# SICK AG WHITEPAPER

## SENSORS AND AI FOR FACTORY AUTOMATION

**Dr. Ola Friman**
Head of Research & Development Machine Vision, Leader I4.0 Deep Learning Initiative
SICK AG

SICK
Sensor Intelligence.

## TABLE OF CONTENTS

SUMMARY

Machine Learning methods based on Deep Neural Network models have in recent years made breakthrough strides in automated perception of data, e.g., for interpretation of images, speech and text. In this white paper, opportunities of the Deep Neural Network technology for factory automation are discussed, with an emphasis on applications in or close to industrial sensors. Whereas research results hold a promise of the next leap in automation levels there are also challenges concerning data, engineering, processes and communication that are important for successful adoption. Addressing these challenges will be key to get past technology pilot phases into operational deployments.



8026681/2021_04
Subject to change without notice

## Introduction

Industrial factory automation relies on sensors such as cameras, LiDAR, light grids, RFID and encoders to provide the perception capabilities necessary for decision-making and control, e.g., for sorting, robot picking and quality inspection. Computational algorithms broadly referred to as Artificial Intelligence (AI), process the raw sensor data to extract relevant information and to form decisions. Traditional algorithms consist of rules and mathematical operations designed and parametrized by human experts. A simple example application is to discard a produced item if a hole dimension is not within a given tolerance threshold, where both the mathematical operations to extract the hole dimension from an image and to set the tolerance threshold value are design questions for human domain experts.

Machine Learning offers a different algorithmic approach to the above inspection problem in which the human handcrafting is replaced by an optimization of the parameters in a Machine Learning model that maps the raw sensor data as input to the desired output decision to reject the item or not. What is ultimately left for the human is to give examples of correct mappings, i.e., to supply training images of holes with the right and wrong dimensions respectively. One advantage of the Machine Learning approach is that the underlying mathematical optimization procedures can handle millions of model parameters, which is impossible to handle for a human. It can thereby also find solutions that are not obvious to a human. A consequence of this advantage is however that the Machine Learning solution often becomes a black box where the inside decision mechanisms cannot be understood, with consequences for life cycle management and general trust in the system.

In recent years, the use of so-called Deep Neural Network models have been shown to outperform human handcrafted algorithms within the machine vision and speech understanding domains. For factory automation, one application of Deep Neural Networks is to mimic the outstanding human visual perception.

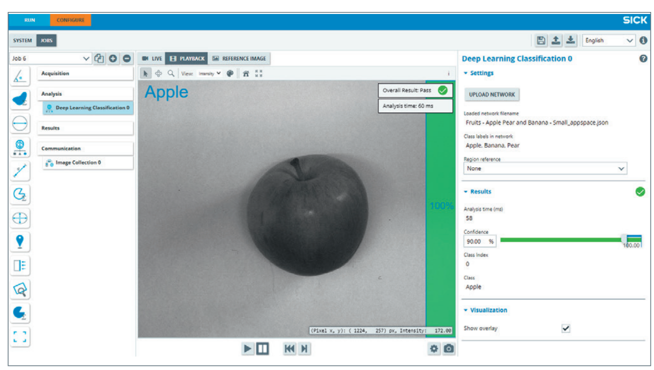


Fig 1: Deep learning technology is well-suited for applications where the task is to identify slight differences within one class, such as grown objects. This example shows the user interface from the Intelligent Inspection SensorApp from SICK.

This is achieved by optimizing the neural network to reproduce human responses to visual data for tasks such as visual defect inspection, localizing objects in the camera field-of-view, sorting based on visual appearance or spotting foreign items in food production. In parallel and in conjunction with the advancement in AI technology, related disciplines also experience strong development, including robotics, data connectivity, Internet of Things, miniaturization of computing power and cloud technology.

This paves the way for the next generation digital transformation manufacturing systems with a high degree of automated and optimized decision-making, leading to improved production flexibility, resource utilization, waste minimization and product quality. This paper discusses opportunities and challenges for the next decade 2020-2030 related to sensor development and AI in form of Deep Neural Networks towards digital transformation production systems. The following sections highlight opportunities and challenges in adopting the Deep Neural Network technology for factory automation.

