AI Magazine Volume 29 Number 1 (2008) (© AAAI)

Fish Inspection System Using a Parallel Neural Network Chip and the Image Knowledge Builder Application

Anne Menendez and Guy Paillet

■ *A generic image learning system, CogniSight, is being used for the inspection of fishes before filleting offshore. More than 30 systems have been deployed on seven fishing vessels in Norway and Iceland over the past three years. Each CogniSight system uses four neural network chips (a total of 312 neurons) based on a natively parallel, hard-wired architecture that performs real-time learning and nonlinear classification (RBF). These systems are trained by the ship crew using Image Knowledge Builder, a "show and tell" interface that facilitates easy training and validation. Fishers can reinforce the learning anytime when needed. The use of CogniSight has significantly reduced the number of crew members needed on the boats (by up to six persons), and the time at sea has been shortened by 15 percent. The fast and high return of investment (ROI) to the fishing fleet has significantly increased the market share of Pisces Industries, the company integrating CogniSight systems to its filleting machines.*

The fish industry is very competitive. Fleet owners are very interested in filling their boats as fast as possible with the fewest and most qualified personnel, thus reserving maximum occupancy for their refrigerated storage. During an expedition, which can last between one to two weeks, the fish processing machinery operates around the clock (figure 1). Typically, fishes are brought on the boat and dropped into metal pockets that convey them through cleaning, cutting, and filleting machines. Anomalies, which must be detected at the beginning of the chain, include a fish of the wrong species or a damaged fish. Such anomalies must be rejected immediately. In addition, the presence of more than one fish in a pocket or the improper orientation of a single fish must be detected quickly to avoid jamming the cutting or filleting machines. This type of real-time inspection is not easy to deploy with conventional image-processing tools since the size, shape, and scales of fishes are difficult to model mathematically. Their features can also change depending on the location of the expedition as well as the season of the year. Figures 2–4 depict well-positioned and damaged fish. Finally, and most importantly, the inspection system must be very easy for crew members to use, since there is no room on board for a software programmer to fix a software bug or change an image-processing algorithm.

Several attempts have been done by Pisces Industries to solve this problem with traditional combinations of cameras, frame grabbers, PCs, and image-processing software. None of these attempts have led to a usable offshore system because of the high nonlinearity of the problem.

A neural network approach was the only possible way to deliver a system that could be both adaptive and trainable by the fishers themselves. A hardware neural network was the best way to deliver a reliable and fast system that featured both a small footprint and affordable cost. Typical fish species to be recognized include ill-defined herrings or mackerels.

Silicon Neural Network Justification

Due to the highly variable aspect of a fish, mathematical modeling was not an option. In addition all the "by catch" (which are random species) must be rejected, but random species cannot be learned due to the infinite number of shapes and textures. In order for the inspection system to operate properly, the concept of "uncertain" and "unknown" was critical. In addition, due to exacting conditions at sea, providing reliable continuous operation within a minimum space and without mechanical components (such as a PC fan) was mandatory. Speed was also essential—360 to 600 fishes per minute. Computerized statistical analysis was not a viable solution because of the need for a small footprint, high speed, and low cost. In real-life situations, the key is not necessarily to have the absolutely best classifier but instead to have a solution that can solve

 ISSN 0738-4602

Figure 1. The Engey (Iceland).

This boat is operating seven CogniSight systems.

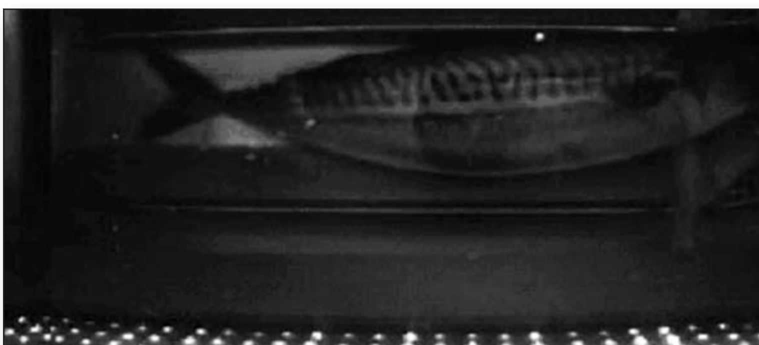

Figure 2. A Well-Positioned Mackerel.

