

AI-based quality control on every PC for every employee: BMW Group is banking on Intel® OpenVINO™



Cracks, scratched paint or incorrect labelling: The BMW Group relies on systems for automated image processing (machine vision) to detect defects in production and quality control. These solutions use AI-based deep learning to efficiently process and analyze visual data, mostly based on specialized hardware or in the cloud. To not only enable data scientists but employees in manufacturing as well as to use AI workloads for quality control on a classic desktop PC, the BMW Group uses an application developed by Robotron with Intel® OpenVINO™. Intel's toolkit accelerates and improves machine vision with deep learning, bringing the BMW Group closer to its goal of an "AI for everyone".

With its BMW, MINI, Rolls-Royce and BMW Motorrad brands, the BMW Group is one of the world's leading premium manufacturers of automobiles and motorcycles as well as a provider of premium financial and mobility services. The BMW Group production network comprises of 31 production and assembly facilities in 15 countries while the global sales network has representations in more than 140 countries. In 2020, the BMW Group achieved worldwide sales of more than 2.3 million automobiles and more than 169,000 motorcycles. The Group currently employs around 120,000 people worldwide.

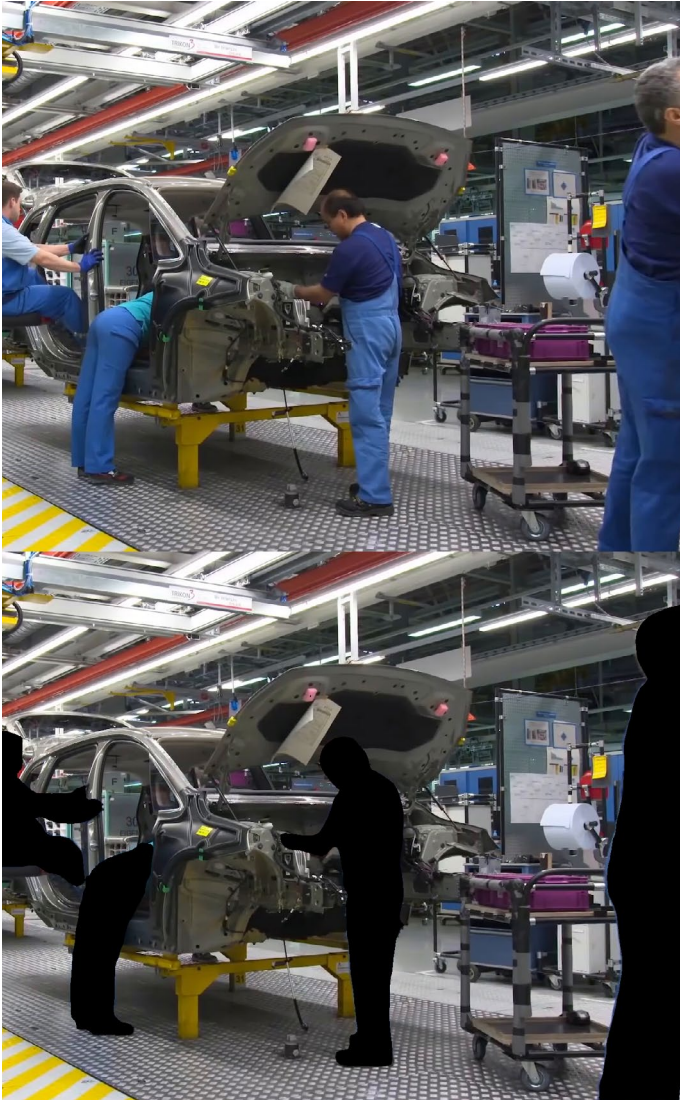
Automated image processing in production

Optical quality control of components using image processing based on Artificial Intelligence (AI) has become standard in the industry. Since 2018, the BMW Group has been using various AI applications for production purposes. One focus is on automated image recognition methods. Experience shows that systems with machine vision are superior to human examiners in repetitive testing tasks. Machine vision works faster, more accurately, and more reliably while it never grows tired. It can check hundreds of parts per minute on the production line around the clock, with consistently reliable results. In addition, with proper resolution and optics, such tools are able to recognize details that the human eye does not see.