## Opportunities

This section identifies different types of opportunities of the Deep Neural Networks with a special focus on sensor technology. An emphasis is on opportunities that can be expected to be realized in some form already in the short-to-midterm time frame based on already demonstrated algorithmic improvements. Long-term predictions in the AI domain are difficult given both the complexity of the problems and the current development pace.

**Opportunity 1: Sensor perception**

The most obvious way Deep Neural Networks may contribute to more efficient production processes is by automating tasks that have not been tractable by means of conventional algorithms. Until now, such tasks either have required the interpretation skills of a human or were simply not possible at all.

Fig. 2: Object classification is one example of a machine vision recognition task. In this user case, new and used bottles are classified in the infeed zone at the bottle-washer machine.

Deep Neural Networks excel in applications made difficult by large variations in the objects-of-interest appearance or by changing light or background conditions. Given the results of past and current intensity of research within the computer vision field, we are likely to see a strong increase of Deep Neural Networks for camera and machine vision tasks over the next couple of years. In the midterm development, we see the application of Deep Neural Networks to other sensor modalities. Compared to regular 2D cameras, LiDAR sensors bring the additional advantage of reliable 3D distance measurements, which are useful for 3D shape recognition as well as for autonomous ground vehicle navigation and obstacle detection. Processing of 3D point clouds from LiDAR sensors using Deep Neural Networks are also likely to be seen in factory and logistic automation systems in the short-to-midterm. Applications where camera and LiDAR sensors are used are characterized by the need to recognize spatial relationships such as shape and distances. In other applications, temporal relationships may carry the information, e.g., a vibration pattern. Deep Neural Network architectures for such applications can draw on results and research within the speech recognition field. There is a number of industrial sensors producing measurements that can be analyzed as a time series, including for example sensors for acceleration, motion, flow, temperature, pressure, distance and proximity. Where Deep Neural Networks provide additional information extraction capacity compared to traditional arithmetic for such sensors still needs to be explored. The key properties to look for are on the one hand an information richness and complexity that implies a robustness aspect in the interpretation of the data, on the other hand the possibility to get a sufficient amount of ground truth data, keeping in mind that human recognition may not be adapted to this kind of data the same way it is to visual and speech data.

**Opportunity 2: Measurement utilization**

While predictions around AI typically often revolve around solving new automation applications, an overlooked aspect is to utilize improved perception skills to simplify existing applications by a more efficient utilization of the measurement data. A straightforward example would be to replace high-resolution 2D cameras with lower resolution ones, but one can also foresee examples where a 2D camera plus improved AI can accomplish the same task as a larger 3D camera. Trends in this direction can be seen within the robotics domain where Deep Neural Networks trained on CAD models can estimate the six-dimensional pose, 3D location and 3D orientation, of an object from a 2D image of it. A practical consequence may be that one can make lighter and small-sized sensor solutions fitting more narrow spaces.

**Opportunity 3: A new configuration paradigm**

A key property of Machine Learning and Deep Neural Network approaches is that they are configured in a fundamentally different way compared to a traditional algorithmic approach, i.e., through a well-defined procedure from collecting the raw data, annotate the raw data, train and finally deploy a neural network.


Fig. 3: A typical workflow starting with collection of the raw data by the sensor using a SensorApp, the training takes place in the cloud service and the trained Deep Neural Network is finally uploaded to the sensor again to be used to directly classify images in the factory.

While there is a great deal of knowledge involved in the above steps, much of the complexity can be hidden in software tools that streamline and simplify the procedure to the extent that it becomes accessible for a wide group of people. That is, an algorithm expert will not be needed to solve highly complex machine vision automation tasks and also more simple tasks can be predicted to be addressed using Deep Neural Networks as the software support matures.