Figure 3. A Well-Positioned Herring.

the problem at hand at a reasonable cost. In 2003 the ZISC[1] was the only available neural network chip that could meet these constraints. Its restricted coulomb energy (RCE) (Reilly, Cooper, and Elbaum 1985) feed-forward network offered a highly nonlinear classifier and allowed for unknown and uncertainty detection.

Learning and Reinforcement

An important feature of the system was its ability to refine the learning "on board," since a fishing vessel can stay out for more than two weeks without ever going ashore (figure 5). An additional requirement was the ability to adjust the "throughput versus accuracy" of the recognition engine by increasing or decreasing the severity on the fish's quality depending upon the situation or season. This was achieved in CogniSight by editing a single parameter, the value of the maximum influence field of the neurons, which controls their level of conservatism during the learning process. The higher this value, the more liberal is the recognition; the smaller the value, the more conservative is the recognition.

The Solution

The CogniSight system (figure 6) is composed of a vision sensor, a silicon neural network, and a recognition engine on a field-programmable gate array (FPGA) acting as the glue logic that extracts a feature vector from each video frame and reads the response of the network after the broadcast of each signature. CogniSight is mounted in a waterproof enclosure above the pocket conveyor just before the filleting machine.

The autonomy of the recognition relies on the parallel neural network, which is capable of learning by example and generating models automatically. The neural network can recognize patterns that are identical or similar to the models stored in the neurons. It produces three types of responses: (1) a status response that indicates whether the classification is identified, uncertain, or unknown; (2) a global response, which is the category of the first neuron with the best match, if any, to an existing model (that is, with the smallest distance to the input vector); and (3) a detailed response that is the category and distance value of all the "firing" neurons read in increasing order of distances. The detailed response of the network can be useful in building hypothesis and leveraging uncertainties. In the case of the fish inspection, the global response has proven to be sufficient because the teaching was done easily but thoroughly on many examples, thanks to the Image Knowledge Builder tool that was delivered with the system. As a result, the system produces few uncertainties, and the best match response gives 98 percent accuracy.

On the filleting line, the number of classifications is limited to four categories: accept, reject, recycle, or empty. However, categories can include multiple visual criteria. An "accept" is a pocket containing a fish of the right species (herring for example) that is not damaged and is in the proper position to enter the filleting machine. A "reject" is a fish of the right species but damaged or a fish of the wrong species. A "recycle" is a pocket showing a fish of the right species, not damaged, but improperly oriented. A "recycle" might also be a pocket with more than one fish of the correct species. The idea of the "recycle" category is to eject the fish on a vibrating table so it goes back to the beginning of the line to be dropped into a new pocket. When teaching the system, crew members must know only to which category a fish belongs. They do not have to worry about describing the rules for a decision. They simply have to reinforce learning by clicking the appropriate button on the touch-screen panel if they see that the camera is making mistakes (figure 7).

The ease and freedom with which tutoring is done was a key selling point of the system. The neural network learns the examples and builds the decision space accordingly, whether highly nonlinear or not. The possible drawback of too much freedom in tutoring is that a fish that looks damaged to one fisherman might pass as acceptable to another fisherman (or the same fisherman, but on another day). Such contradictions deteriorate the knowledge by creating "degenerated" neurons. To circumvent this risk, CogniSight is delivered with a software tool for training and validation that allows teaching on many images collected with the sensor and reviews the consistency of the recognition. This software is called Image Knowledge Builder. It features a very simple and practical user interface that can test the accuracy and throughput of the recognition on many images before loading it into the sensor.

Image Knowledge Builder runs a simulation of the silicon neural network. It can save the contents of the neurons into an image knowledge file format readable by CogniSight. This file transfer is equivalent to a knowledge transfer. Once the silicon neurons are loaded with reference patterns and associated categories, CogniSight can then execute the same recognition as Image Knowledge Builder, only at full speed on live images with no connection to a PC. The recognized category is transmitted to output lines that indicate whether the content of the pocket is an "accept," "reject," "recycle" or "empty." This signal is sent to a programmable logic controller (PLC), which itself controls two brushes sending the fishes to a reject or recycle bin as applicable.

CogniSight can be connected to an Ethernet local-area network installed on the boat for four

Figure 4. A Damaged Mackerel to Be Ejected.

