Pre-processing, scaling, region-of-interest selection, decision making

* 16 nm FinFET

SYNOPSYS®

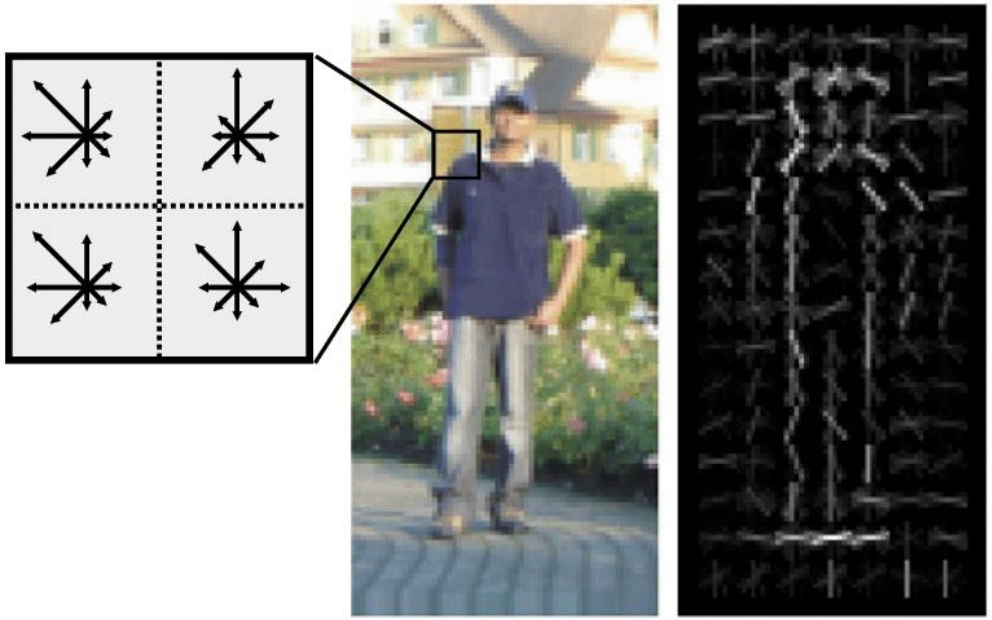
Traditional Computer Vision for Object Detection

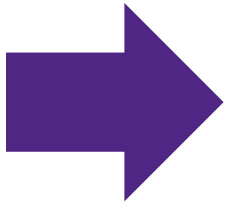
Histogram of oriented Gradients (HoG) example

In the past, most pattern recognition tasks were performed on vector processing units with programs hand-tuned for feature extraction followed by shallow learning

Histogram of Oriented Gradients (HoG)

Object appearance and shape within an image can be described by the distribution of intensity gradients or edge directions.

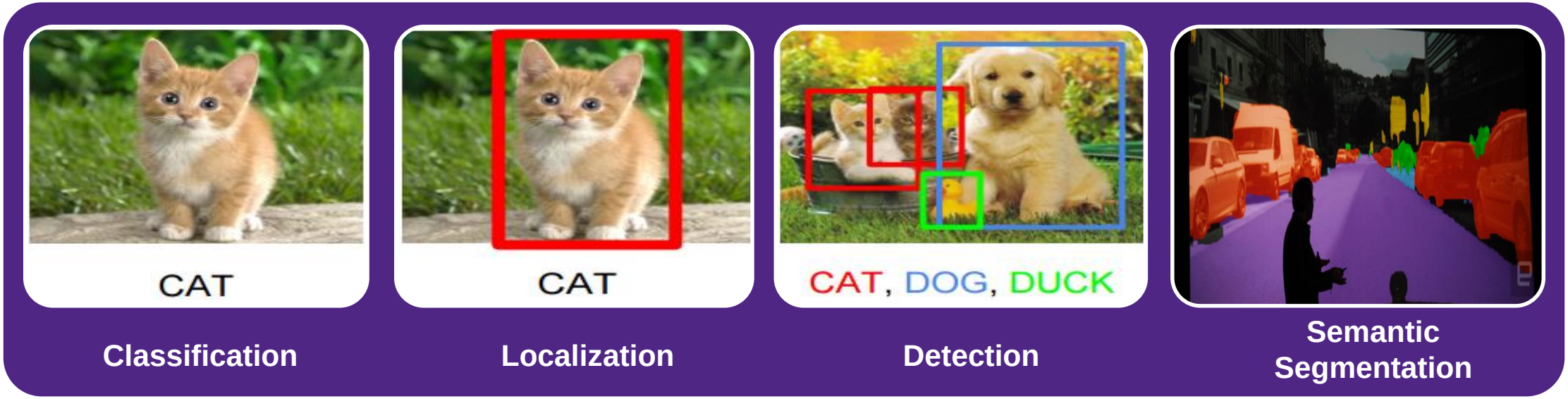




Shallow Learning
e.g. SVM

Recent EV Market Trends & Perspective

- Displacement of traditional vision algorithms with deep learning for improved accuracy
  - i.e. Pedestrian Detection (HoG), Face detection (Viola-Jones) moving to Convolution Neural Network (CNN) based deep learning approaches
- Rapid evolution in deep learning technology to a wide set of applications