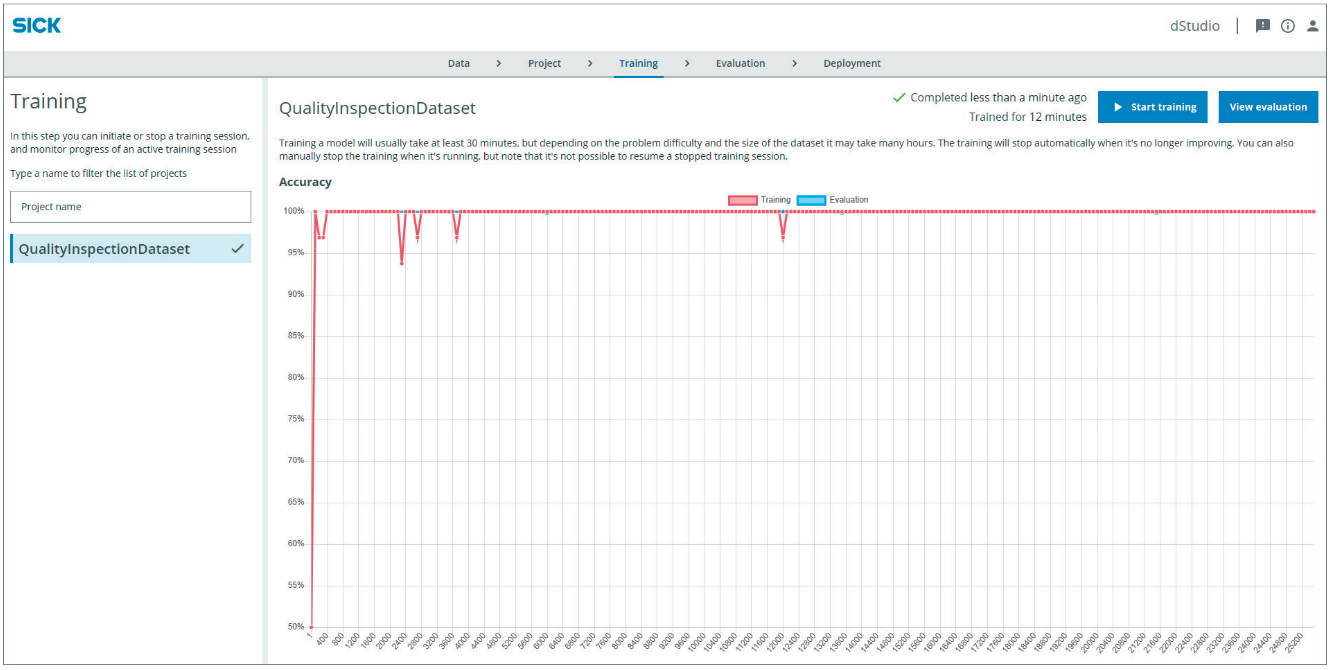


Fig 4: Much of the complexity can be hidden in the AI software tools. The example shows a trained neural network with data set from the cloud-based service dStudio from SICK.

**Opportunity 4: Post-deployment improvement**

Following Opportunity 3 above, there is also a clear path of how the performance of a deployed Deep Neural Network system may be improved by continuously increasing the training data set and retrain the network models. This can be especially useful in cases where there is an imbalance in the training data. For example, in a machine vision defect inspection system it is common to have an abundance of training data from the 'ok' class but fewer examples of defects as they occur less frequently. A deployment can in such case be made using an initial training data set and then subsequently be improved upon as more examples of the defect class are collected. It needs to be noted though that current Deep Neural Network systems do not adaptively self-tune during operation but that a retraining with new human-labeled data is needed. This limitation is further discussed in the Challenges section.

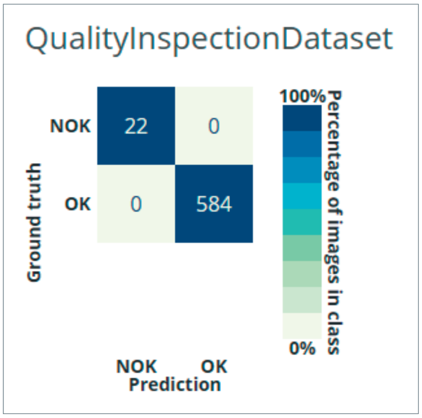


Fig 5: Evaluation matrix showing a typical machine vision inspection where there are more training data from the OK class but fewer examples of defects.

