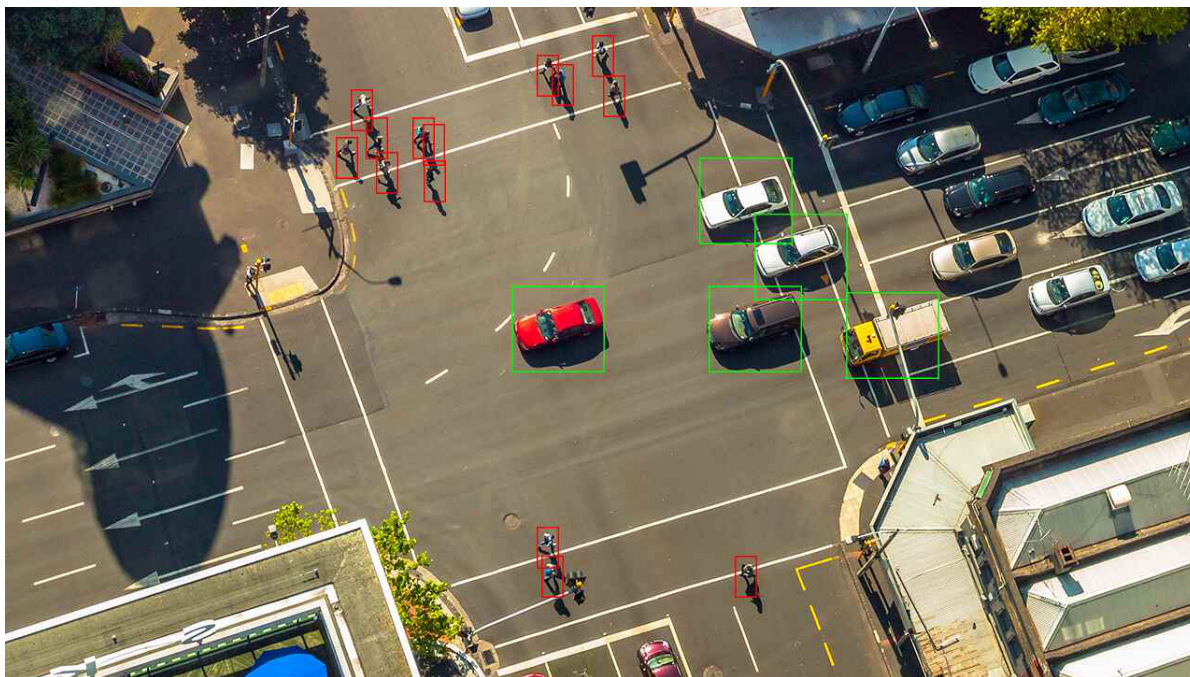*Technical Article*
What's the best type of computer vision for AI applications?

TEXAS INSTRUMENTS

Reese Grimsley

Computer vision refers to the technological goal of bringing human vision to computers, enabling applications from assembly-line inspection to driver assistance and robotics. Computers lack the ability to intuit vision and imagery like humans. Instead, we must give computers algorithms to solve domain-specific tasks.

This article explores artificial intelligence (AI) from the perspective of making computers "see" to perceive the world more like humans. I'll briefly compare each type of computer vision, especially in embedded systems that collect, process and act on data locally rather than relying on cloud-based resources.

What is computer vision?

In the 1960s, computer vision performed such tasks as reading text from a page (optical character recognition) and recognizing shapes such as circles or rectangles. Computer vision has since become one of the core domains of AI, which encompasses any computer system that perceives, synthesizes or infers meaning from data.

There are three approaches to computer vision:

- **Conventional computer vision** refers to programmed algorithms that solve tasks such as motion estimation, panoramic image stitching or line detection. Conventional computer vision uses standard signal processing and logic to solve tasks. The engineer selects functions that extract meaning from images by hand, and uses the resulting features within an algorithm that solves the task. Algorithms such as Canny edge detection or optical flow can find contours or vectors of motion, respectively, which is useful for isolating objects in an image or motion tracking subsequent images. Parameters that need calibrating for this task or environment are tuned by hand or through a supporting algorithm.

- **Computer vision with classical machine learning** requires an expert to "craft" the feature set on which the machine learning model is trained. Many of these features are common to conventional computer vision applications. Not all features are useful, thus requiring analysis to prune the uninformative ones; the machine learning algorithm is trained on these features to find patterns, which may be nontrivial to isolate by hand. Implementing these algorithms effectively requires expertise in image processing as well as machine learning.
- **Computer vision with deep learning** is machine learning, but on very large neural network models operating on largely unprocessed, "raw" data. Deep learning has had a significant impact on computer vision by pulling feature extraction operations into the model, such that the algorithm learns the most informative features without requiring expertise to hand-craft the feature set. Deep learning is even better able to isolate subtle patterns, yet has higher computational and memory requirements.

So which type of computer vision is best?