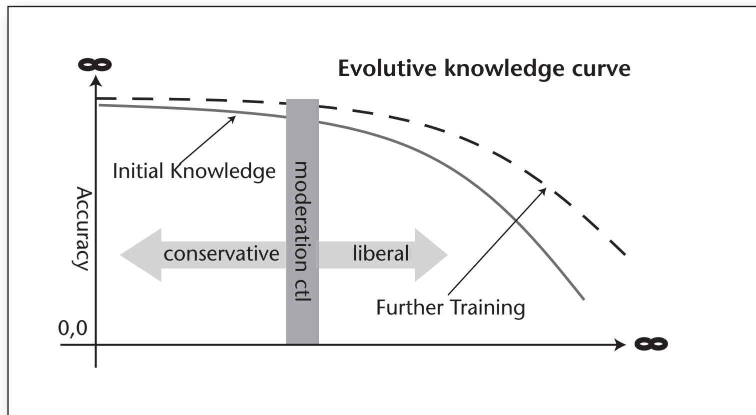

Figure 5. Classification Moderation in CogniSight.

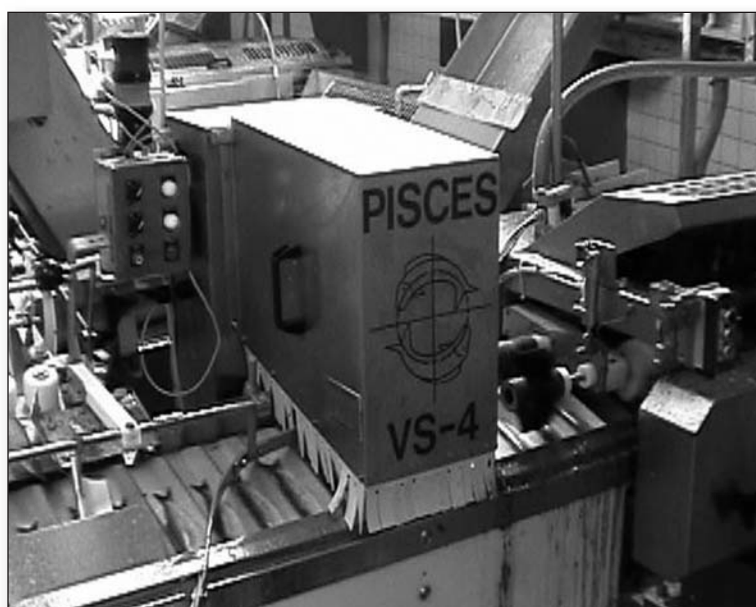

Figure 6. The CogniSight System Installed on the Filleting Line.

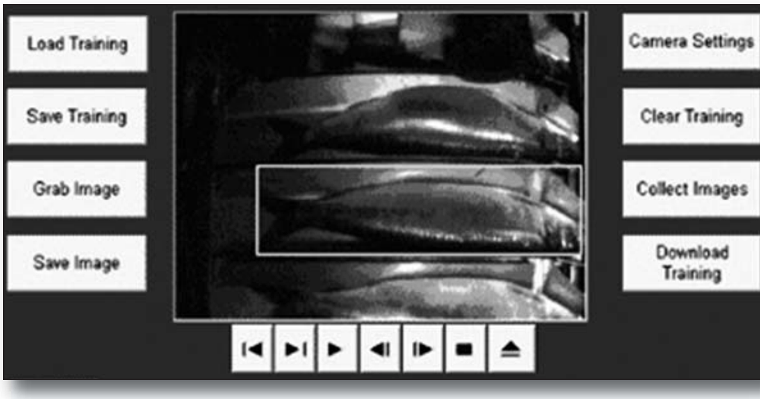

Figure 7. Control Panel for Online Learning and Additional Tools.

types of operations: (1) view the images on screen to adjust the camera settings at the time of installation; (2) collect images to proceed with the training of the neural network with Image Knowledge Builder; (3) load an image knowledge file (IKF file) into the neural network of the camera; and (4) view statistics accumulated on the camera and report the number of acceptable and nonacceptable fishes

Use of AI Technology

Feature extraction is executed by a field-programmable gate area (FPGA), which converts the pixel values received from the charge-coupled device (CCD) sensor (640 x 480 pixels) into a feature vector of 256 bytes. It is calculated over a region of interest specified in the video frame, which can range from 16 x 16 pixels to the full frame. This transform is a spatial and gray-level integration based on a "best fit" function of 256 blocks within the selected region. The resulting 256 components are transmitted to the neurons, and the response is read back before the next frame starts (figure 8).