**Opportunity 5: Sensor data quality**

Deep Neural Network applications mentioned so far operate on output data from an industrial sensor. Inside most sensors there is a low-level measurement core where processing is carried out to produce the output measurements. 3D reconstruction in time-of-flight and stereo cameras are high-level examples of such processing. As low power edge computing becomes more powerful, Deep Neural Networks and other Machine Learning models can be applied deeper into the sensors with the aim of producing higher-quality measurements. Further use could be sensor self-monitoring to detect operation anomalies such as dirt or moisture on a camera lens.

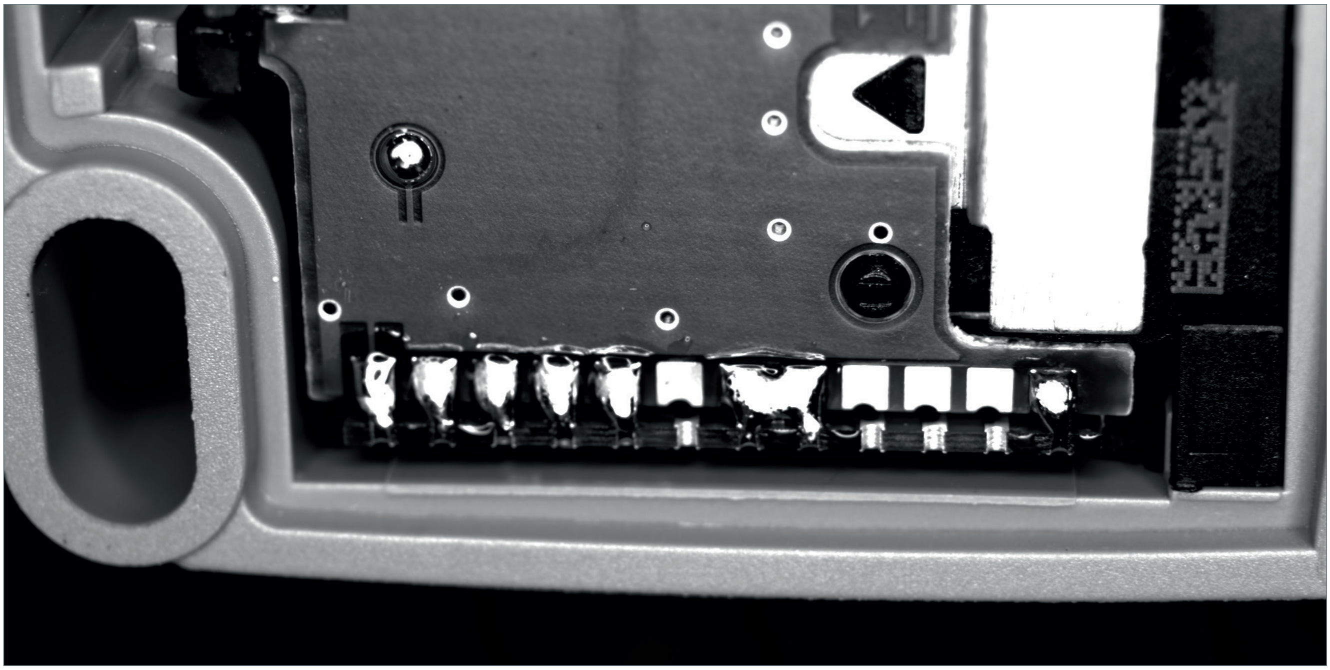

Fig. 6: Classification of good and bad soldering joints using a Deep Neural Network. Patterns like these can be difficult to describe and discriminate between using conventional algorithms.

## Challenges

Reaping the benefits of Deep Neural Network AI does not come without challenges. This section identifies issues to address in order to move beyond pilot studies and to deploy Deep Neural Networks for factory automation at large scale.

**Challenge 1: Managing expectations**

The scientific Artificial Intelligence field has in its history experienced several so-called 'AI winters' during which funding and interest went down significantly due to overinflated promises and expectations that could not be realized. While the latest AI results are directly applicable in industrial manufacturing systems, a challenge for the coming decade is to maintain the readiness to invest by communicating balanced predictions and descriptions that set realistic expectations. As an example, it is close at hand to make analogies with human learning when describing Deep Neural Networks, which if overemphasized may lead to false assumptions of what they, at least currently, are capable of doing. While humans can learn from a single sample instance, say a single image, a Deep Neural Network will need a multitude of examples also for quite trivial tasks. On a similar note, Deep Neural Networks will likely not respond to unfamiliar or unexpected situations in a similar way a human would. In fact, the performance of Deep Neural Networks in situations which were not covered in the training data will most likely be perceived as unsatisfactory. A description such as 'self-learning system' is ambiguous and should be used with care as current deployed Deep Neural Networks are not self-learning in the sense that they adaptively update themselves automatically as new data is observed; they need annotated training data, typically human-produced, to improve.

