

Using computer vision for defect classification and anomaly detection

Lighting manufacturer Signify explores the use of the Intel® Distribution of OpenVINO™ toolkit to create a new approach to quality control on the production line.

Authors

Philip Newton
Scientist, Signify

Roy Engelen
Scientist, Signify

Ivo Durlinger
Principal Scientist, Signify

Marc de Samber
Fellow Scientist, Signify

David Farrow
EMEA Industry Technology Specialist,
Intel Corporation

Jamal El Youssefi
CE Application Engineer, Intel
Corporation

Erik Janssen
ADP Engineer, Axis Communications



Executive Summary

Signify's quality control process was highly dependent on human inspection. It's increasingly difficult to find people with the skills and experience to detect the tiny defects in the lamps Signify makes. So, the company carried out research to see how computer vision might help.

The machine learning model was trained using images of 11 defective and 40 good lamps. Using an Industrial Axis camera, images were captured from all around the lamps. The images from defective products were labeled by human experts, while the good products were labeled as OK automatically.

The Intel® Distribution of OpenVINO™ toolkit was used for anomaly detection and for defect classification. The study found that the Intel® Movidius™ Myriad™ X vision processing unit (VPU) offered a good performance level, the Intel® Core™ i7 processor offered better performance, and the best performance was achieved by the integrated Intel® Iris® Xe Graphics processor.

Before being deployed on the production line, the model needs to be refined with more training data to avoid the model activating on variations in the lamps that were not faults.

Business Challenges

It's often hard to find defects as products roll off the manufacturing line. That's why Signify wanted to explore the potential of artificial intelligence (AI) and machine learning to streamline the process.

Signify makes lighting products, including High Pressure Sodium (SON-T) lamps marketed under the Philips brand. They're used in greenhouses to help plants flower.

The first step in the quality control process was for electrical and geometrical tests to be carried out in-line automatically. Products that passed these tests were then inspected offline by experienced quality operators. Defects occur rarely, and the types of defects vary. Some defects are extremely difficult to see and rarely occur, but can lead to the product failing or being rejected in the field.

The manual inspection process incurs delays. Qualified people are becoming harder to find, and it's difficult to transfer their experience and skills to others. As is common with processes depending on human judgment, there can be errors. There might be false positives where products are erroneously treated as faulty, or false negatives where faulty products are cleared.

Signify wanted to use computer vision to improve the efficiency and accuracy of quality control, preventing false positives and negatives. The company also wanted to provide a method that not only automates the pass/fail decision, but also provides insight into failure types and frequencies so the production process can be optimized.

Technical Challenges

Signify would need to overcome a number of technical challenges in order to use computer vision for lamp inspection.

The burner, which is the light-emitting part of the lamp, needs to be inspected from all sides. It is shaped like a tube and is made of diffuse-transparent ceramic.

Defects can be smaller than a millimeter. They show up as visual features, of various types and with various aspect ratios and sizes. They can be at various locations on the outside of the tube, on the inside of the tube, or inside the ceramic material of the tube. Defects range from small cracks to small residues (fibers or particles), which show up as black or gray discolorations in images of the burners.

Training the AI Model

The first step in training the AI model was to capture video footage from the burners as they were rotated, so that they could be inspected from all sides.

Signify planned to use images that showed the entire product to save time capturing and processing images. One of the challenges was to capture images that could

show the whole product but also show the small defects in sufficient detail. Initial trials were performed using a range of standard high-definition cameras and lenses, and also using a midrange 3D camera. These cameras could not capture the entire product without significant shape distortion and did not have sufficient resolution to identify the defects. To overcome these challenges, Signify switched to using an Axis P1367 Network Camera. It is a day and night camera typically used for surveillance.

Day and night cameras provide color images by day and black and white images when the light fades below a certain level. The camera is more light sensitive in night mode because the infrared filter is removed. In this mode, the camera captures images more quickly. The anomalies in the images don't have much color, so the black and white images are adequate.

Light from a certain direction will emphasize some anomalies but not others, so the direction of light is important. Back lighting from behind the burner travels directly into the camera. This makes the overall scene very bright but yields a darker subject because the camera adjusts. In extreme cases the dynamic range of the camera can be exceeded, leading to a likely loss of detail. Light from behind the burner should be limited to just enough to highlight details without affecting exposure too much.

The Axis camera processes images with an emphasis on preserving detail to maximize the forensic value of the image. This is also effective for visualizing the important details in this industrial application.