RBF Classifier with Automatic Model Generator

The incoming feature vector is broadcast to a hardware neural network of 312 neurons. This network is composed of four semiconductor chips (ZISC) with 78 neurons per chip connected in parallel. The neurons "react" to the incoming feature vector by evaluating the similarity with their reference vector (stored in their memory during the training). Their network implements a radial basis function (RBF) classifier, derived from the compound classifier (Batchelor 1974) and the RCE (figure 9).

When committed (or trained), a neuron evaluates whether an incident pattern is similar enough to its stored pattern (for example, a model) by calculating an $L1$ distance. If similar, it will output its response to the global response bus. The similarity domain of a neuron is self-mapped during the training process and does not require manual adjustment. If many neurons respond to the same stimuli, a winner takes all (WTA) patented scheme allows retrieval, in a fully parallel manner, of the best responses of these neurons. The classification can then be determined as either positively identified or uncertain. If all the neurons firing with the same smallest distance value are in agreement with the identified category, the classification is said to be "identified." If, on the contrary, these firing neurons return different category values, the classification is possible, but with some level of uncertainty. Uncertainty can be waived using different methods, but in the case of the fish inspection, a conservative training scheme has allowed good accuracy by simply reading the response of the first neuron on the list (figure 10).

Generation of Portable Knowledge

The neural network embedded in CogniSight is easily trained using the Image Knowledge Builder software. This application features batch utilities to automatically extract feature vectors from annotated images and broadcast these features to the neural network repetitively until the knowledge is stable. Stability is achieved when the learning of the vectors no longer creates any new neuron. Reporting utilities help identify difficult or complex classifications, pinpoint possible erroneous or inconsistent annotations, and evaluate compromises between throughput and accuracy (figure 11).

An image knowledge file (IKF) can either be stored for use on the same boat during a next fishing trip or transferred for use on other CogniSight systems installed on other boats running the same type of expeditions. Currently, Pisces has installed systems on a fleet fishing for herrings. The knowledge file for these systems has been built over several expeditions taking place throughout the year. The accumulated knowledge can be labeled as "year-around" knowledge. It is composed of fewer than two hundred models and has already inspected several million herrings. In addition, when new conditions are encountered during a fishing campaign, the crew can perform reinforcement learning to refine the knowledge.

Introducing Silicon Neural Networks (SNN)

Neural networks have been extensively discussed since their appearance in the late 1980s. DARPA's (DARPA/TTO 1988) recommendation after conducting an extensive survey at that time was to "go silicon." As neural networks are inherently a group of elements with the same basic behavior, they are indeed candidates for massively parallel architecture. While it has been claimed by many that neural networks benefit from being parallel, most development has been done on standard comput-

