



Pick & Place with artificial intelligence without a PC

AI is fast, robust, works almost error-free and does not need to pause. This means that it is already superior to humans in many areas where work processes have to be carried out continuously with consistently high performance and quality. Reasons why one would like to use AI in the machine vision environment in interaction with robotics to make processes more efficient and cost-effective. The "Vision Guided Robot" use case shows how typical pick & place tasks can be intelligently automated with a robot and an embedded AI vision camera, and how a PC can even be made redundant.

For "smart gripping", different disciplines must work together optimally. If, for example, the task is to sort products of different size, shape, material or quality using robots, they must not only be gripped, but also identified, analysed and localised beforehand. With rule-based image processing systems, this is often not only very time-consuming, especially in small batch sizes, but also hardly economically feasible. But in combination with AI-based inference, industrial robots can already be equipped with the necessary skills and product knowledge of a skilled worker. In fact, the wheel no longer needs to be reinvented for the individual subtasks. It is enough to have the right products work together effectively in an interdisciplinary manner as a "smart robot vision system".

EyeBot Use Case

In a production line, objects are randomly scattered on a conveyor belt. The objects must be detected, selected and, for example, placed in a packaging or passed on in the correct position for a further processing or analysis station. The software company urobots GmbH has developed a PC-based solution for detecting objects and controlling robots. Their trained AI model was able to recognise the position and orientation of objects in camera images, from which grip coordinates for a robot were determined. The goal was now to migrate this solution to the AI-based embedded vision system from IDS Imaging Development Systems GmbH. For urobots, two things were most important for the solution:

1. The user should be able to easily adapt the system himself for different use cases without special AI expertise. This means even if, for example, something changes in production, such as the lighting, the appearance of the objects or even if additional object types are to be integrated.
2. The overall system was to be completely PC-less through direct communication of the device components in order to be both cost-effective and light and space-saving.

Both requirements are already available by IDS with the IDS NXT ocean inference camera system.



Alexey Pavlov (Managing director of urobots GmbH) explains:

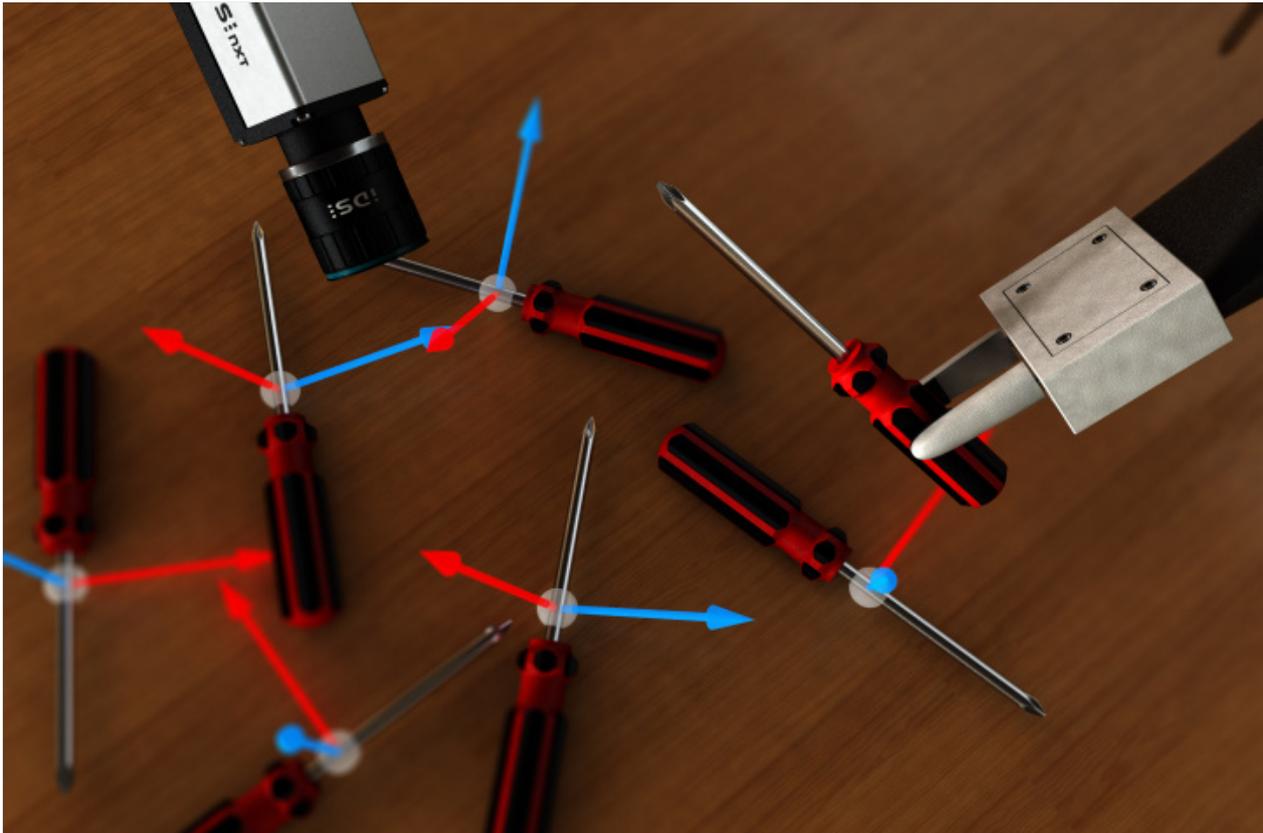
"All image processing runs on the camera, which communicates directly with the robot via Ethernet. This is made possible by a vision app developed with the IDS NXT Vision App Creator, which uses the IDS NXT AI core. The Vision App enables the camera to locate and identify pre-trained (2D) objects in the image information. For example, tools that lie on a plane can be gripped in the correct position and placed in a designated place. The PC-less system saves costs, space and energy, allowing for easy and cost-effective picking solutions."