At the BMW Group, AI solutions relieve production staff during monotonous routine activities such as checking whether the warning triangle or first aid kit is in the right place in the trunk or if the windscreen wiper cover cap has been fitted correctly. At the Dingolfing plant, an AI application also performs more demanding tasks. In the final inspection phase, the application compares the order data of a vehicle with the live image of the model lettering of the newly produced automobile. These letterings such as "X3" or "xDrive" as well as all possible combinations are stored in an image database. In the absence of a lettering, the final inspection staff receive a notice to that effect.

Another example is an AI application from the pressing plant in which components for the cars body are formed from boards. There, the BMW Group, together with Robotron Datenbank-Software GmbH, a specialist for Industrial Computer Vision and IoT, based in Dresden, developed a solution that eliminates so-called pseudo-errors, i.e., deviations from the set value, even though in reality no deviation exists. "Metal chips or oil residues that remain on the body parts after forming can easily be confused with very

fine cracks,” explains Michael Baling, Head of the Industry Division at Robotron. “With the AI solution and the possibility of the neural network to access about 100 real images per characteristic pseudo-errors are minimized – with 100 images of the flawless component as well as 100 images with dust grains, oil drops on the component or with cracks. This can be used to identify borderline cases.”



© BMW Group

Based on deep learning

The described inspection solutions for quality and conformity testing in manufacturing use deep learning to process the huge amount of image data. Deep learning is a subcategory of AI or machine learning based on artificial neural networks that have multiple layers and are similar to the human brain. After training, special algorithms autonomously extract samples from raw data (in this case image data) and improve their skills over time. As a rule, humans no longer have to be a part of the actual learning process.

Image processing is based on a pattern comparison. In order to obtain reliable and repeatable results, the AI system must quickly and accurately match trained patterns with the actual objects on the assembly line. In the current production, the solution evaluates images of a component and aligns them in milliseconds via the algorithm, which draws upon the trained

knowledge from hundreds of other images of the same component. This way, the artificial neural network determines deviations from the standard in real time and checks whether, for example, all necessary parts are installed or mounted in the right place.

Deep Learning Inferencing, i.e., the concrete use of the trained AI model for the evaluation of the visual data, is so far only possible to a limited extent on special hardware.

Goal: Democratize AI with Intel OpenVINO

The target of the BMW Group is to democratize AI inferencing and a “no code” solution. This means that in the future, not only data scientists or machine learning specialists should be able to develop and use AI workloads for quality control. With Intel OpenVINO, every employee in manufacturing should have this tool at his or her disposal, not only running on expensive special hardware but on every standard PC.

In order to achieve this goal, the BMW Group, together with Robotron and Intel, has created a general AI-platform or rather an inference API. Using it, every user will be able to build and use an AI application in the future for object recognition on his work PC. This open-source project, which the BMW Group provides for the entire industry under <https://github.com/BMW-InnovationLab/BMW-IntelOpenVINO-Inference-API> is based on the Intel OpenVINO Toolkit. The solution runs on both Windows as well as Linux operating systems.

Intel OpenVINO is a special toolkit for deep learning frameworks such as TensorFlow, MXNet and Caffe. It includes the Intel® Deep Learning Deployment Toolkit with a Model Optimizer and an Inference Engine, as well as optimized libraries and features for OpenCV and OpenVX. OpenCV is a free software library with algorithms for image processing and machine vision; OpenVX is an open standard for cross-platform acceleration of machine vision applications.

The OpenVINO Toolkit accelerates workloads for machine vision with deep learning and optimizes these applications across different Intel platforms. These include CPUs with built-in graphics, hardware accelerators for cameras or other image processing devices in edge computers where the AI application works close to the video camera. This ensures that the algorithm runs optimally on each PC.