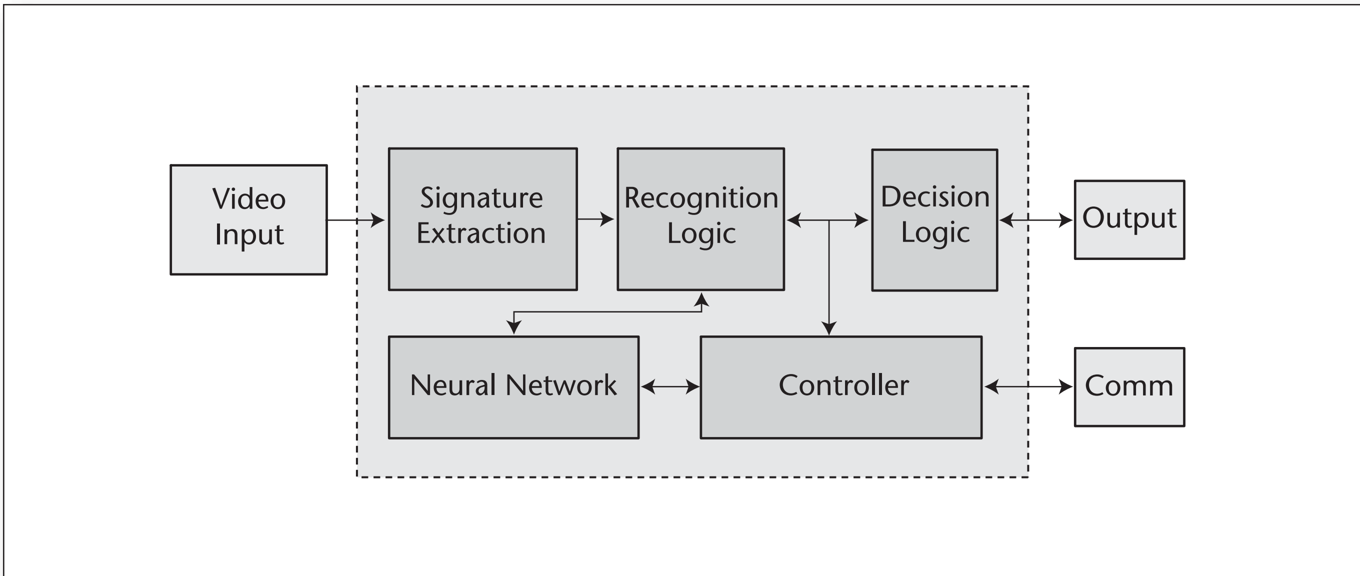

Figure 8. CogniSight Image-Recognition Engine Architecture.

ers, which suffer three basic limitations. First, they execute one instruction at a time (sometimes four, if the computer has quad cores). Second, a good part of their data bandwidth is dedicated to fetching and decoding instructions before actually executing them. Third, data is routed through a single bottleneck: the memory bus, which, in most cases (except for Harvard RISC), is also the vehicle for the instructions. As a result, a neural network implemented by software on a standard computer cannot be defined as parallel.

Going forward using the massively parallel architecture concept demonstrated that, in the case of multiple programmed processors, synchronization between them could become a serious hurdle. The way to overcome these limitations was to design a neuron entity with all the "genetic" material to learn and recall without the need of running program code. In addition, this architecture would have to be fully distributed (no supervising unit) and have theoretically unlimited expansion capability. Another constraint was to have a fast and constant learning recognition time, regardless of the number of connected neurons. This was achieved by the ZISC architecture, which was described by the coauthor and jointly developed and patented[2] with IBM. The first SNN was the ZISC36 in 1993, followed by the ZISC78. Connections of up to 5,000 neurons (on a multiple peripheral connect interface [PCI] board architecture) were demonstrated. Today the CogniMem chip (that is, cognitive memory), successor of the ZISC, contains 1024 neurons in parallel with a memory size extended from 64 bytes to 256 bytes. The cascadibility of multiple chips is easily made through a bus of 27 wires, and we can envision the

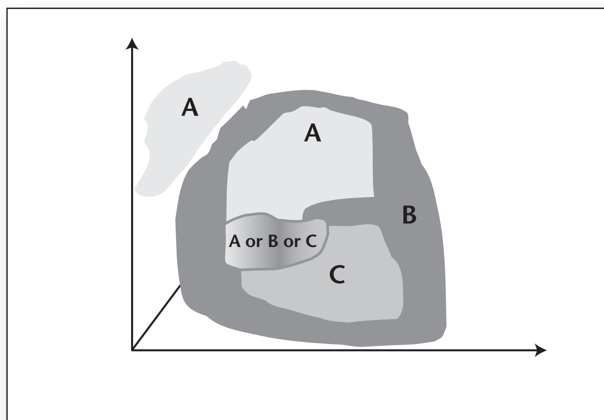

Figure 9. RCE Decision Space Mapping.

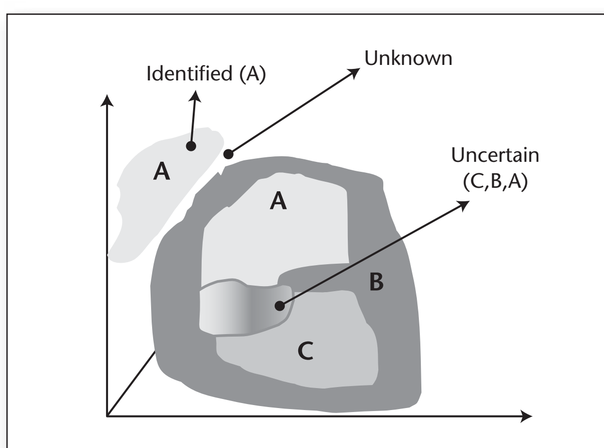

Figure 10. Decisions into Feature Space.

