# Deep Vision Becomes Reality with CNN

### Neural networks execute complex image processing tasks

Deep learning with neural networks will strongly influence the future of image processing, as this approach provides a number of significant advantages regarding classification and analysis results as well as final image quality. Since small neural networks suffice for many typical vision applications, processors such as FPGAs can be implemented effectively for convolutional neural networks (CNNs). This results in a wide application field far beyond current classification tasks and efficient use within embedded vision systems is also possible.
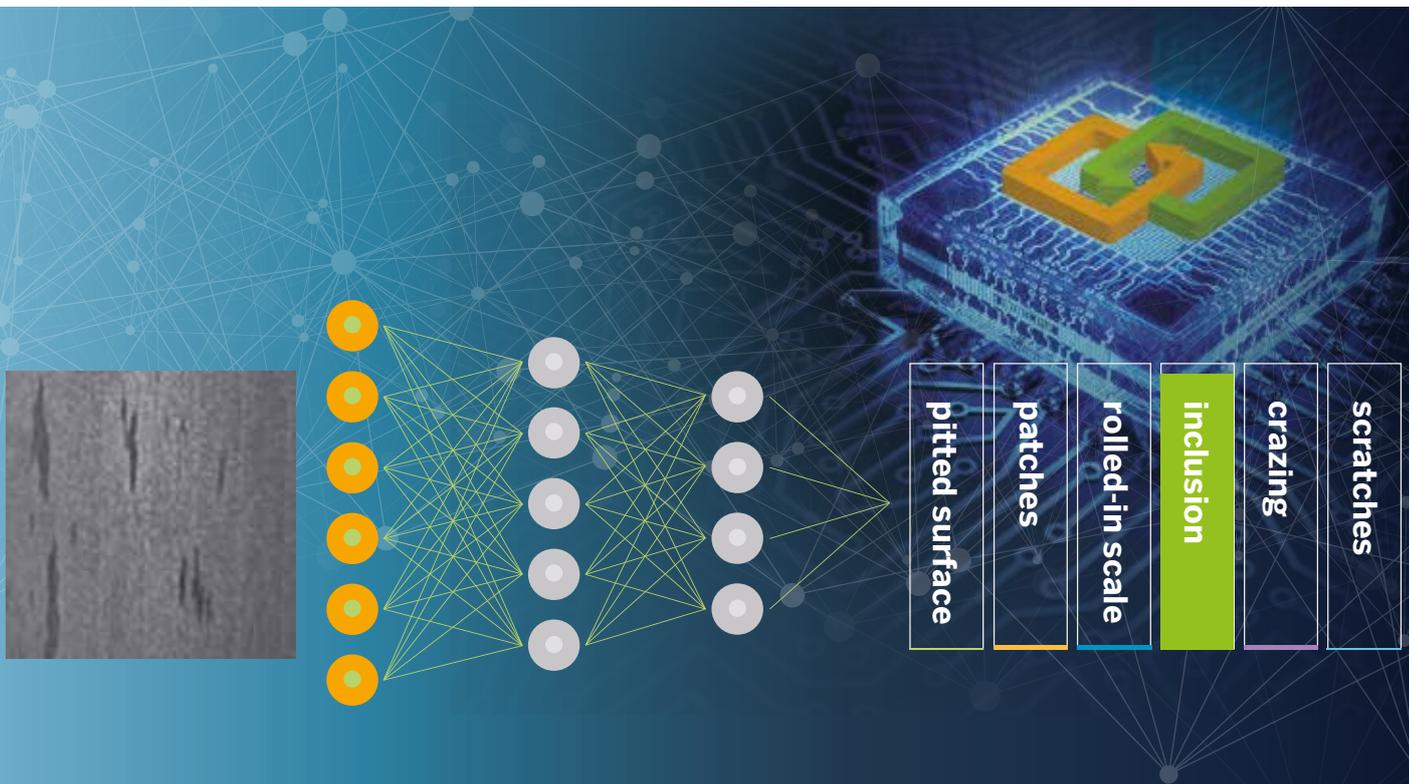
Source: Silicon Software GmbH