Current project: Anonymization of image data

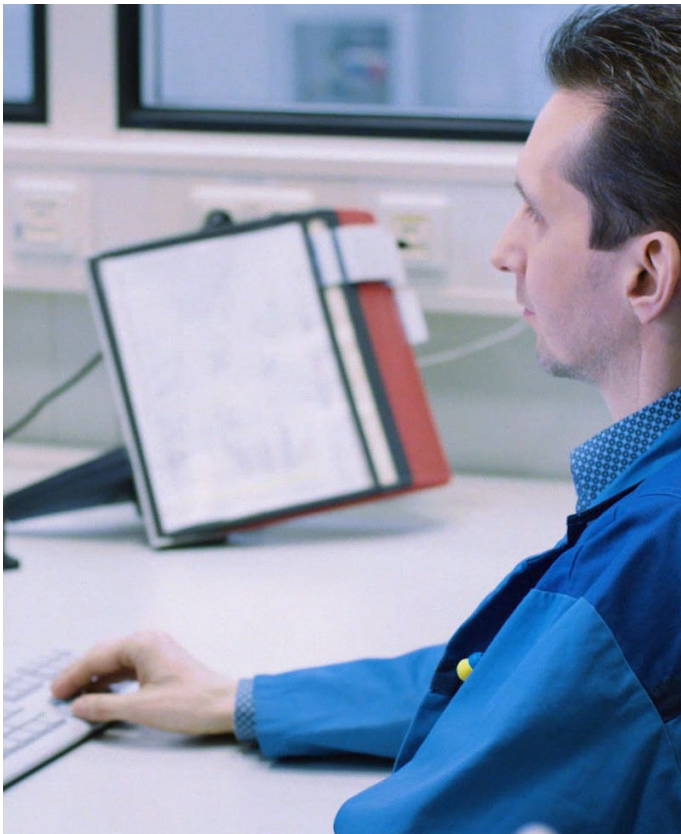
Currently Robotron is working on a project with Intel OpenVINO for the BMW Group to anonymize faces or components.

“Sometimes the cameras used for quality monitoring also make production processes or the faces of employees visible. The work council and the General Data Protection Regulation require anonymizing in order to maintain compliance. Thus, AI helps machine vision to become GDPR compliant,” explains Michael Baling of Robotron.

In order to automatically anonymize people or components, the faces visible on the image/video should be pixelated or concealed with the new AI module. The same applies to areas or products that should not be caught and recorded on the image/video. For this purpose, models are trained with image files to detect whether there are any people or components in the image so that they can filter out or pixelate them later on. The anonymization process, in which the model or algorithm

is applied to the image or video, runs on a standard Office PC. Thanks to OpenVINO, no servers or other special hardware are necessary anymore.

Michael Baling sums up: "With OpenVINO, the inferencing for the anonymization workflow can be carried out directly on standard hardware and optimized perfectly for this purpose. Thanks to our purpose-oriented partnership with Intel, we were very quickly able to achieve good results for the BMW Group. This means that now potentially every employee is able to use AI applications for quality control."



© BMW Group

Components of the Intel OpenVINO Toolkit

The **OpenVINO Toolkit** includes a Model Optimizer, an Inference Engine, optimized libraries and features for OpenCV and OpenVX, and the web-based Deep Learning Workbench graphical environment. Here is a brief description:

- The **Model Optimizer** is a python-based command line tool running on Windows, Linux and MacOS operating systems. It supports more than 100 already trained, publicly accessible models from popular deep learning frameworks such as Caffe, TensorFlow, Apache MXNet or Open Neural Network Exchange (ONNX). The Model Optimizer performs analyses and adapts the models for optimal execution on different end devices. This way, algorithms can be deployed to various Intel hardware platforms, saving time and development resources.
- The **Inference Engine** provides inference solutions via a common API on the desired platform, i.e., processor, graphics processor (GPU) etc. Inference refers to the use of the trained model to draw conclusions or make predictions from given data. This also makes it possible to run different layers on different hardware, such as CPU and GPU, as well as optimize workloads for, among other tasks, graph analysis, scheduling, or model compression.
- **Open CV and OpenVX:** The toolkit also offers optimized libraries and functions for OpenCV and OpenVX. OpenCV is a free software library with algorithms for image processing and machine vision for the programming languages C, C++, Python and Java. OpenVX is an open standard for cross-platform acceleration of machine vision applications.
- **Deep Learning Workbench:** This web-based graphical environment allows users to visualize the performance of deep learning models and datasets on different Intel architecture configurations (CPU, GPU, VPU). It provides important metrics such as latency, throughput, and performance for each layer of the selected neural network. This allows inference experiments to be configured to deliver optimal performance.

**BMW
GROUP** *robotron®*

Solution Provided By:

