Artificial Intelligence in Image Processing: Deep Learning Compared with Conventional Methods

In the area of Machine Vision, deep learning-based methods have been established as part of the standard portfolio offered by most vision system providers. Many manufacturers dedicate a large part of their development capacities to the expansion of their deep learning-based features. Have the “conventional” methods used for image processing and image analysis become obsolete, and will they eventually be completely replaced by deep learning-based functions?

This White Paper presents an overview of the possibilities in industrial image processing currently offered by deep learning-based artificial intelligence, what the advantages and disadvantages are compared to conventional methods with and without artificial intelligence and what roles the different methods might play in this area in the near future.

The term “conventional methods” includes simple, rule-based methods in image analysis as well as methods through which image data are analyzed with the aid of machine learning algorithms that are not based on deep learning.

Content

1. Application areas of deep learning in industrial image processing

The typical application areas of deep learning are anomaly detection, image classification, image segmentation and object recognition. Well-trained artificial neural networks (ANN) are characterized by higher precision and greater flexibility compared to conventional image analysis methods.

- Anomaly detection refers to the division of entire images into “good” or “bad”, e.g. the ability to detect faulty objects and structures.

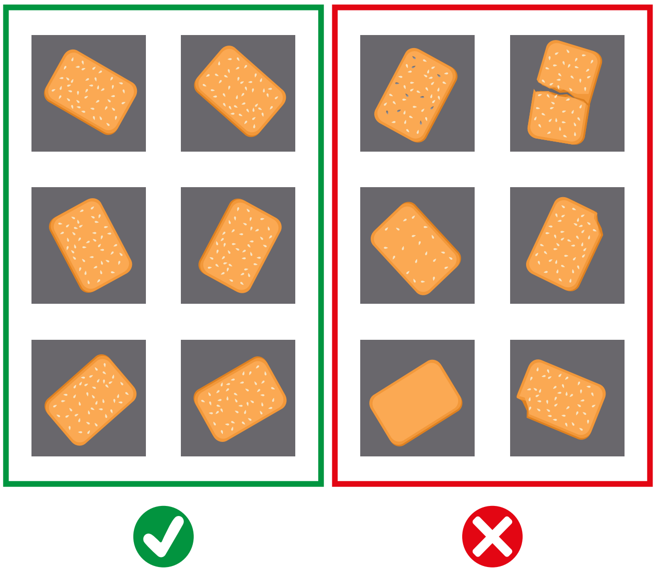

Figure 1 – Anomaly detection and classification

- Image classification refers to the classification of entire images into defined classes, e.g. assigning images that contain only a single object.
- Image segmentation refers to a pixel-by-pixel classification and can be used to detect and localize faulty surface structures, for example.

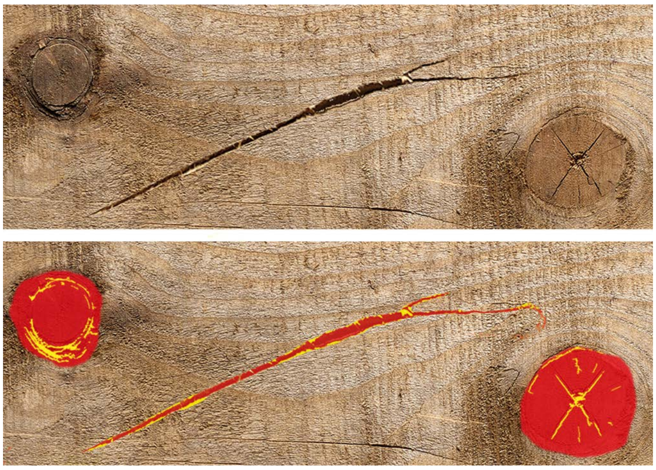

Figure 2 – Image segmentation

- Object recognition comprises the classification and localization of one or multiple objects in an image, for example to count particular objects or determine the position and orientation of certain objects.