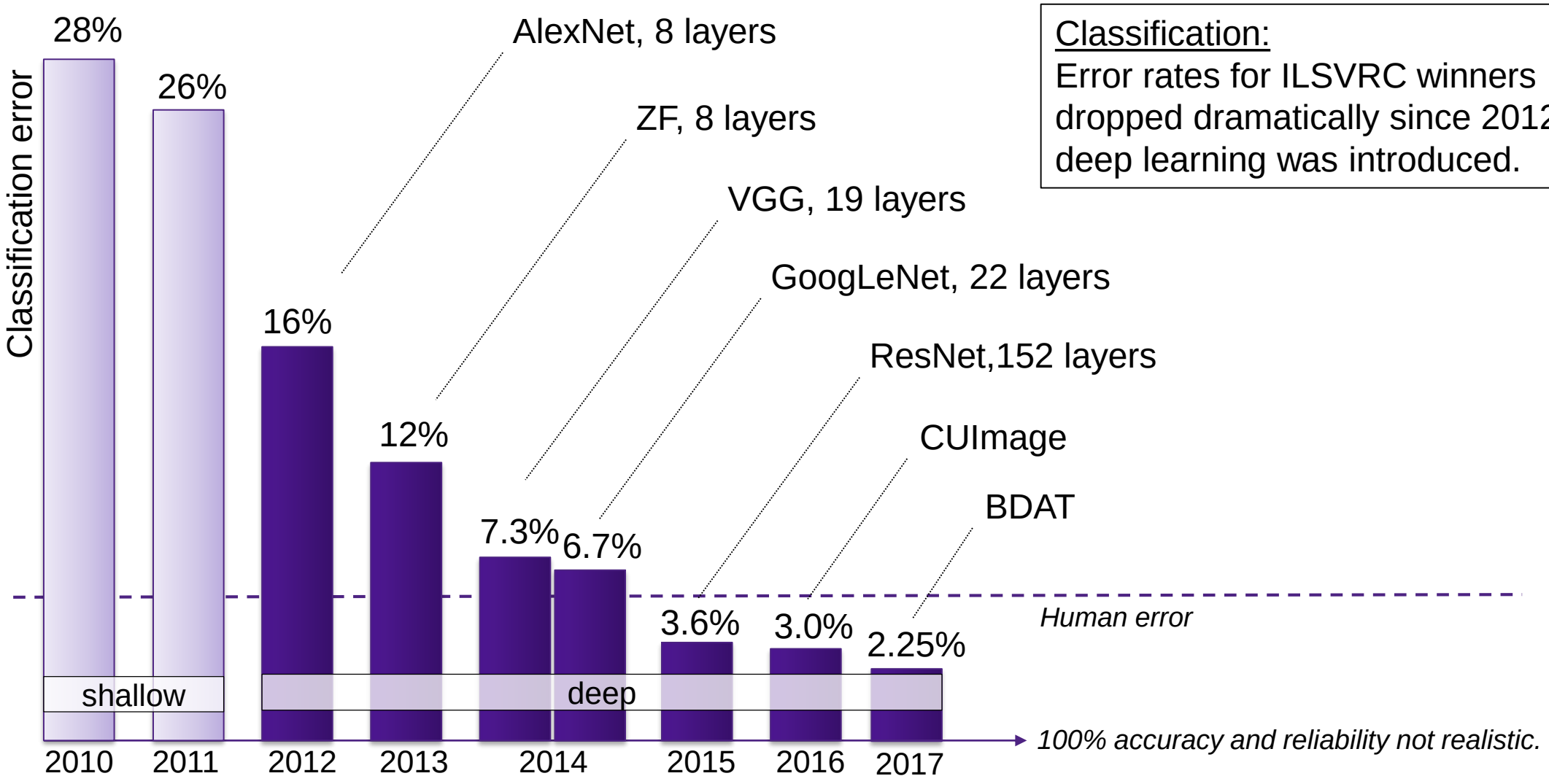
SYNOPSYS®

CNN-based Object Classification


```
Top 5 classes:
1: moped
2: motor scooter, scooter
3: barrow, garden cart,
   lawn cart, wheelbarrow
4: tricycle, trike,
   velocipede
5: crash helmet
```