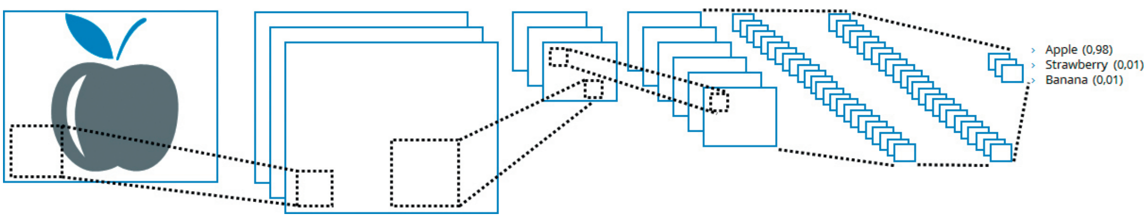


Fig. 7: Deep Learning finds pattern in data using neural networks with multiple layers. A Deep Learning Neural Network needs a multitude of examples also for quite trivial tasks like recognizing an apple.

**Challenge 2: Energy efficiency**

Deep Neural Network models are computationally intensive and therefore consume significant amounts of energy to train and use. To scale up AI in a sustainable way and to be able to deploy Deep Neural Networks within small embedded hardware, attention to the energy efficiency is required. There are two main ways to address this, both of which are being pursued currently: To shrink the Deep Neural Network sizes, both in terms of the number of parameters and in terms of numerical precision, and to design dedicated energy efficient hardware, ASIC, FPGA and similar, for Deep Neural Networks. We will likely see both approaches realized in deployed industrial systems in the short term.

**Challenge 3: Labeled data**

Deep Neural Networks systems are configured mainly by training the data sets. Data management and especially human labeling of the data are therefore bottlenecks in the adoption of AI in general and Deep Neural Networks in particular. Research efforts into reducing the amount of data needed for training Deep Neural Network models are ongoing, an example being to use a network pre-trained for some other task and tune it for the new task. Reducing the model network sizes as discussed in Challenge 2 works towards this end too. It should be noted though that even if the amount of labeled data needed for network training can be reduced, there is also the need for labeled test data to verify the accuracy of the trained neural network. As example, say one seemingly manages to train a Deep Neural Network for a machine vision inspection task using only fifty labeled example images, the network still needs to be verified using many more test images before deciding to deploy it into real operation. Another challenge is to acquire sufficient of training samples of unusual situations. A case already discussed is the typical imbalance in machine vision quality inspection where samples of defects are rarer than the normal class. Simulating training data can complement real data, reducing the manual labeling effort and also providing a method to

Fig 8: InspectorP61x ultracompact 2D vision sensor fits into tiny mounting spaces. Deep learning is available as licensed option.

generate examples of unusual situations. Creating realistic simulated data requires technical expertise though and the variations in factory automation setups and applications are large. Simulating data for each new case therefore comes with a significant cost and it still cannot fully replace real data for the pre-deployment testing as discussed above.

