# COMPUTER VISION SEES A BRIGHT FUTURE

BADRU AGARWALA, GENERAL MANAGER, MENTOR, A SIEMENS BUSINESS

WHITEPAPER

Mentor®
A Siemens Business

HIGH-LEVEL SYNTHESIS

## INTRODUCTION

Computer vision attempts to replicate the human perception and associated brain functions in order to acquire, process, analyze, understand, and then act on an image. But replicating this process is extremely challenging. Why? Let's take an example. While walking down the street, you find yourself suddenly hungry. As you pass by businesses along the way, your brain automatically applies filters to keep out thoughts unrelated to satiating your hunger and puts thoughts of "restaurant" top of mind and then categories of your favorite foods pulled from memory are applied, such as "pizza" and "hamburgers." Once you see a "restaurant" and match that to one offering "pizza" your sense of smell and fond memories takes over and you decide to enter the pizzeria.

Now, think about the hardware and software required to perform this same task. The seemingly-simple concept of isolating an image in order to identify it has taken years of research and development to accomplish. Today, teams using computer vision hardware and software algorithms coupled with deep learning are seeing success in identifying objects. But as of now, computer vision systems cannot be pointed at a random object and asked, "what is that?" and have it answer with 100 percent reliability.

What business comes to mind when you hear the term computer vision? Security, robotics, self-driving cars, or farming? Yes, farming. When your lifeblood demands a lot of manual labor and includes a daily fight against Mother Nature, you are looking for ways to automate the farm. The Grail for a farmer is to sit at the kitchen table monitoring a fully-automated farm and making remote adjustments (Figure 1).

*Figure 1: A vision of a fully-automated farm*

Computer vision-based companies recognize the plight of the farmer and they offer a wide range of solutions. For example, AgEagle® offers a drone that stitches together aerial images of the farm, providing a geo-tagged representation of the farmer's land and crops. Prospera® recently raised investments for their computer vision product that analyzes crop images for diseases, nutrient deficiency, irrigation, and crop rotation using artificial intelligence. Blue River Technology® offers a lettuce bot that uses computer vision and machine learning to identify and spray weeds and to thin crops. Energid Technologies® sells a citrus harvesting system that uses six-camera computer vision technology to pick fruit. Case IH® is developing self-driving tractors to prepare land and to harvest crops. John Deere® has created JDLink™ that tracks and analyzes farm machines, monitors maintenance issues, and connects directly to a local dealer for service. And, if farmers need a unique part for a piece of equipment, they can send a picture of it to PartPic® which uses algorithms to identify the part and finds a vendor for it.

There are as many applications and markets out there for computer vision products as there are uses for the human eye (Figure 2).

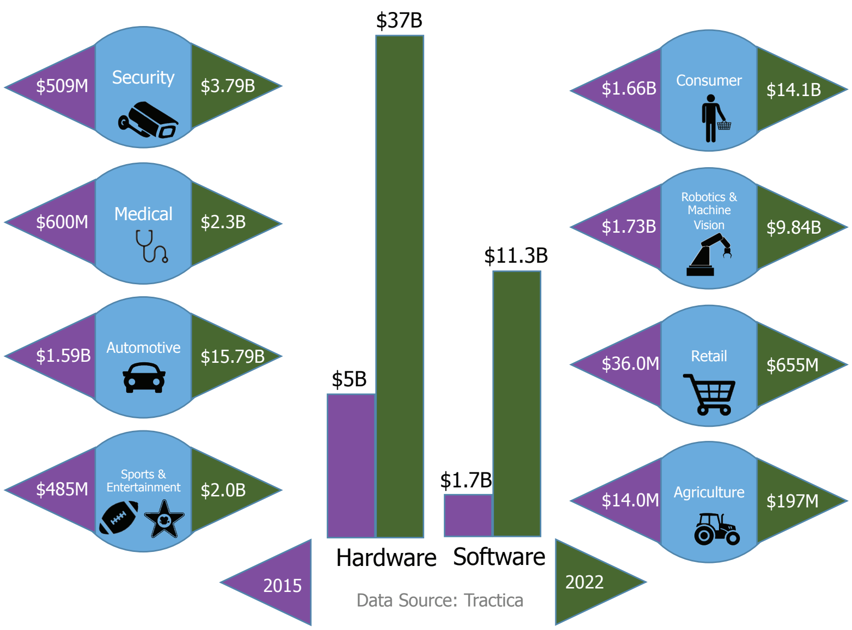


*Figure 2: Hardware and software markets for computer vision*

Semiconductor and component manufacturers, software companies, and product developers are all making major investments in developing computer vision products in these major markets:

- Security: live camera feeds for surveillance, people tracking, and crime prevention.
- Medical: generating 3D images from body scanners.
- Automotive: driver assistance and autonomous vehicle systems.
- Sports and Entertainment: technology and effects for sporting events, movies, and TV content.
- Consumer: smartphones, cameras, biometrics, and virtual reality.
- Robotics and Machine Vision: industrial automation, inspection, and drones.
- Retail: customer tracking and purchasing analytics.
- Agriculture: systems for crop inspection, harvesting, and land preparation.