Deep Learning Approaches Human Levels of Accuracy

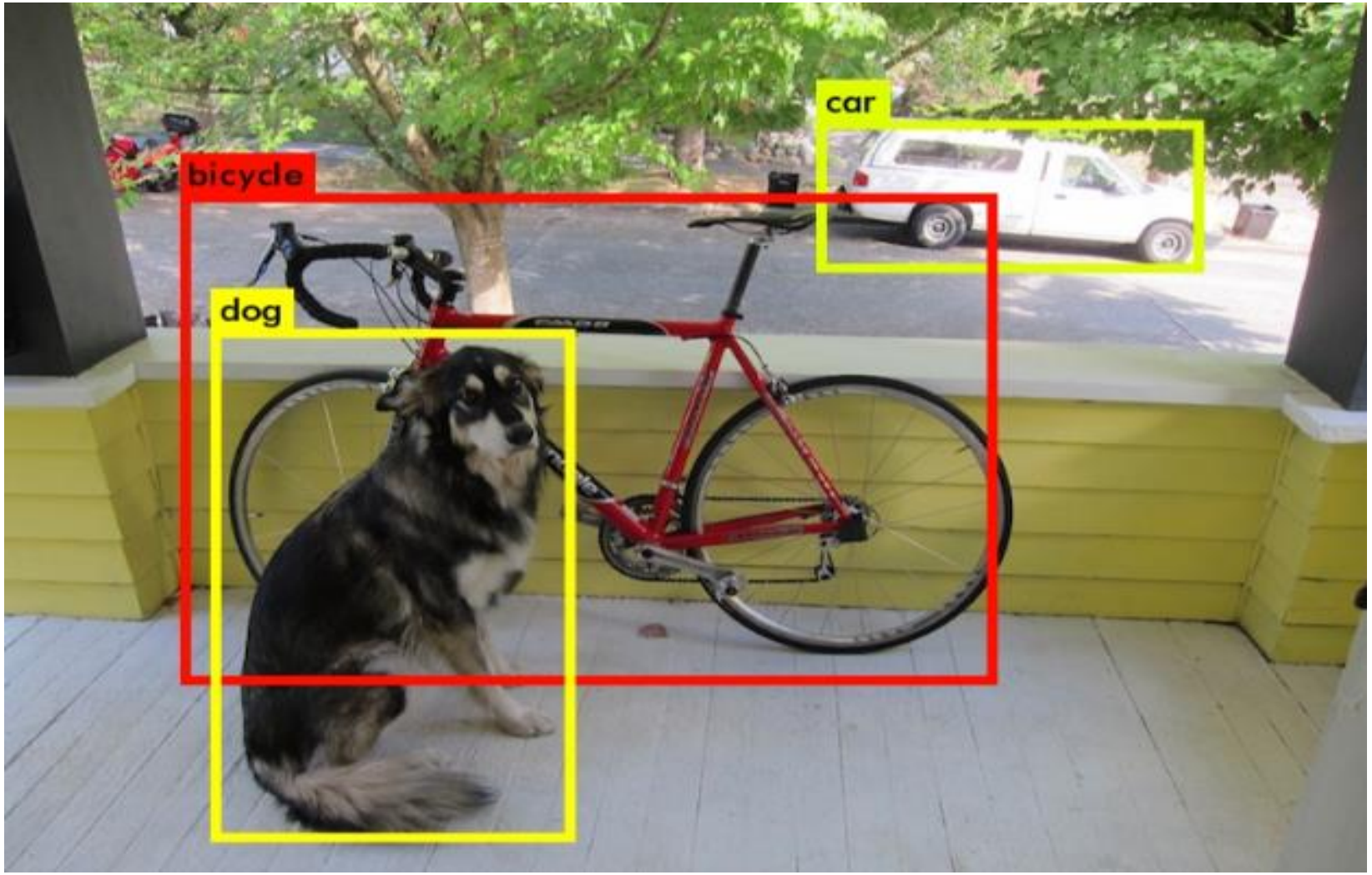
ImageNet Large Scale Visual Recognition Challenge Results