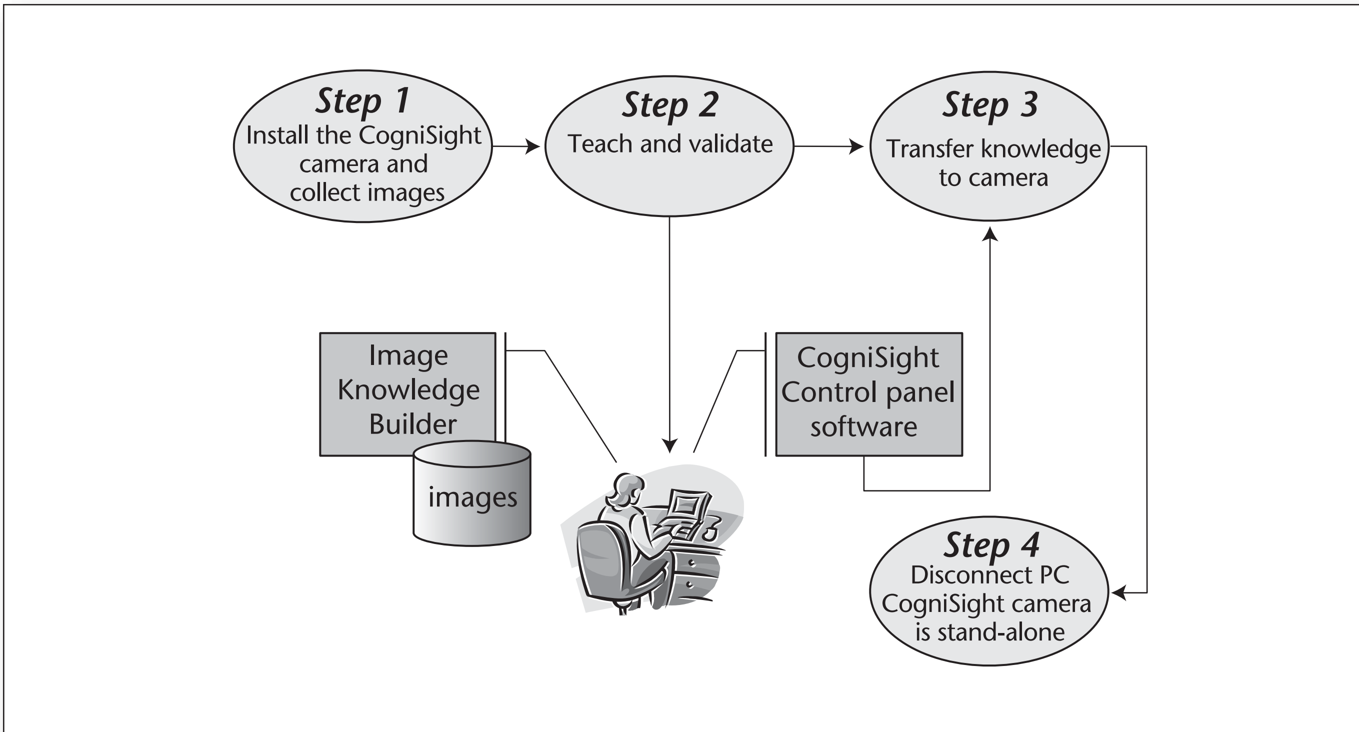

Figure 11. Methodology Flow Chart.

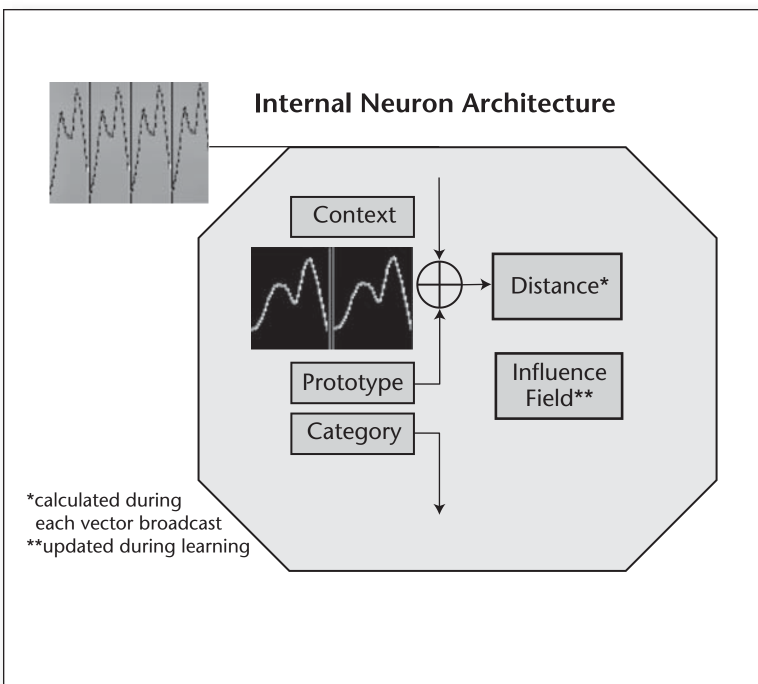

Figure 12. Neuron Structure.

design of a system with 10,000 neurons fitting in a Rubix cube.