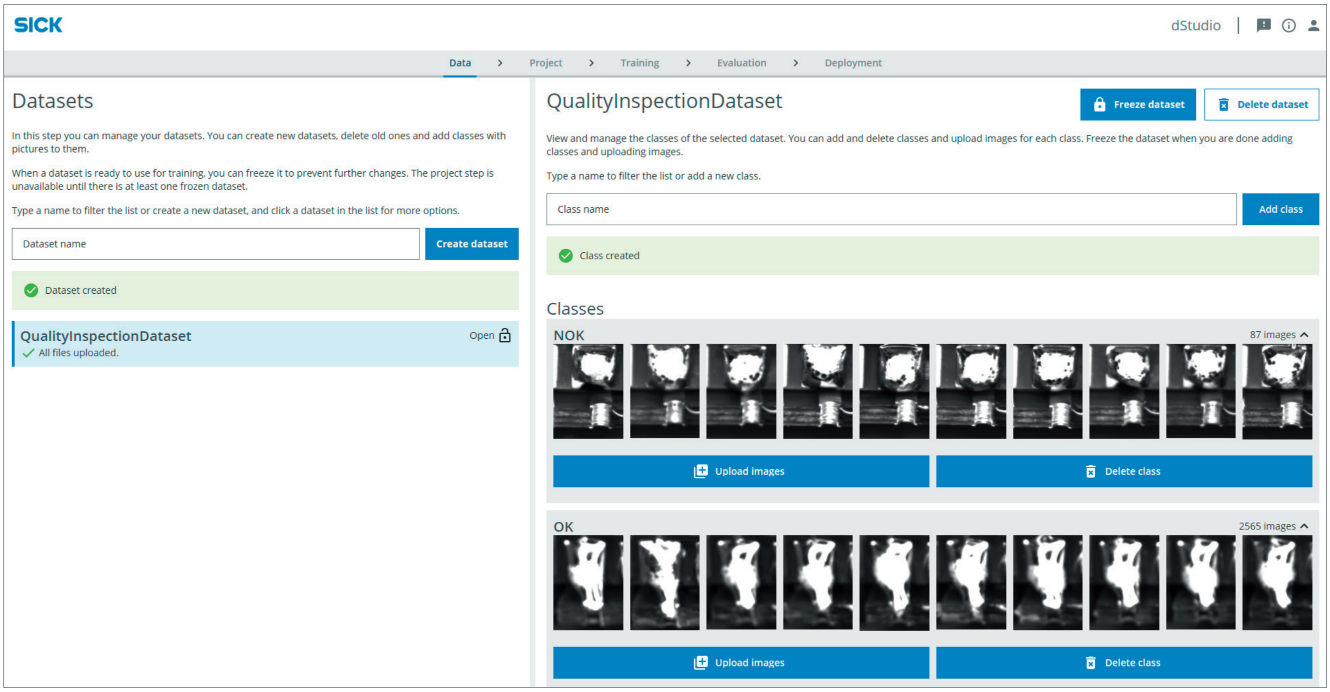


Fig. 9: Deep Neural Network systems are configured mainly by training data sets. This example shows a labeled data set in the cloud-based training service dStudio from SICK.

Finally, most current Deep Neural Network systems attempt to mimic human perception skills by learning from human-labeled example data. When moving into domains not covered by basic human recognition, e.g., radar, RFID, vibrations and other multi-dimensional measurement signals, specialists will be required to provide properly labeled data. A similar situation can be found in the medical imaging domain where radiology specialists, trained for years to recognize subtle variations within medical images, are needed to create the training data sets. Getting the time and willingness from domain specialists to hand-label large amounts of data is a major challenge.

**Challenge 4: Fusion with other models**
In industrial automation systems there are already plenty of geometric and mathematical models in use that describe prior knowledge of sensors and the production process. Examples include CAD models, camera calibration models or simply the laws of physics. It is currently an open research question how a Deep Neural Network, itself being a generic mathematical model, can incorporate the information represented within other more specific models. For example, how can a Deep Neural Network be used to find deviations in a produced item relative to a reference CAD model? Currently a Deep Neural Network must re-learn the information encoded in other models through observations, which is inefficient and prone to errors.

Fig. 10: Coping with the product variability in the batch size one production line is a challenge for AI and Deep Neural Networks. Example of industries; textile, car manufacturer and electronics.

**Challenge 5: Batch size one**
The ability to manufacture individually designed products with the same efficiency and quality as in mass-production lines is a central vision in the digital transformation. Using Deep Neural Networks in such a setting raises the question of coping with the product variability: how can quality assurance be automated when every manufactured object is different? Obtaining the necessary training data that lets the Deep Neural Network experience the full product variability will be more challenging

than in mass-production lines. Deploying Deep Neural Networks in batch size one systems will likely require solutions to both Challenge 3 to reduce the requirements on the amount of labeled data, and to Challenge 4 to be able to incorporate CAD models or similar information that may be encoded in digital twin models.