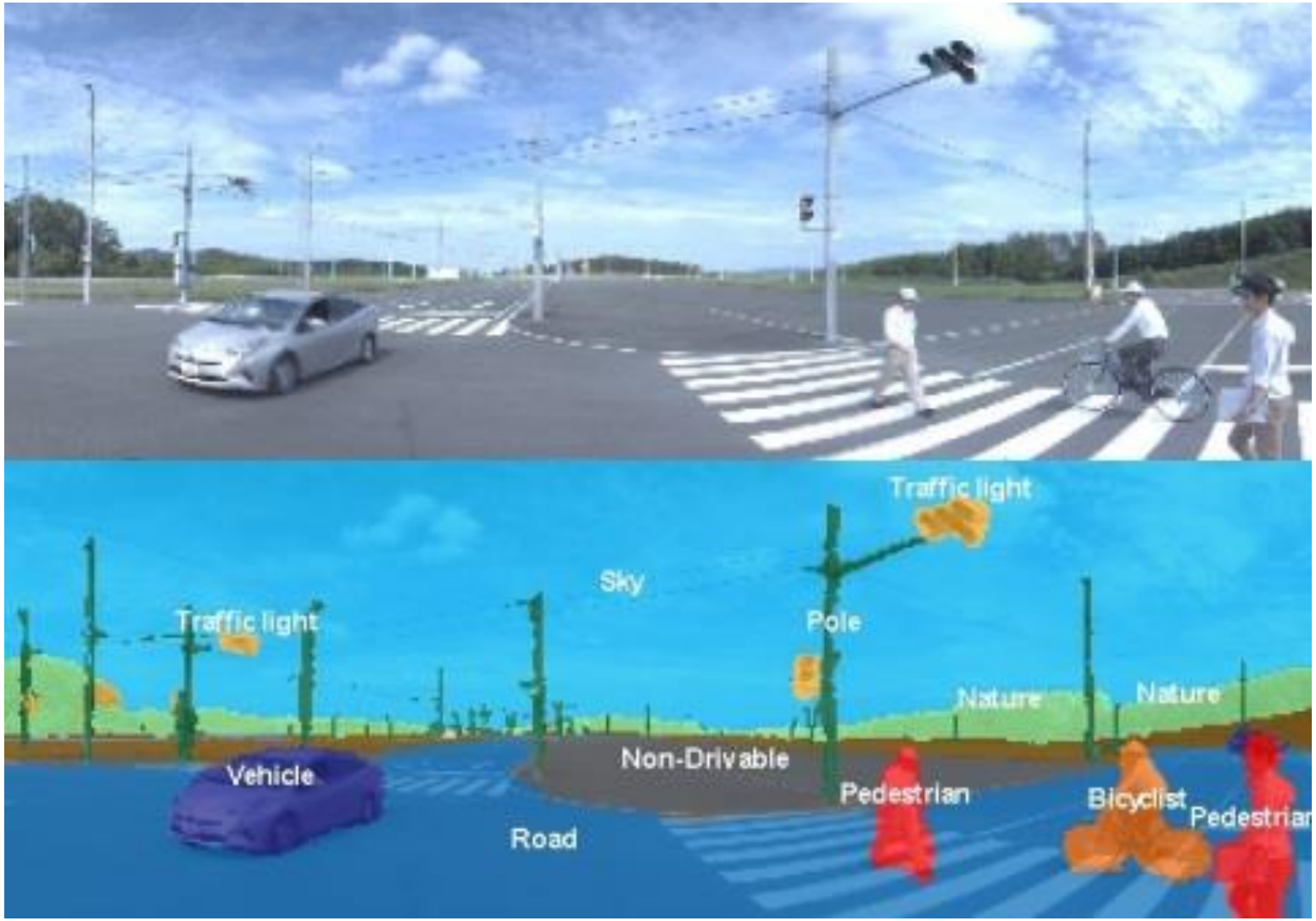
Classification:
Error rates for ILSVRC winners have dropped dramatically since 2012 when deep learning was introduced.