The Parallel Digital Neuron

The digital neuron consists essentially of a memory storing the learned pattern prototype (or kernel), a hardwired distance evaluation unit, a learning logic, and an associative logic. Depending upon its content, the neuron has three different states: (1) idle (does not participate), (2) ready to learn (next in line to learn a pattern), and (3) committed (learned a pattern associated to a category, and has an influence field) (figure 12).

Neuron Memory: While the first neurons used 64 bytes of memory, advances in semiconductor technology now allow 1024 neurons with 256 bytes each. The information the new 256 bytes of memory contains is indeed richer than previously and tends to reduce the number of neurons needed for a given problem. For the inspection of herrings, the original knowledge was composed of 300 neurons of 64 bytes. That has been reduced to 120 neurons of 256 bytes and achieves a similar accuracy. The learned pattern is called a *prototype,* as it is a significant representation of one model of the population.

Distance Evaluation Unit: The distance evaluation unit computes the *L*1 distance (accumulation of absolute differences) between the incoming vector (up to 256 components) and the stored pattern. This occurs in parallel for all the committ ed neurons at each feeding of a vector component.

Associative Logic: The associative logic triggers the output of the category if the evaluated distance falls into the influence field of the neuron.

Learning logic: The learning logic enables a committed neuron to autonomously reduce its influence field (generalization capability) to accommodate the creation of a new neuron if applicable. This is self-adaptation. If no committed neuron identifies the taught category, the ready-to-learn (RTL) neuron automatically commits and adjusts its influence field to the distance to its closest neighbor (K-nearest neighbor [KNN]). All these operations occur inside the neuron and are not under control of an external logic.

Network behavior: While each neuron is fully independent during the learning and the recognition process, all the neurons can "see" the global results. The "search and sort" patented method allows it to find the closest distance, "winner takes all," among all the responding neurons in 19 clock cycles (that is, less than one microsecond) regardless of the number of neurons. This provides the unique ability to learn without the need of program instruction. The hardware network topology is illustrated in figure 13.

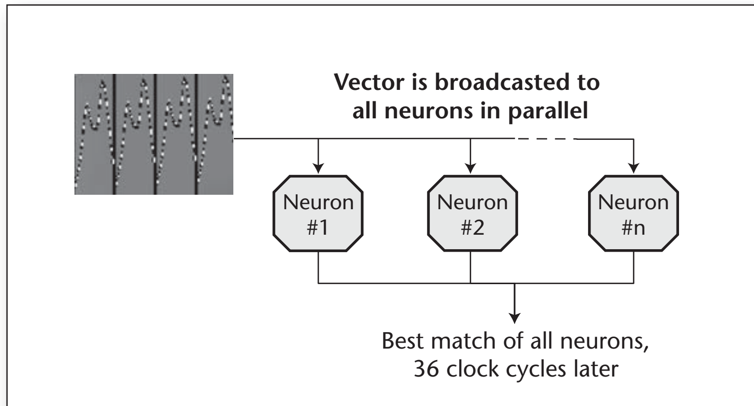

Figure 13. Hardware Network Topology.

Use and Payoff

Pisces Industries,[3] manufacturer of fish processing equipment, has presently installed more than 30 systems on five different fleets in Norway, Iceland, Scotland, and Denmark. So far, most expeditions have been for herrings and mackerels. The camera inspects at a speed of 360 pockets per minute on the herring lines, but can go faster. An accuracy of 98 percent was verified for the classification of 16 tons of fishes with knowledge based on 80 neurons. Inspection is a tedious job for a human operator, especially considering that he or she must also supervise multiple noisy heavy machinery stations such as feeders, ejectors, vibrating tables, the filleting machine, and so on. The use of the CogniSight systems for inspection has helped shorten the duration of the expeditions. As a result, a boat can fill its cargo in five days instead of seven days. Fishers appreciate these shorter expeditions, and also don't mind sharing bigger catches with fewer coworkers.