Position detection and direct machine communication

A trained neural network identifies all objects in the image and also detects their position and orientation. The AI does this not only for fixed objects that always look the same, but also when there is a lot of natural variance, such as with food, plants or other flexible objects. This results in a very stable position and orientation recognition of the objects. urobots GmbH trained the network for the customer with its own software and knowledge and then uploaded it to

the IDS NXT camera. To do this, it had to be translated into a special optimised format that resembles a kind of "linked list". Porting the trained neural network for use in the inference camera was very easy with the IDS NXT ferry tool provided by IDS. In the process, each layer of the CNN network becomes a node descriptor that precisely describes each layer. The end result is a complete concatenated list of the CNN in binary representation. The CNN accelerator IDS NXT ocean core, which was specially developed for the camera and is based on an FPGA, can then optimally execute this universal CNN.

The vision app developed by urobots then calculates optimal grip positions for a robot from the detection data. But this did not solve the task. In addition to the results of what, where and how to grip, direct communication had to be established between the IDS NXT camera and the robot. This task in particular should not be underestimated. This is often the crucial point that determines how much time, money and manpower has to be invested in a solution. urobots implemented an XMLRPC-based network protocol in the camera's vision app with the IDS NXT Vision App Creator in order to pass on the concrete work instructions directly to the robot. The final AI vision app detects objects in about 200 ms and achieves a positional accuracy of +/- 2 degrees



The neural network in the IDS NXT camera localises and detects the exact position of the objects. Based on this image information, the robot can independently grasp and deposit them.