SYNOPSYS®

Yolo: Object Detection and Localization



Scene Segmentation



Source: Press Release by Toshiba and Denso, 17 Oct. 2016

SYNOPSYS®

CNN-based Denoiser

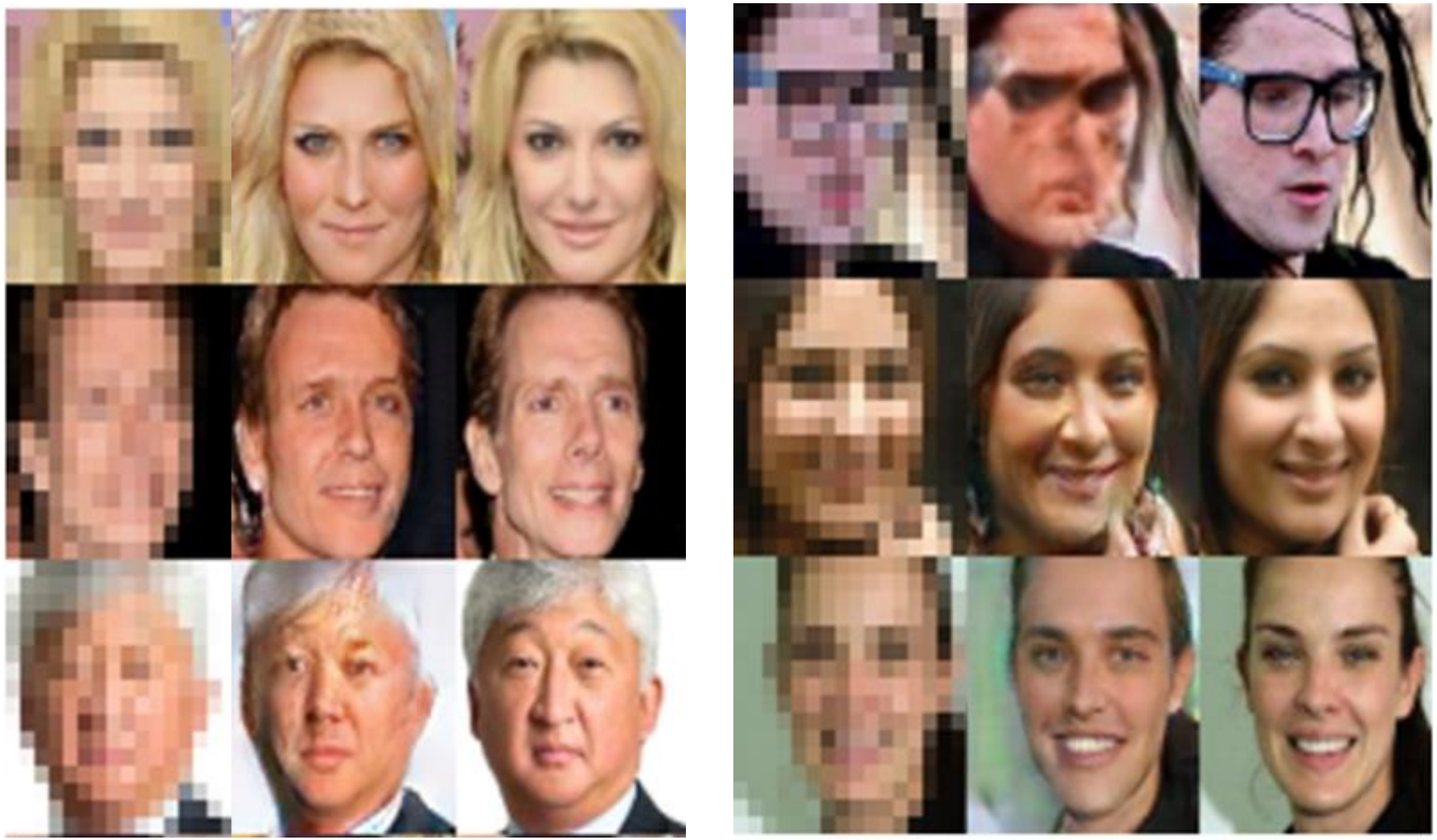
After

Before

SYNOPSYS®

Super resolution using CNN → 600 GMAC per 4K frame

Source CNN Reference Source CNN Reference



"Image Super-Resolution Using Deep Convolutional Networks (2016), C. Dong et al."

Neural Networks for Radar Waveform Recognition

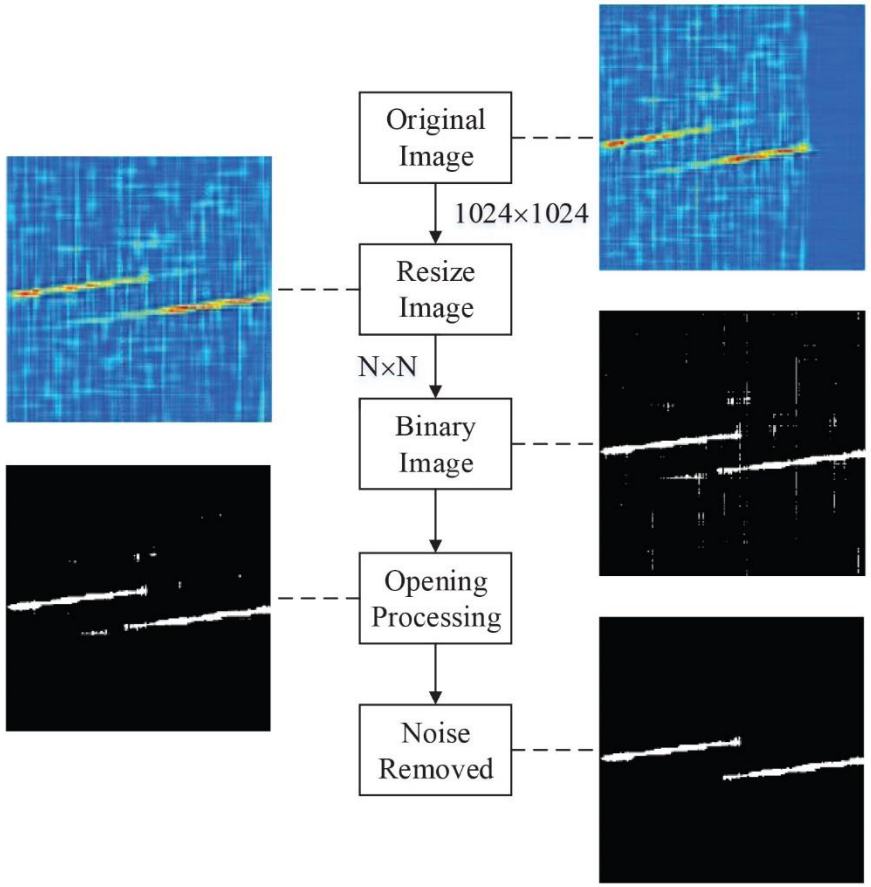
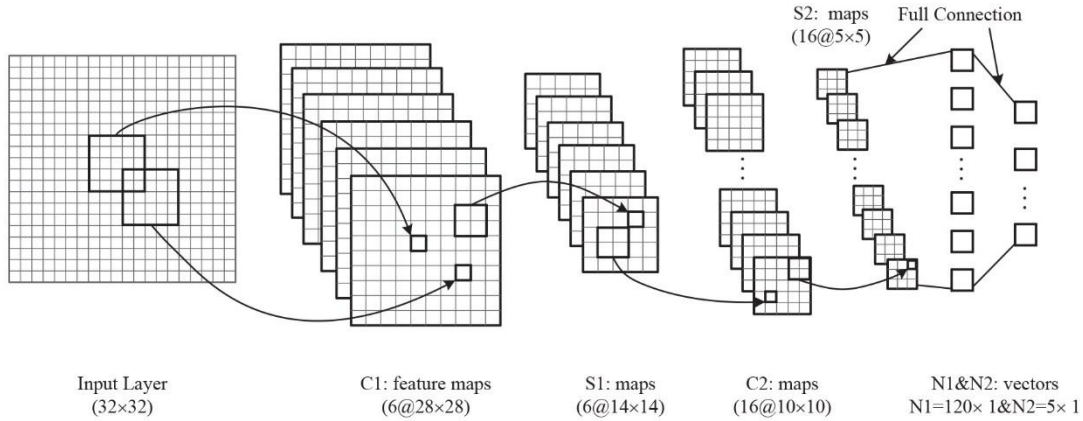


Figure 4. In this figure, we exhibit the processing with P3 code at an signal-to-noise ratio (SNR) of −4 dB.