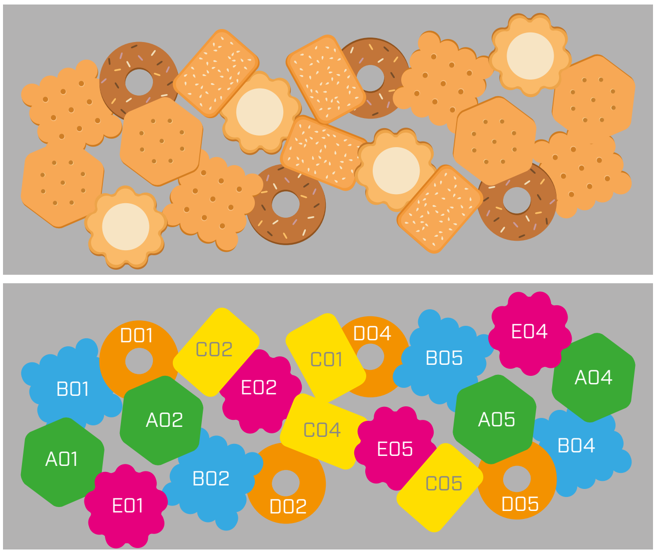

Figure 3 – Object recognition

2. Advantages of deep learning-based methods

In addition to improved accuracy, deep learning-based methods offer Machine Vision applications a robustness and flexibility that would not be feasible with traditional methods, or only with great effort. Compared to traditional machine learning methods, the step of feature engineering is eliminated. With feature engineering, a machine learning algorithm for the detection of relevant properties is set up manually before it can be trained to detect entire objects. What distinguishes deep learning networks is that they independently learn the relevant properties during the training.

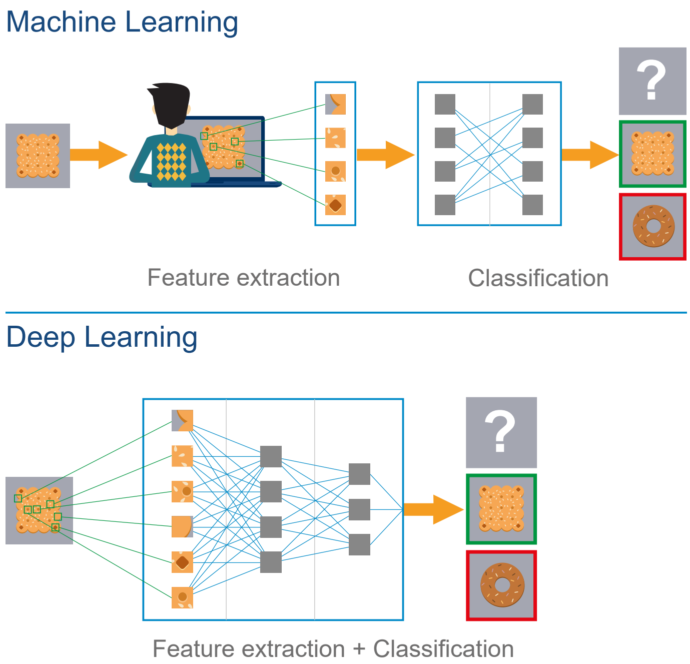

Figure 4 – Deep learning without feature engineering

Detection of varying objects and features: An advantage of deep learning methods is the ability to detect objects or properties that occur in varying shapes, such as scratches on surfaces, differently shaped natural products or handwriting. For anomaly detection, for example, it is sufficient to train a neural network only with images of flawless objects in order to be able to identify damages to objects during use.

More robust object recognition:Objects in variable environments, for example those with changing backgrounds, in different resolutions or varying lighting, can be recognized well with appropriately trained neural networks without the need to individually train and adjust them for every possible specialization.

3. Costs of using deep learning

In order to benefit from the advantages of purely deep learning-based methods in an application, compared to conventional methods, additional investments in the hardware used, and a high number of sample images to train the neural network are usually required.

Additional hardware: Complex and powerful deep learning architectures are distinguished by a high number of neural network layers. Executing a deep learning application thus requires a large memory and computing capacity. Only by outsourcing the computing tasks to an additional processor for simultaneous executions of the computations – for example on GPUs (graphic cards) – is it generally possible to achieve an acceptably short processing time.