**Challenge 6: Life cycle maintenance**

Industrial systems have a long life span compared to consumer electronics and applications, more than ten years of operation is not unusual. Considering the fast-moving pace AI is currently developing at including the architectures and software used for training and deploying Deep Neural Networks, a question that needs attention is how AI solutions deployed today can be maintained over a ten-year period or more? Questions that arise include: Can the same Deep Neural Network format be deployed on new hardware after, say, five years of operation or have formats and software evolved to the degree that a retraining of the network is necessary? If a retraining of the network is necessary, is the original training data available and in a format that is still usable? Is there a sufficiently detailed description of how the original network was trained? What does it take to retrain a Deep Neural Network and obtain the same result each time, given that randomness is an essential part of Deep Neural Network training, e.g., to initialize the network weights and to randomly perturb and augment the available training data for increased reliability. A similar but more general question is how to handle an increasing number of black box solutions over long time spans. Compared to handcrafted algorithmic implemented as human-readable logic, Deep Neural Network solutions may prove to have a different set of life cycle challenges we have not yet experienced the implications of.

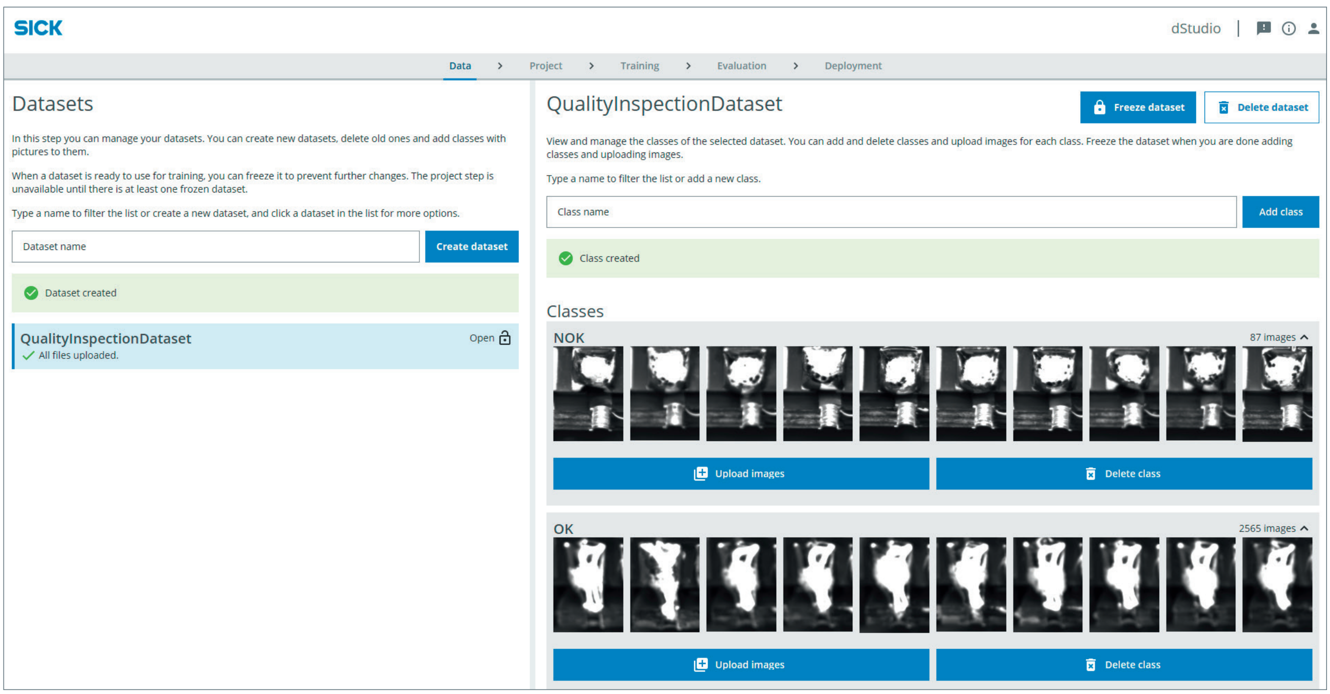


Fig. 11: Data management and especially human labeling of the data are bottlenecks in the adoption of AI in general and Deep Neural Networks in particular.

