Using deep learning to push recognition rates to a maximum and exception jobs to a minimum

A NEW GENERATION OF OCR ENGINES BASED ON CONVOLUTIONAL NEURAL NETWORKS AND ARTIFICIAL INTELLIGENCE.

Game changing technology
Camco is renowned for its highly performing OCR engines and has proven solutions serving many terminals worldwide. In our constant striving to further increase the accuracy of our systems, we are investing significantly in the continuous improvement of recognition rates.

As a market leader keen on innovation, Camco decided 3 years ago to extend its Artificial Intelligence (AI) know-how with Deep Learning techniques to further enhance its systems. Our software team has successfully implemented deep learning models into our OCR camera software, providing excellent results.

In traditional OCR systems, programmers use hand-coded rules to define specific patterns or features. The OCR engines scan images to search and recognize these patterns or features. Each pattern, font, font size, etc. needs to be programmed in the software to achieve good recognition rates and high confidence levels. However, in some cases the traditional OCR engines generate false positives: a result that the OCR engine claims to be correct, but is actually not.

For the development and deployment of artificial intelligence and deep learning solutions, Camco relies on a team of 6 computer vision experts, who have integrated this technology into the latest OCR engines.

Developing systems with deep learning functionality is a highly complex matter needing extensive knowledge and expertise and requires a strong and solid architecture of software and hardware components. Camco’s 19 years of experience allows to get the maximum out of the combination of traditional OCR engines and Artificial Intelligence.

Deep learning and images
Deep learning is one of many approaches to machine learning and includes several technologies. When working mostly with images, convolutional neural networks (CNN) is one of those groundbreaking technologies.

Explaining convolutional neural networks is a challenge. This new generation of image recognition algorithms was inspired by studies on the visual cortex of animals. These networks mimic how the brain extracts abstract information from the signals generated by the optical nerves in the eye. This information could be shapes like squares or triangles etc, but also the silhouette of a person, a car or any object.

Just like the human brain, convolutional neural networks learn by experience. Therefore, developers no longer need to handcraft patterns or features to look for in the image. But provided

FIGURE 1: Example of container classification, illustrating the end-to-end learning principle. An image goes in, a result comes out. In between, the network has self-learned the meaningful features of a certain class, based on the thousands of images that were fed. In the case of a container door, these are the company logo, bars and hinges, container number and other textual information. The center heatmap illustrates that our network has learned to extract these features very well.
with a large number of labeled images - the training set - the network self-learns the features that comprise an object. What these features are exactly, depends on the dataset and the architecture of the network. A car detector for example, might learn the individual parts that make up a car, such as tires, radiator and learn to recognize the shapes and textures of each of these parts and combine them.

The process of supervised learning requires that Camco feeds the networks with lots of high quality images and annotations (objects contained in an image and their location within the image). The process of collecting a high quality dataset and optimizing both the data, the models and the training procedures are complex processes, requiring expertise and resources. But the results are quite rewarding.

It has taken Camco a lot of investment and effort to reach the benchmarks and targets we set out when we started developing this technology. But through combination of high-tech engineered sensor solutions, running on hardware optimized for deep learning (NVIDIA TX2) and software developed, reliable results are obtained.

How do terminals benefit?

Our extensive and augmented datasets - containing images from all over the world - are the basis of our generic convolutional models that produce better recognition rates, lowering the number of exception jobs.

Following examples illustrate what is difficult to define using traditional OCR software, while accurately read using our latest models: partly occluded characters (1), dirty and damaged characters; mixed font sizes and font colors (2); puzzle characters (3), shadows; damage (4).

<table>
<thead>
<tr>
<th>License Plate OCR rates</th>
<th>Before</th>
<th>After</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site 1 (USA)</td>
<td>84.8%</td>
<td>99.4%</td>
</tr>
<tr>
<td>Site 2 (South America)</td>
<td>84%</td>
<td>98%</td>
</tr>
<tr>
<td>Site 3 (Europe)</td>
<td>76.9%</td>
<td>91.6%</td>
</tr>
<tr>
<td>Site 4 (Europe - Belgium)</td>
<td>93%</td>
<td>99.8%</td>
</tr>
</tbody>
</table>

TABLE 1: This table illustrates how various customer sites experiencing issues with license plate recognition, significantly improved after Camco’s License Plate OCR software was upgraded to the newest one, using deep learning models.

The results show that there is an enormous step forward in the rates for license plate OCR accuracy. In one project, deep learning moved accuracy rates for license plate recognition from 76.9% to over 91%. The same generic models are used for all sites.

Deep learning in other applications: damages, crane spreader flight path

Current results are very good, and we see more areas where AI will contribute to the efficiency and effectiveness of terminals. One key area is the enhancement of damaged characters and data. AI helps in “repairing” damaged data seen in pictures. Like the human brain, AI helps the OCR engines to add missing pieces of characters, allowing to extract information.

Recognizing damages on containers is another area where deep learning may prove very useful. Although this is very complex matter, Camco is investing in developing technology to automatically detect damages. However, this will require significant time and resources.

For the crane systems, AI helps in predicting the movement of objects. By looking at the start of the container flight path during the vessel operations, AI will contribute to predict where the container will pass the portal legs. This calculated prediction will be used to move the BoxCatcher to the right window faster and more accurately.

FIGURE 2: Examples of numbers accurately read using Camco’s deep learning models.

There is also an indirect impact on operations. Based on some straightforward classifications, such as determining the door side of a container, our engine can decide if seal detection is necessary or where it has to look for the container number. See FIGURE 1.
When tracking and predicting the flight path of the container, this data is valuable to improve the operational skills of the quay crane driver. Cutting travel distance of the spreader/crane trolley shortens the cycle time of the quay crane and improves the hourly performance of the quay crane resulting in shorter vessel turnaround times.

**Hardware and software tuned for optimal results**

The use of deep learning and AI requires a lot of processing capacity. AI and deep learning use complicated algorithmic models and processing of matrix data. The hardware and the software need to work in perfect harmony.

Fortunately, these deep learning algorithms can be processed very efficiently on GPUs (Graphical Processing Unit) of computers. Actually, these algorithms run easily up to 100 times faster on a GPU compared to a CPU. Companies such as NVIDIA make special GPU based hardware specifically for AI processing.

For the processing of the images there are four options: processing on a central server, processing on the local camera CPU, on a GPU cloud (several companies as Amazon, Google and Microsoft have large farms of GPU processors) or a local GPU.

Camco’s AI applications are high demanding in function of speed (real-time application), latency (we can’t cope delays) and reliability. To meet those requirements Camco integrated a performant NVIDIA GPU processor inside a new series of intelligent cameras. This is a unique approach as image sensor and image analysis hardware are only a few centimeters separated from each other. The new cameras are part of the third generation OCR portal and BoxCatcher products.

**Camco has developed new cameras, with integrated NVIDIA TX2 module**