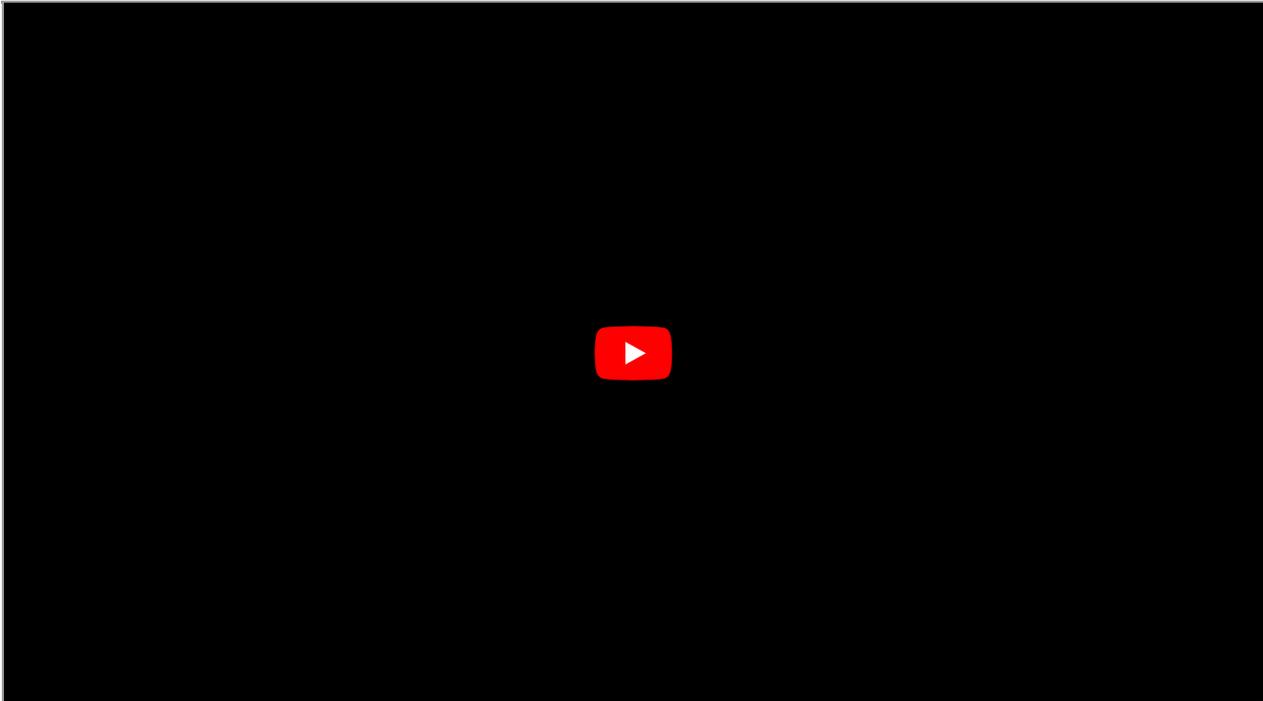
More than artificially intelligent - PC-less

It is not only the artificial intelligence that makes this use case so smart. The fact that this solution works completely without an additional PC is also interesting in two respects. Firstly, since the camera itself generates image processing results and does not just deliver images, the PC hardware and all the associated infrastructure can be dispensed with. Ultimately, this reduces the acquisition and maintenance costs of the system. Quite often, however, it is also important that process decisions are made directly at the production site, i.e. "in time". Subsequent processes can thus be executed faster and without latency, which in some cases also enables an increase in the clock rate.

Another aspect concerns the development costs. AI vision or the training of a neural network works in a completely different way than classical, rule-based image processing, and this also changes the approach and handling of image processing tasks. The quality of the results is no longer the product of manually developed programme code by image processing experts and application developers. In other words, if an application can be solved AI-based, IDS NXT ocean can also save costs and time of the corresponding experts, because with the comprehensive and user-friendly software environment, each user group can train a neural network, design the corresponding vision app and execute it on the camera.

The EyeBot use case has shown how a computer vision can become a PC-less embedded AI vision application. The expandability through the vision app-based concept, the application development for different target groups and an end-to-end manufacturer support are also advantages of the small embedded system. With EyeBot, the competences are clearly distributed in an application. The user's attention can stay with his product, while IDS and urobots concentrate on training and running the AI for image processing and controlling the robot. An additional advantage - through Ethernet-based communication and the open IDS NXT platform, the vision app can also be easily adapted for other objects, other robot models and thus for many other similar applications.

Video showing the EyeBot from urobots



Further information

- Learn more about the IDS NXT embedded vision AI platform on the [product website](#).
- Discover the easy entry into Deep Learning technology with the IDS NXT ocean inference camera solution in the technical article "[AI for everyone](#)".
- For more information on the advantages of the deep ocean core as an FPGA solution on a camera platform, see the technical paper "[Sustainable Edge Intelligence](#)".