Heat development, power consumption: The high required computational power of large neural networks increases an application's power consumption and thus the heat generation. This can be problematic in particular for embedded systems, which are often expected to be optimized for low power consumption and heat management.

High amount of training data: The robust recognition of objects requires a large number of training images, which depict and mark all of the objects and properties to be recognized in as many different variations and environments as possible. The greater the amount of different image data used for the training, the more easily a deep learning network can learn to recognize objects. Acquiring hundreds and sometimes even several thousands of required images often creates a difficult task in the development of a Machine Vision application.

4. Optimization of deep learning networks through a hybrid approach

Image processing in artificial neural networks can be combined with conventional methods thereby applying a hybrid method to process the image data. Conventional methods first preprocess the image. An artificial neural network then delivers the desired results with the preprocessed data. The advantage: High performance with low memory and power requirements.

Additionally, this hybrid approach makes it possible to reduce the number of images that must be acquired to train a deep learning network.

Optimizations can be achieved with conventional methods by reducing the image size, minimizing the variances of the inspected images or increasing the number and variance of the training data via simple means (keyword augmentation).

4.1 Reduction of the image size

Reducing the resolution of the input images means reducing the data volume to be processed. A neural network can thus work with fewer layers, requiring less memory and computational effort. Cutting down the relevant image section (ROI--region of interest) can result in a lower resolution with the same depth of detail so that only the selected part of the image is processed further.

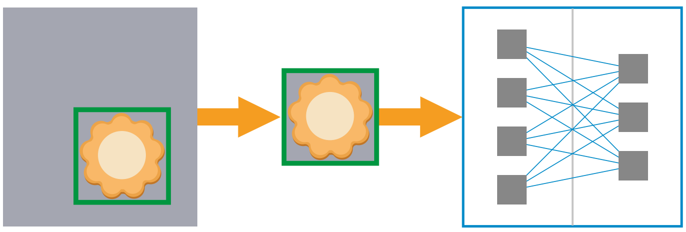

Figure 5 – Reduction of the image size

With a homogeneous background, the relevant image section can often be detected through simple color or grayscale comparisons.

4.2 Minimizing the variances of the inspected images

If the avoidable differences to the inspected images are reduced, less image data is needed to train a neural network. Static ambient conditions such as stable lighting, a consistent monochrome background, fixed positioning of the inspected objects and unchanging orientation make it easy to detect objects and structures with a relatively low training effort.

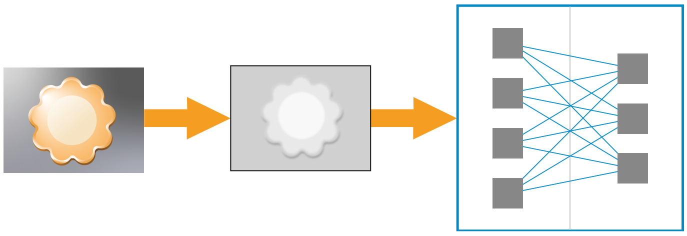

Figure 6 – Minimizing the variances

By removing irrelevant information with the help of simple filters and contrasting, conventional methods can additionally help a deep learning network to accurately recognize certain objects and structures with less training data and a smaller network.

4.3 Increase in the number and variance of the training data

The quantity and variance of the available training data can be increased by generating additional training data, or so-called augmentation. For such augmentation, new variants of the original images are created from existing images with the help of conventional methods, for example by rotating, translating or skewing the original image. This allows the neural network to learn how to recognize an object independently of its orientation and positioning in the image, for example.

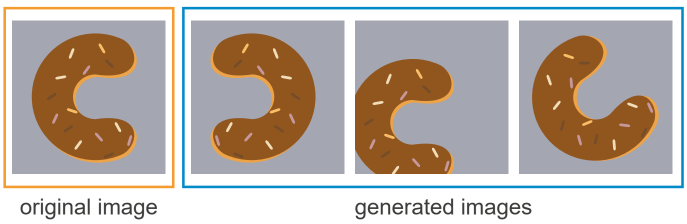

Figure 7 – Augmentation

5. Application areas for solutions without deep learning

In many use cases, deep learning-based methods offer no additional benefits that would justify the added costs. It would make no sense, for example, to set up and train a deep learning network only to differentiate between red and green objects. This might just require a simple rule-based color comparison, which would also most likely be far more suitable and robust than a deep learning network.