- An automatic radar waveform recognition system to detect, track and locate low probability of intercept (LPI) radars.
- The detected signals are processed into binary images which are resized for CNN
- The finished binary images are used in CNN and feature extraction

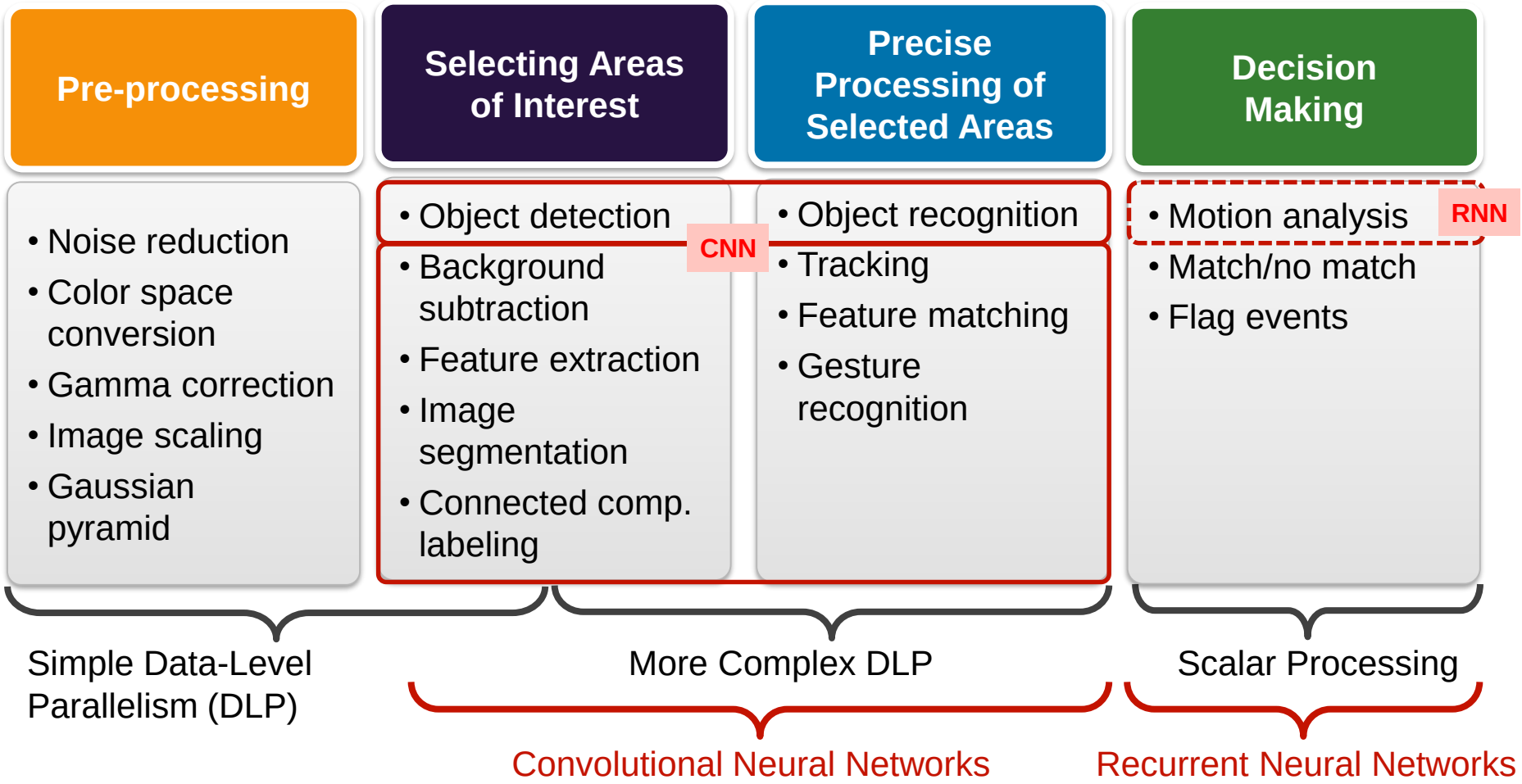


Ming Zhang, Ming Diao, Lipeng Gao * and Lutao Liu
http://www.mdpi.com/2073-8994/9/5/75