## WHY IS COMPUTER VISION HOT?

Hardware and software improvements have driven the computer vision market as the race to achieve 100% accuracy is underway. Convolutional neural networks (CNN) are the technology of choice for image recognition and classification and many advances in this area by companies such as Google® and IBM® are achieved every day. It is only within the past 5 years that the ecosystem (Figure 3) has evolved to accelerate computer vision product development.

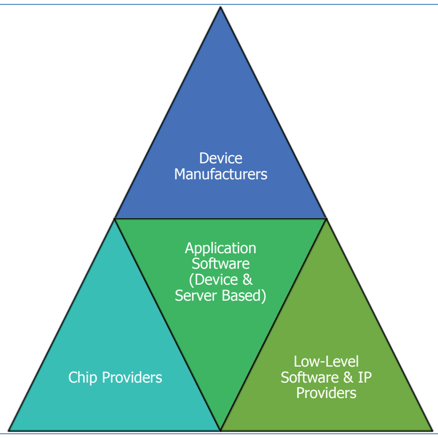


*Figure 3: Computer vision product development ecosystem (Content source: Tractica)*

Many technical advances contribute to this acceleration:

- Deep learning advances for statistical analysis of images to successfully identify and classify objects.
- Wireless networks are available to the masses.
- High-bandwidth availability to transmit images for processing and analysis.
- Massive data storage and access is available for CNN training.
- Gigantic image databases available for training networks.
- Extremely powerful heterogeneous systems tailored to massively-parallel computing exist.
- Software stack and open source libraries are available to build differentiated products without re-inventing the infrastructure.
- Chip and memory prices are falling, making advanced computer vision product development affordable.
- CMOS image sensors are now on par with CCD sensors and they can be fabricated using the same technology used for chipsets for less money.

There are also key market barriers to overcome that provide opportunities for companies to invent new solutions:

- 100% accuracy of computer vision is still not attainable.
- New businesses centered on computer vision need to make strong business cases for novel solutions and this takes time.
- Most of the computer vision advances take place in research labs of universities and big companies with deep pockets and skilled employees, meaning development costs could be a barrier to entry.
- There is a shortage of skilled engineers with computer vision and deep learning expertise.

- There are many more software engineers in the workforce than hardware engineers.
- There is no recognized "killer application" yet for computer vision products. Some believe this will be virtual reality or self-driving cars.
- Complex CNN algorithms are computationally intensive and consume a lot of power for CPUs and GPUs, limiting onboard (smartphone, drones, etc.) analysis of computer images.
- New algorithms and training solutions change and evolve constantly, making hardware implementations very difficult to complete before the next, better idea comes along.

The computer vision market is wide open for startups to deliver new products and applications. As a sign of this, artificial intelligence (which includes neural networks, deep learning, and CNN) startups have received $3.48 billion in venture capital (VC) between Q3 2015 to Q3 2016 (Venture Pulse – KPMG International® and CB Insights®).

The most active corporate investors from 2011 to June 2016 in artificial intelligence (AI) were Intel Capital®, Google Ventures®, GE Ventures®, Samsung Ventures®, and Bloomberg Beta® (2016 CB Insights). And in that same time period, nearly 140 VC-backed AI companies were bought out by Google, IBM, Intel®, Apple®, Salesforce®, and Samsung® (2016 CB Insights).

## COMPUTER VISION SOLUTIONS

The basic architecture of a computer vision system consists of a lens, hardware, and software (Figure 4). Developers of these systems decide what algorithms are implemented in hardware and which remain in software.