That ultimately depends on a few factors, outlined in Table 1. These are broad generalizations, where metrics such as accuracy and task complexity depend on the use case.

Table 1. Comparison of computer vision technologies

Aspect	Conventional computer vision	Classical machine learning	Deep learning
Accuracy	Moderate	Moderate	High
Expertise required	High	High	Low
Effectiveness on complex tasks	Low	Moderate	High
Computational intensity	Low	Moderate	High
Generality or simplicity of tuning to specific tasks or environment	Low; requires expert tuning	Moderate; tune with more data	High; tune with more data
Explainability	High	Moderate to low	Low to none
Sample or training data needed	Low to none	Moderate	High
Growth and research and development interest	Low	Low	High and accelerating

Computer vision with classical machine learning rests between the conventional and deep learning methods; the set of applications that benefit compared to the other two approaches is small. Conventional computer vision can be accurate and highly efficient in straightforward, high-throughput or safety-critical applications. Deep learning is the most general, the easiest to develop for, and has the highest accuracy in complex applications, such as identifying a tiny missing component during printed circuit board (PCB) assembly verification for high-density designs.

Some applications benefit from multiple types of computer vision algorithms in tandem, to cover each other's weak points. This approach is common in safety-critical applications with highly variable environments, such as driver assistance systems. For example, you could employ optical flow using conventional computer vision methods alongside a deep learning model for tracking nearby vehicles, and use an algorithm to fuse the results to ascertain whether the two approaches agree. If they do not, the system could warn the driver or start a safety maneuver.

An alternative is to use multiple types of computer vision sequentially. A barcode reader can use deep learning to locate regions of interest, crop those regions, and then use a conventional computer vision algorithm to decode.

Benefits of deep learning in computer vision applications

Compared to conventional computer vision and classical machine learning, deep learning has consistently higher accuracy and is rapidly improving, as it is immensely popular in the research, open-source and commercial communities. Figure 1 summarizes the difference in data flow from a developer's perspective for the three technologies.
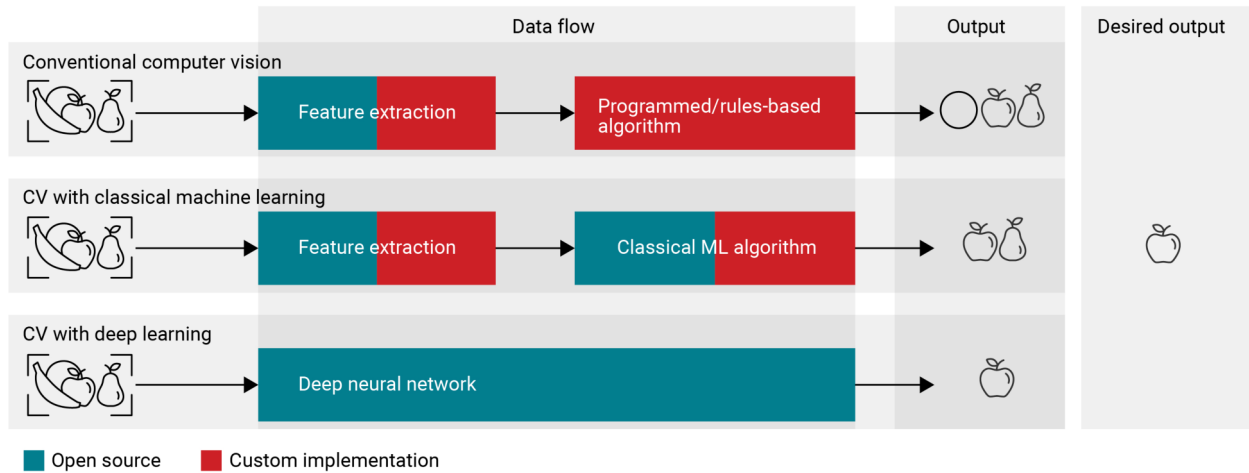
Figure 1. Data flow for each computer vision approach

Deep learning is computationally intensive. However, improvements in processing power, speeds, accelerators (such as neural processing units and graphics processing units), and improved software support for matrix and vector operations have mitigated the increased computational requirements, even on embedded systems. Microprocessors such as the AM62A7 leverage hardware accelerators to run deep learning algorithms at high frame rates.

Computer vision in practice

Processors in the TI AM6xA portfolio such as the AM62A7 contain deep learning acceleration hardware and supporting software for conventional and deep learning computer vision tasks. Digital signal processor cores such as the C66x and hardware accelerators for optical flow and stereo depth estimation also enable high-performance conventional computer vision tasks on processors such as the TDA4VM and AM68PA.

With processors capable of both conventional and deep learning computer vision, it's possible to build tools that rival sci-fi dreams. Automated shopping carts will streamline shopping; surgical and medical robots will guide doctors to early signs of disease; mobile robots will mow lawns and deliver packages. See TI's edge AI vision page to explore how embedded computer vision is changing the world.

IMPORTANT NOTICE AND DISCLAIMER