Maintenance and Upgrades

Over the past three years, no maintenance has been requested. Tuning of the image recognition for different types of expeditions has been handled by the crews. Except for possible electrical failure of the camera, the only serious problem envisioned is an insufficient number of neurons. This might occur if the fleet decides to teach a more complex or extended image knowledge file, such as one capable of sorting multiple species during the same expedition or an image knowledge file sorting damaged fishes into different subcategories so they can be diverted towards different recycling processes. If and when this occurs, the new generation of CogniSight cameras (figure 14) has a flexible architecture that allows neuron expansion cards to be added at will, in increments of 1024 neurons. Existing knowledge files can be loaded on the new model and expanded as long as neurons are available. Figure 14 depicts a new trainable sensor.

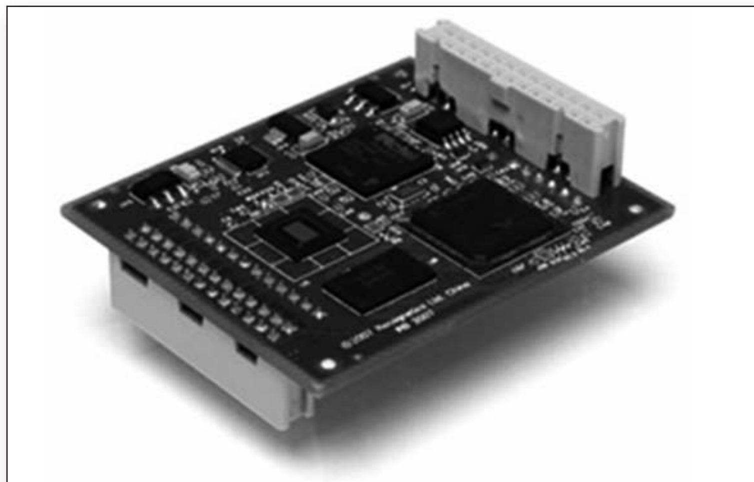

Figure 14. New CogniSight Miniature Trainable Sensor.

Conclusion and Perspectives

It has been proven with this application that AI can provide high return on investment for small entities unrelated to the field. The evolution of hardware neural networks toward a 1000-neuron chip allows significant cost and size reduction for the present installation. At this time, AI has become a "commodity" for Norwegian fishers. The size and cost reduction of the next CogniSight systems should enable additional control points.

The key to the widespread use of these systems is cost reduction (to less than US$500), robustness (the ability to be deployed in harsh and changing conditions), and ease of training (of the operator of the machine). The hardware neural network allows for a dramatic footprint reduction (close to the size of a matchbox) providing the speed of multiple workstations in parallel. After this first success story, there is a strong possibility that the CogniSight technology connected with Image Knowledge Builder will make a significant contribution toward turning AI into a commodity in many domains related to vision machine learning.

Notes

1. ZISC (zero instruction set computer) is a registered trademark of IBM Corporation.
2. U.S. patent numbers 5,621,863; 5,701,397, 5,710,869; 5,717,832; 5,740,326; 6,606,614; European patents 694854; 694853; 694856; 694855; and equivalent patents in Japan, Canada, and Korea.
3. Pisces Industries, www.pisces-ind.com.

References

Batchelor, B. G. 1974. *Practical Approach to Pattern Classification.* New York: Plenum.

Reilly, D. L.; Cooper, L. N.; and Elbaum, C. 1985. A Neural Model for Category Learning, *Biological Cybernetics* 45: 35–41.

Tactical Technology Office, U.S. Defense Advance Research Projects Agency (DARPA/TTO). 1988. *DARPA Neural Network Study.* Washington, D.C.: Government Printing Office.

Anne Menendez (anne@general-vision.com) is the president of General Vision, Inc., a company that she founded in 1987. After deploying machine vision applications mostly based on LabVIEW imaging tools, she has refocused the company in 1999 towards the use of neural networks for image recognition. She is now developing the CogniSight technology with possible implementation in software, FPGA, and ASIC. Menendez graduated from ESIEE Paris and brings significant expertise in advanced vision algorithm, user interface, and HDL logic design.

Guy Paillet (guy@general-vision.com) is the chief executive officer of General Vision, Inc. He joined the company in 1999. Paillet is the inventor of the 1993 silicon neuron technology ZISC (zero instruction set computer), copatented with IBM. While in France he developed high-speed target tracking and recognition embedded cameras using the ZISC made by IBM. Paillet brings a long experience in hard-wired artificial vision systems and parallel pattern recognition architecture to the company.