Classic image processing applications reach their limits when test objects are deformed and reveal irregular shapes or large object variations, if the lighting enivronment is not suitable or if lens distortions are present. If, as in these cases, the framework conditions for the image acquisition cannot be controlled, even individual algorithms for feature description are often barely possible. CNNs, on the other hand, define characteristics through its training method, without using mathematical models. This makes it possible to capture and analyze the image in difficult situations, such as reflecting surfaces, moving objects, face detection, robotics, as well as easier classification
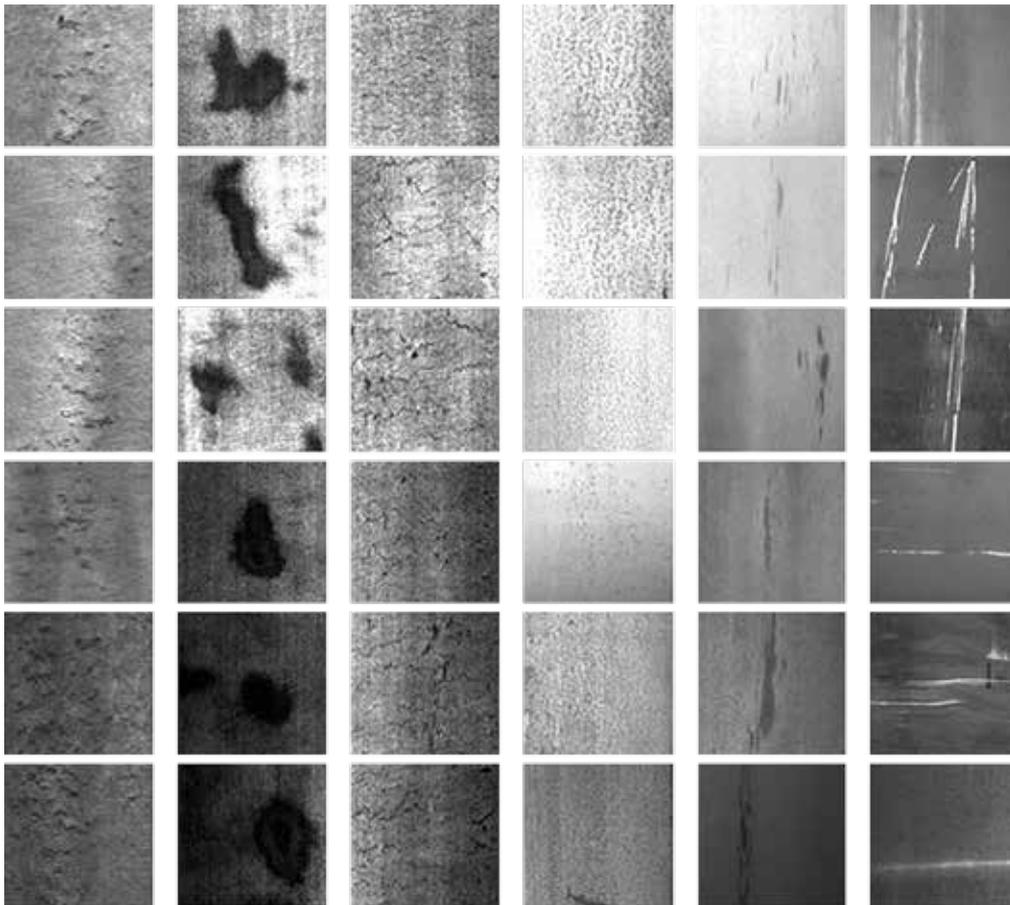
Detected error classes (from left to right): rolled-in scale, patches, crazing, pitted surface, inclusion, scratches/ Source: NEU Surface Defect Database

of image data directly from the preprocessing to the classification result, which is required in many applications. Nevertheless, CNNs cannot cover all areas of classic image processing, such as precise location determination of objects - here new and advanced CNNs must be developed.