The recognition of barcodes, QR codes and data matrix codes also generally requires no resource-devouring deep learning network. Sophisticated methods without deep learning have proven reliable in such cases, offering high performance and accuracy with relatively low requirements for memory and computation.

Other use cases that are typically handled with conventional methods include precise size and distance measurements. Conventional methods have proven reliable here as well, so that an implementation with deep learning would be laborious without offering a significant advantage.

In general, complex machine vision applications include components that use conventional methods as well as those using deep learning-based methods. In many cases, different test steps are performed simultaneously in an inspected image, for example to conventionally read a barcode and at the same time perform object recognition and a structural inspection of the object with the aid of a deep learning network.

6. Outlook

Many current research projects and new developments using deep learning network architectures focus on optimization strategies to achieve better performance. The goal is often to reach real-time capability, for example to use in systems for autonomous driving. The detection of the relevant image detail (ROI) and corresponding image data reduction before the object identification are usually integrated into the deep learning architecture through various model approaches. Prominent examples include “Faster RCNN” or “Mask RCNN”, which build on a two-stage concept, or so-called “single shot detectors”, such as “YOLO” and “SSD”. The results of the various architectures differ in their speed and the accuracy of their recognition.

New insights about the functioning of neural networks deal with the initialization values of the network structures before training, which have been randomly assigned up to now. The research results suggest that with appropriate initialization values, the networks actually need only a tenth of the actual network sizes, with equal accuracy. It can thus be expected that efficient and complex deep learning networks that require much less memory and computational effort can be implemented in the future.

Summary

Compared with a variety of conventional methods, deep learning offers many advantages while also increasing the possibilities of analytical image processing, by enabling applications that would not be possible with traditional methods with a reasonable effort.

In order to benefit most efficiently from the advantages of deep learning-based methods, it can be expedient to use a hybrid approach that combines conventional methods and deep learning.

Such a combination makes it possible to reduce the memory requirements, power consumption and training effort of a deep learning network while improving the execution speed at the same time.

In some use cases, it actually makes more sense to apply established conventional methods without deep learning, for example with code readers or a simple classification by color. Here the use of deep learning methods would be too laborious without offering significant advantages.

Even if deep learning networks are becoming increasingly efficient and productive in the future, there are still good reasons for the use of conventional methods. They are useful for providing support in the preprocessing for deep learning networks as well as in use cases in which the application of deep learning, even with high performance and efficiency, offers no advantage.

Author

Tilmann Zuper
Product Manager Software

Tilmann Zuper, Product Manager in Marketing at Basler AG, is in charge of Basler's software for the integration and use of cameras in Machine Vision applications.

During over 20 years of professional experience in leading companies in the area of software product development, he worked as a project manager, product manager and consultant, including for image-processing and networked systems. He graduated from the English Open University in Milton Keynes with a Master of Science in Software Development.

About Basler

Basler is an internationally leading manufacturer of high-quality cameras and accessories for applications in factory automation, medicine, traffic and a variety of other markets.

The company's product portfolio encompasses line scan and area scan cameras in compact housing dimensions, camera modules in board-level variants for embedded vision solutions, and 3D cameras. The catalog is rounded off by the user-friendly pylon SDK and a broad spectrum of accessories, including a number developed specially for Basler and optimally designed for the Basler cameras. Basler has 30 years of experience in the area of computer vision. The Basler Group is home to approximately 800 employees at its headquarters in Ahrensburg, Germany, and its additional sites in Europe, Asia and North America.

Contact

Tilmann Zuper – Product Manager Software
Tel. +49 4102 463 402
E-Mail: tilmann.zuper@baslerweb.com

Basler AG
An der Strusbek 60-62
22926 Ahrensburg
Germany