SYNOPSYS®

Emerging Embedded Vision Solutions

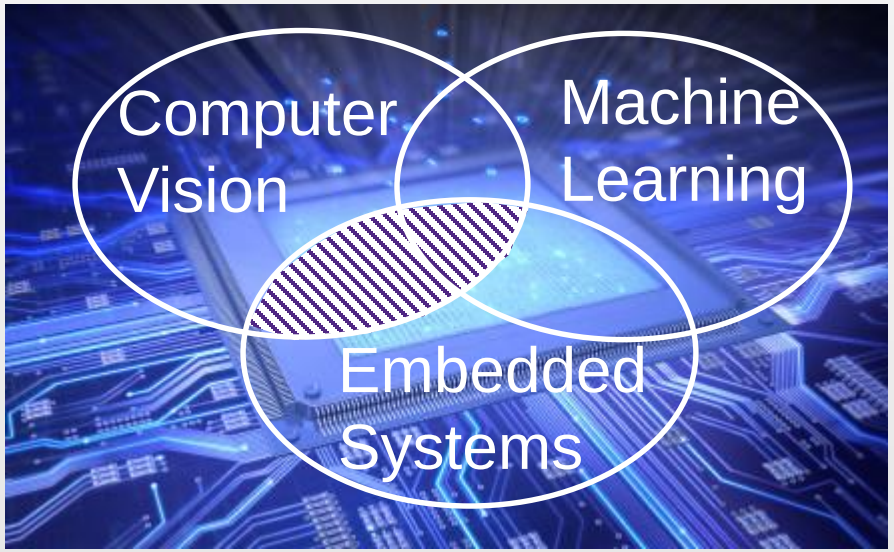
- Combining the best of traditional vision and deep learning approaches
- Combining scalar, vector processing with specialized CNN engine



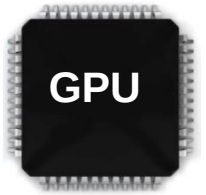
SYNOPSYS®

Choosing a Processor for Embedded Vision

Embedded Vision



- On-chip
- Cost efficient
- Energy efficient
- Real-time



Performance Power Area

Performance Power Area

Performance Power Area

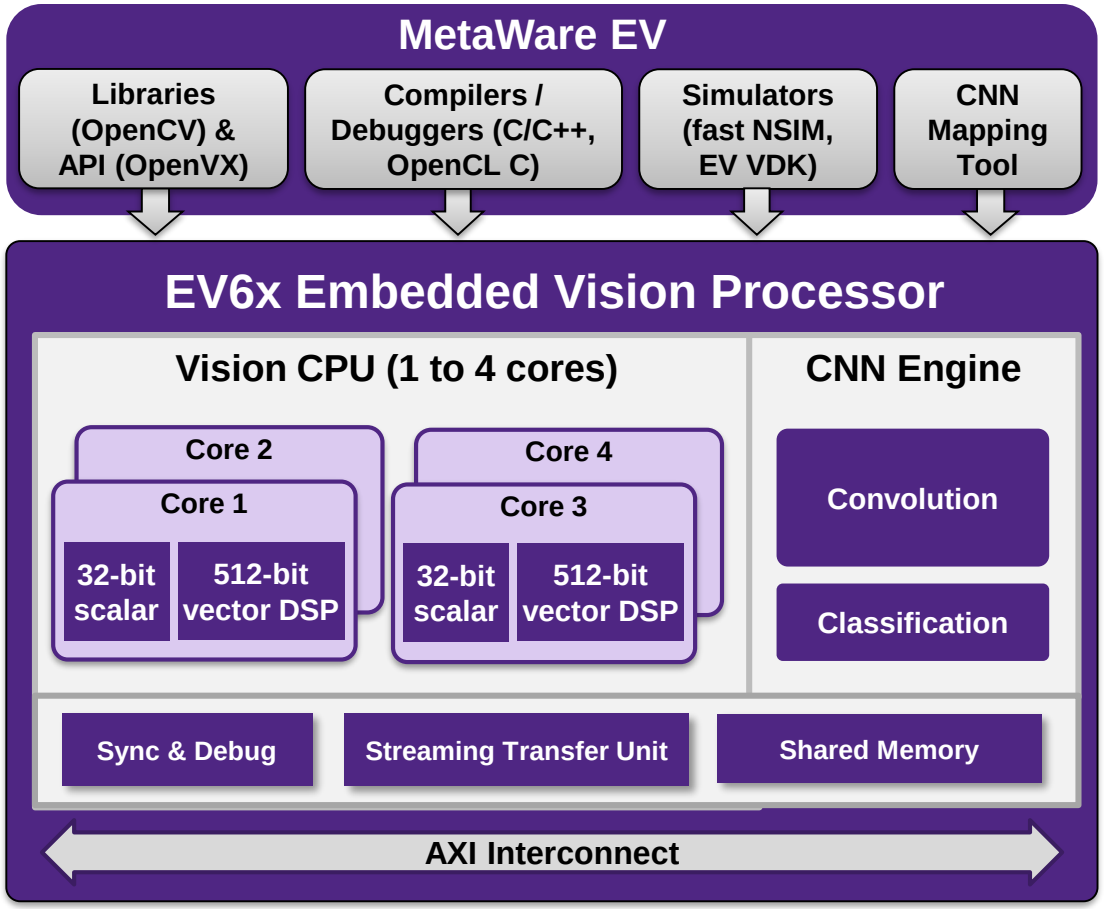
Performance Power Area

source: GPU-Based Deep Learning Inference: A Performance and Power Analysis, NVIDIA, 11/15