**Challenge 7: Safety-critical systems**

Similar to self-driving cars, there are industrial applications that include aspects of human safety, for example anti-collision systems for mobile or collaborative robots. Given the black box nature of Deep Neural Networks, how are they and tested certified, both from hardware and software points of view, for safety-critical applications? One aspect of this issue is to get standardized qualification processes in place. Another more fundamental issue is to make the Deep Neural Networks actually pass the tests applied by making them fault tolerant to unusual events and disturbances in inputs or in hardware. A special kind of vulnerability is so-called adversarial attacks, which are intentional modifications of the input data. One example is to put specially designed patterns in the view of a camera that are known to confuse an already trained Deep Neural Network to make erroneous predictions. Another example is to contaminate the training data set with the purpose of making a Deep Neural Network trained on it misclassify or ignore certain patterns in the data. Conclusion, keeping the training data secure and reviewed is therefore important.


Fig. 12: Self-driving cars utilize Deep Learning to find objects in the environment. Technology has been available for decades. With availability of data, computational power and open source algorithms adoption has increased.

## Summary and Outlook

Over the next years, we will see current Deep Neural Network research results transferred into operative factory automation deployments. The main driver is the improved sensor data perception capabilities through which more tasks and decisions can be automated. To maintain the current development momentum and to realize more of the opportunities brought up in this paper, attention must be paid to the challenges that do exist with the Deep Neural Network technology. The two fundamental challenges towards large-scale adoption are the dependence on labeled training data sets and the black box character of the solution. As example, both these challenges emerge in the quality assurance process: Due to the black box character one can only verify functionality for the particular labeled cases used for testing, there is no way to look into the Deep Neural Network and verify high-level rules the network has inferred. Another example is the resilience to changes in the environment or sensor data acquisition. For a visual inspection task, a human will adjust to variations in lighting, camera angle or exposure time within certain limits, whereas a Deep Neural Network may not, unless these variations are explicitly included in the training data set. Obtaining labeled data that is not too narrow in scope is for this reason a task that must not be underestimated and which is key for a successful transition from pilot investigation to product.

The relatively small labeled data set sizes currently available in the factory automation domain is a challenge both for applications and for applied research in the area. A promising near-time research direction is to utilize unlabeled data in the Deep Neural Network training to augment the smaller amount of labeled data. The unlabeled data does not give information about the task to learn per se, but it does prime the network to the general structure and variations in the data. A related but longer-term and more complex question is how to move towards truly self-learning systems? How can current Deep Learning Network architectures with good perception skills evolve to obtain cognitive skills that allow them to reason and generate new knowledge on their own without, or very little, labeled data. In industrial contexts, one can predict to see such cognitive elements first in the robotics domain using so-called reinforcement learning methods. A notable difference between a robotics application and other applications mentioned in this paper, e.g., a visual defect inspection task using a camera, is that the robot is an actuator that has the possibility to perform actions and observe the result through sensors. A question is to what extent the ability to interact with the environment and sensing the effects is a prerequisite to develop cognitive skills, and how this in such case affects the long-term AI development for factory automation? As a final note, it follows from the predictions in this paper that new types of jobs will be created within factory automation, with the tasks to create labeled data sets, maintain, test, improve and retrain Deep Neural Networks or other Machine Learning models to new situations or products. Industrial automation is characterized by a myriad of specialized tasks; consider for example all the different products being manufactured that can be quality-inspected at various steps of the production, each requiring its own-labeled data set. Compared to other AI success areas such as car driver assistance systems, face recognition and voice recognition, it is not as straightforward to reuse and cooperate to pool labeled data. Bringing out Deep Neural Networks in factory automation will therefore be a more decentralized effort that takes place closer to the individual applications.

8026681/2021-04-14 · BEM · Whitepaper

**FURTHER LINKS**

➜ www.sick.com/ai
➜ www.sick.com/Deep_Learning
➜ www.sick.com/Intelligent_Inspection

SICK AG | Waldkirch | Germany | www.sick.com

Sensor Intelligence.