Optimized CNNs accelerate vision

Practical experience with CNNs from prior years led to mathematical assumptions and simplifications (pooling, ReLu and overfitting avoidance... to name a few) which led to a reduction in computational expense, which in turn enabled implementation of deeper networks. By reducing image depth at the same rate of detection and by optimizing the algorithm, CNNs can be significantly accelerated and are now ideally suited for image processing. CNNs are shift and partially scale invariant thus allowing the use of the same network structures for different image resolutions. Smaller neural networks are often sufficient for many im-age processing tasks.

Due to the high degree of parallel processing, neural networks are particularly well suited to FPGAs (Field Programmable Gate Array), upon which CNNs also analyze and classify high-resolution image data in real time. In machine vision, FPGAs function as massive accelerators of image processing tasks and guarantee real-time processes with deterministic latencies. Until now, the high programming effort and the relatively low resources available in an FPGA hindered efficient use. Algorithmic simplifications now make it possible to construct efficient networks with high throughput rates in an FPGA.

New CNN operators and frame grabber in development

To implement CNNs on FPGA hardware platforms, the VisualApplets graphical environment can be used. The CNN operators in VisualApplets allow
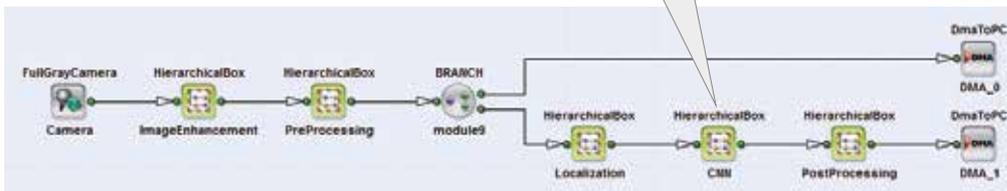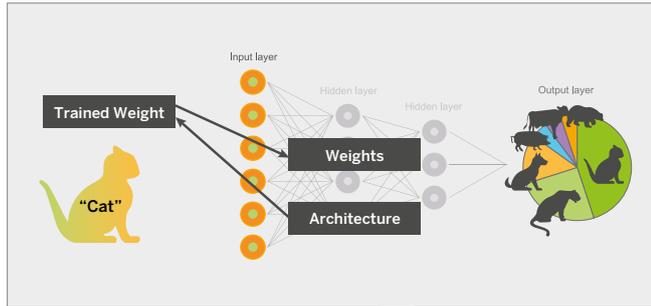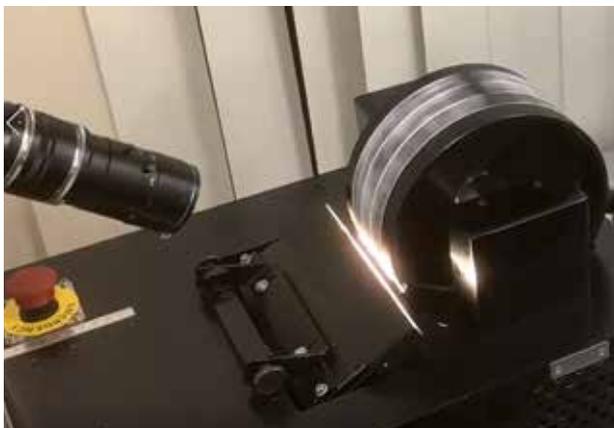
CNN integrated into VisualApplets/ Source: Silicon Software GmbH

users to create and synthesize diverse FPGA application designs without hardware programming experience in a short time. By transferring the weight and gradient parameters determined in the training process to the CNN operators, the FPGA design is configured for the application specific task. The operators can be combined into a VisualApplets flow diagram design with digital camera sources as image input and with further image processing operators for optimizing image preprocessing.