SYNOPSYS®

DesignWare® EV6x Embedded Vision Processor IP

Scalable Hardware-Software Solution for High Accuracy Vision Processing



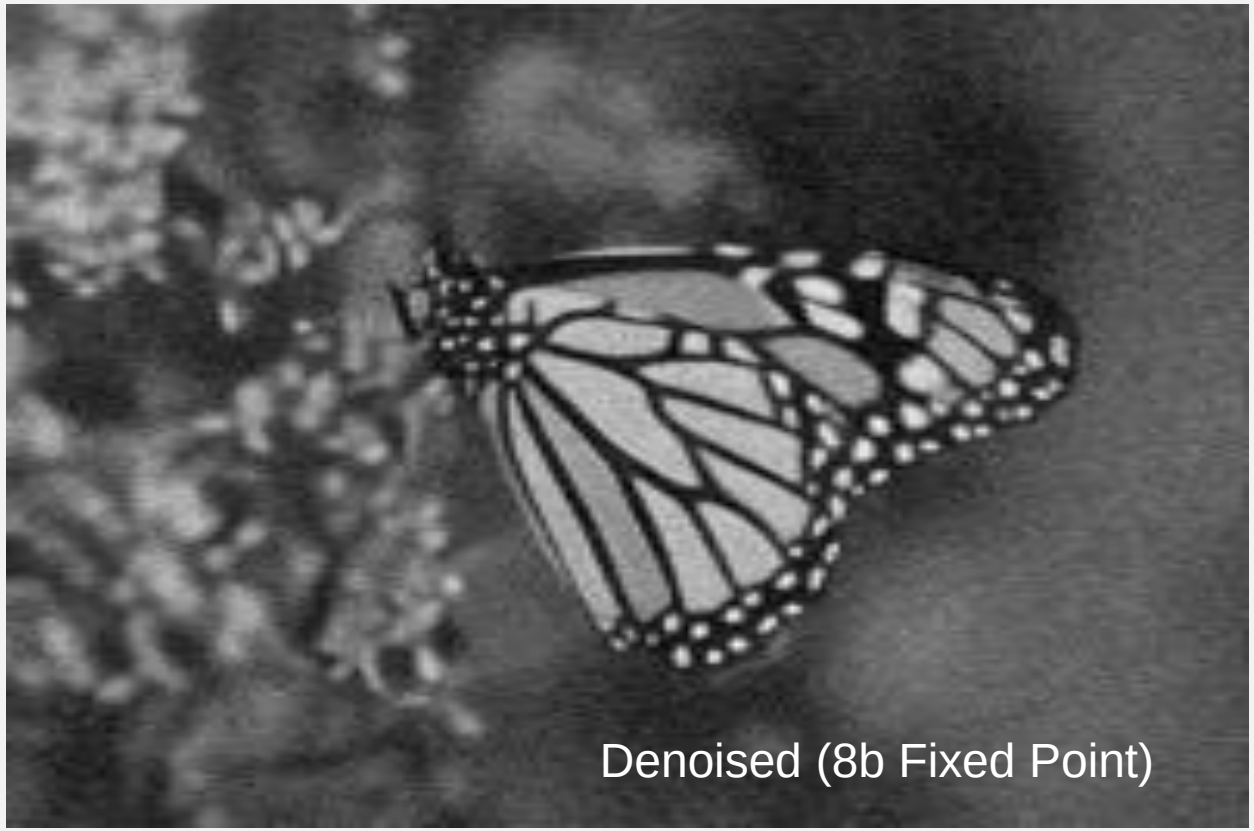
- Integration of traditional CV and CNN
  - Unified scalar, vector DSP and convolutional neural network (CNN) architecture
- High-performance, programmable CNN engine
  - Scalable from 880 to 1760 to 3520 MACs/cycle delivering up to 4.5 TMACs/sec performance*
  - Dedicated memory architecture
  - Dedicated DMA
  - Multi-dimension parallelism
  - Support of 8 bit and 12 bit data accuracy
  - Supports all modern CNN graphs (e.g., VGG16, GoogleNet Inception, ResNet, DenseNet, MobileNet, SqueezeNet, Yolo, Faster R-CNN, SSD, Denoiser, etc.)
- CNN mapping tool from well-known CNN frameworks, including Caffe and Tensorflow

* 16 nm FinFET

SYNOPSYS®

Denoiser Filter Results

Close comparison of 12-bit and 8-bit accuracy



Summary

- Deep learning techniques, like convolutional neural networks, offer the highest accuracy for object classification, detection, and scene segmentation
- CNN replacing traditional computer vision algorithms
- Specialized CNN architecture offers area and power efficiencies, and higher accuracy for image quality improvement applications

DesignWare EV6x Processors

- Unified multicore processor for automotive vision processing
- Scalar + vector DSP + CNN engine
- State-of-the-art convolutional neural network (CNN)

SYNOPSYS®
Silicon to Software™

Thank You