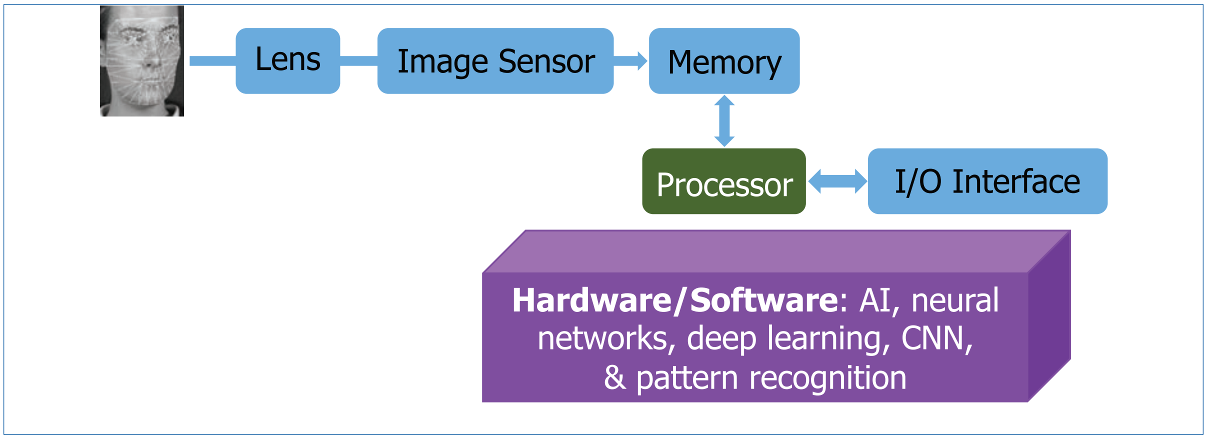


*Figure 4: The basic architecture of a computer vision system*

Companies are developing computer vision systems tied to deep learning systems that live on the edge of the Internet of Things (IoT), in onboard systems, or that perform inference analysis in the Cloud. Of particular interest is the processor (or array of processors) used in the computer vision system. The deep learning algorithms perform billions of computations and need to produce results quickly. These systems can employ solutions using CPU/DSP, GPU, FPGA, embedded FPGA, or ASIC technologies. Each has its advantages and disadvantages, and GPUs are currently the technology of choice. But, GPUs consume a lot of power and they might be unsuitable for onboard systems. In addition, GPUs do not allow flexible hardware configurations tailored to particular algorithms. Thus, FPGA and embedded FPGA technology is starting to make inroads into computer vision processing. If the market demands, implementations using FPGA technologies migrate to ASIC technologies. Teams need a fast way to see how their algorithms perform on any of these technologies before production, without changing their algorithm description.

## THE HIGH-LEVEL SYNTHESIS SOLUTION

Algorithm developers prefer to write code in C++, do not want to learn register transfer languages (RTL) such as Verilog or VHDL and they do not want use the tools and methodologies required for the hardware implementation process.

To address this problem, some algorithm developers write their code in C++ and then use High-Level Synthesis (HLS) tools to quickly make tradeoffs and then automatically generate RTL. These developers can try out multiple implementations of new algorithms, explore the performance and power consumption of implementing these algorithms in hardware or software, and examine the tradeoffs of running on ASICs, FPGAs, CPUs or GPUs. Tradeoffs are also explored by varying the number of processor cores and experimenting with memory sizes and configurations.

Because the world of computer vision coupled with machine learning evolves so quickly, teams need a way to design and verify an algorithm while the specifications and requirements evolve without starting over every time there is a change. HLS flows allow this by using constraints and directives that guide the process, while leaving the algorithm unchanged. For example, at the last minute, a team decides to change the clock speed. That constraint is changed, and the HLS tool re-generates the RTL according to the new clock speed, with no change to the C++ algorithm. This type of late-stage change is not possible in traditional RTL flow, yet with HLS, the RTL is rebuilt with a push of a button.

Fifteen years ago, Mentor recognized that design and verification teams need to move up from the RTL to the HLS level. Working with customers over all those years has resulted in today's Catapult HLS Platform that provides a complete flow from C++/SystemC to optimized RTL (Figure 5).