To implement especially large neural networks for complex CNN applications, a new programmable frame grabber in the microEnable marathon series is being released that has a 2.5-times more FPGA resources compared to the current flagship marathon series, and is expected to be ideal for neural networks with more than 1GB/sec CNN bandwidth. The CNNs are not only capable of running on the frame grabbers' FPGAs but also on VisualApplets compatible cameras and vision sensors. Since FPGAs are up to ten times more energy-efficient compared to GPUs, CNN-based applications can be implemented particularly well on embedded systems or mobile robots with the required low heat output. In future, the application diversity and complexity with neural networks based on newly developed special processors will further increase. In the development of new hardware and software solutions and for the exchange of research results, a cooperation with Professor Michael Heizmann from the Institute for Industrial Information Technology (IIIT) at the Karlsruhe Institute of Technology (KIT) is part of the FPGA-based "Machine Learning for Industrial Applications" project. New advances are expected from here in the future.
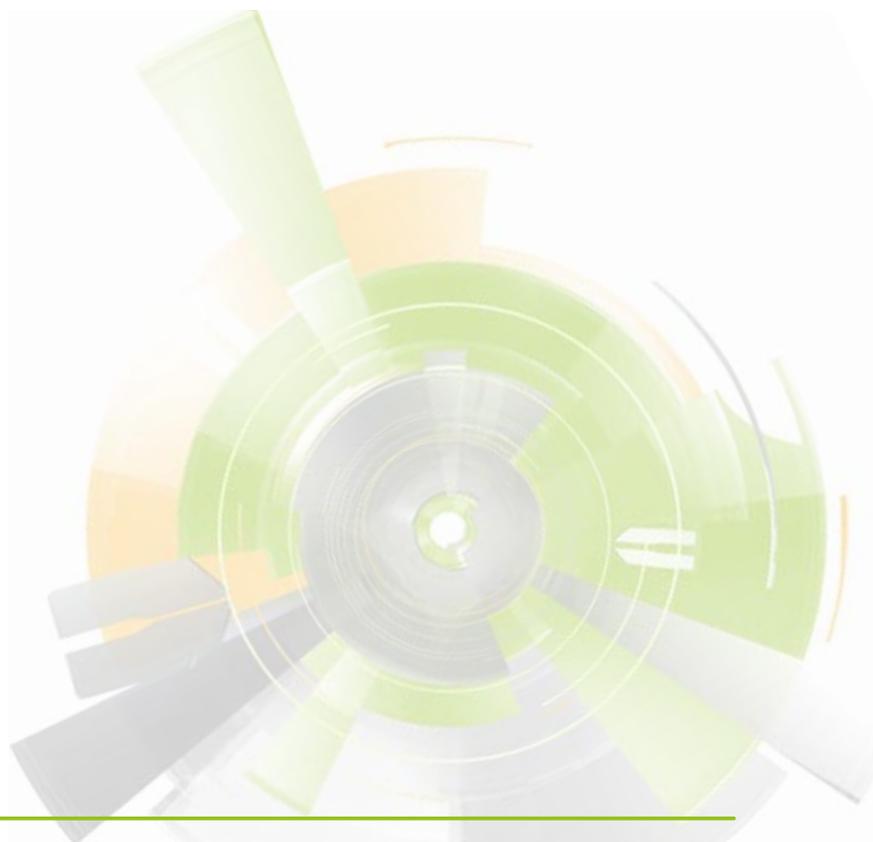
In order to determine the percentage of correctly detected defects in difficult environmental conditions, a neural network with 1,800 reflective metallic surface images was trained on which six different defect classes were defined. Large differences in scratches coupled with small differences in crazing, paired with different surface grey tones from lighting and material alterations, made the analysis of the surface almost impossible for conventional image processing systems. The test results demonstrated that the different defects were positively classified by the neural network an average of 97.4%, which is a higher value compared to

Demonstrator for classification of defects on metallic surfaces by CNN realized with VisualApplets/ Source: Silicon Software GmbH

classic methods. The data throughput in this application configuration was 400MB/sec. By comparison, a CPU-based software solution achieved an average 20MB/sec.

The implementation of deep learning on FPGA processor technology, which is ideally suited for image processing, is an important step. Requirements that the method must be deterministic and algorithmically verifiable, however, will make entry into all areas of image processing more difficult. Likewise, options to document which areas were identified as errors, as well as segmentation and storage, have not yet been implemented.

Outlook: Thus far, training and operational use of CNNs are two separate processes. In the future, new generations of FPGAs with higher resources or using powerful ARM / CPU and GPU cores will enable on-the-fly training on the newly acquired image material, which will further increase the detection rate and will simplify the learning process.

Deep Vision Becomes Reality with CNN