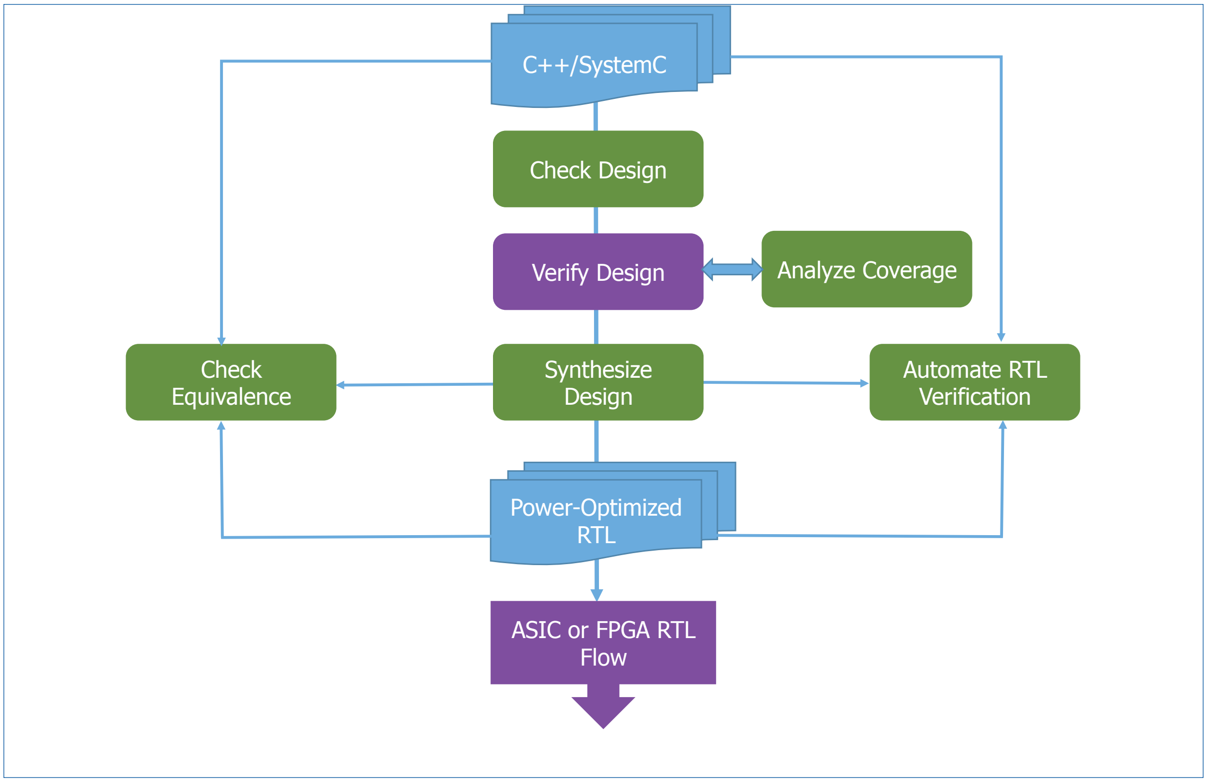


*Figure 5: The Catapult HLS Platform*

The Catapult HLS Platform provides a hardware design solution for algorithm designers that generates high-quality RTL from C++ and/or SystemC descriptions that target ASIC or FPGA implementation. The platform delivers the ability to check the design for errors before simulation, provides a seamless and reusable testing environment, and supports formal equivalence checking between the generated RTL and the original source. The platform flow ensures fast design and verification and delivers power-optimized RTL ready for simulation and RTL synthesis.

By employing these elements of the Catapult HLS Platform solution teams take their computer vision products to market faster and at a lower cost. Imagine getting to market months ahead of competitors who are still using a traditional RTL flow.

## THE TAKE-AWAY

Computer vision product development provides a wide array of opportunities for design automation. From the design of self-contained systems that include analysis of the images, to the chip development for servers in the Cloud that are used to perform deep learning. Custom ASIC, FPGA, and embedded FPGA design flows are well established but are seeing a new audience of computer vision and deep learning developers. High-Level Synthesis solutions are now allowing software algorithm developers to make informed decisions as to which hardware technologies work best for their application without being hardware experts. Yet, hardware designers are also helped by the HLS flow because they can support constantly-evolving architectures. Without a HLS flow, market-leading computer vision innovation is impossible.

To learn more about how HLS is revolutionizing computer vision innovation and to experience customer success stories, view the free video seminar series here.



**Corporate Headquarters**
**Mentor Graphics Corporation**
8005 SW Boeckman Road
Wilsonville, OR 97070-7777
Phone: 503.685.7000
Fax: 503.685.1204

**Sales and Product Information**
Phone: 800.547.3000
sales_info@mentor.com

**Silicon Valley**
**Mentor Graphics Corporation**
46871 Bayside Parkway
Fremont, CA 94538 USA
Phone: 510.354.7400
Fax: 510.354.7467

**North American Support Center**
Phone: 800.547.4303

**Europe**
**Mentor Graphics**
Deutschland GmbH
Arnulfstrasse 201
80634 Munich
Germany
Phone: +49.89.57096.0
Fax: +49.89.57096.400

**Pacific Rim**
**Mentor Graphics (Taiwan)**
11F, No. 120, Section 2,
Gongdao 5th Road
HsinChu City 300,
Taiwan, ROC
Phone: 886.3.513.1000
Fax: 886.3.573.4734

**Japan**
**Mentor Graphics Japan Co., Ltd.**
Gotenyama Trust Tower
7-35, Kita-Shinagawa 4-chome
Shinagawa-Ku, Tokyo 140-0001
Japan
Phone: +81.3.5488.3033
Fax: +81.3.5488.3004

Mentor
A Siemens Business

MGC 03-18 TECH17080-w