Two approaches were used to train a YOLOv3 real-time image detection model (see Figure 1):

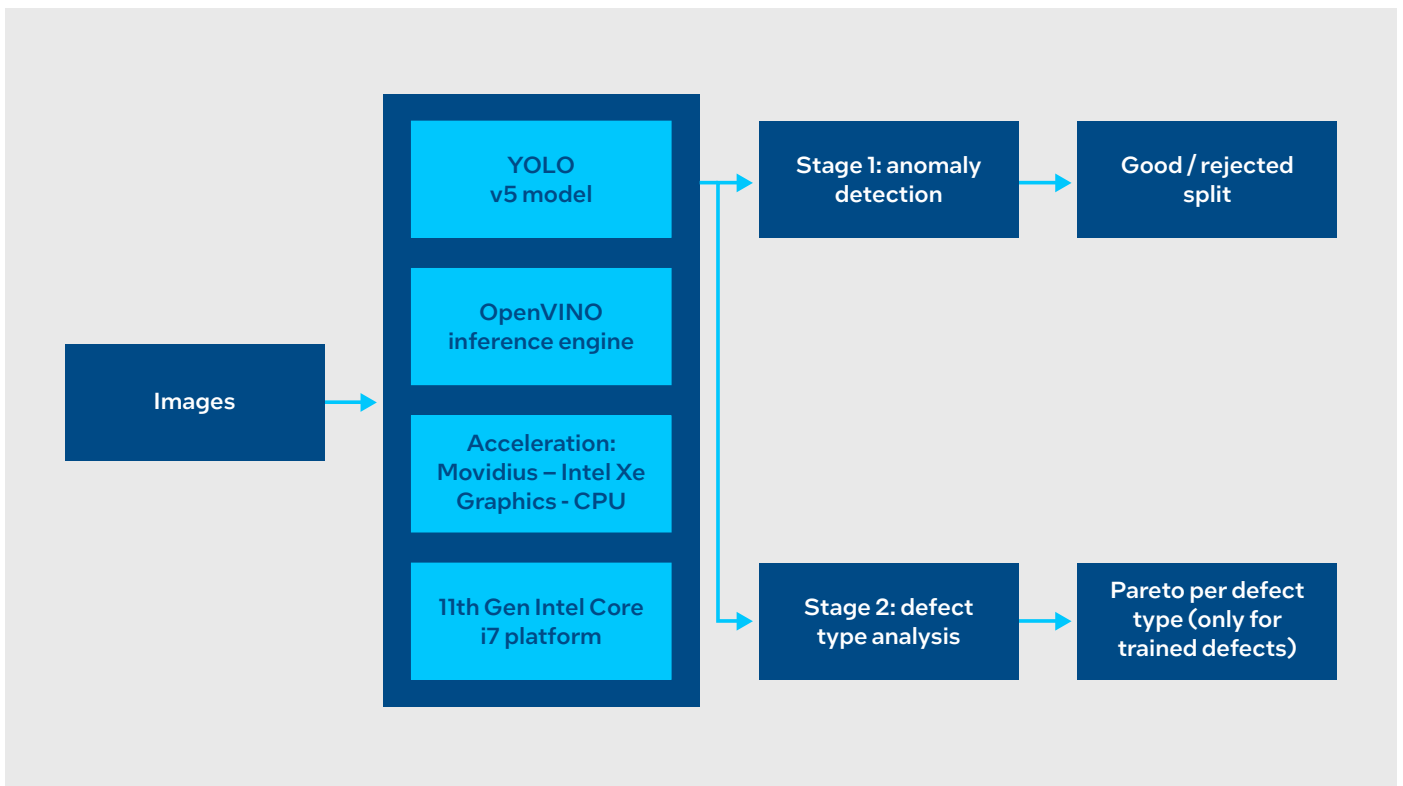


Figure 1. The training and inference processes for detecting anomalies and defects used by Signify.

- **Anomaly detection using unsupervised learning.** In this approach, there is no need to label the samples. It only requires good samples as data input. This approach can provide an indication of whether a burner is good or bad (which means it deviates from what a good burner should look like).
- **Defect detection using supervised learning.** In this approach, each type of defect must be labeled, and there needs to be a large number of burners and defects to use as input. When used to analyze lamps on the production line, this approach is able to provide insights into the defect type(s) identified on the burner. The quality assessment is fed back to the quality, engineering, and production teams.

Videos were captured of 11 defective burners (one for each defect type) and 40 good burners. Two videos were made of each burner using two different light settings (back lighting and side lighting).

From each video, 512 equidistant frames were generated, representing a single full rotation of the burner. As a result, there was a total of 11,264 images from defective burners and 40,960 images from good burners¹.

The images from good burners were automatically labeled as OK, but the images from defective burners were labeled by a human expert. For each defect type and for each light condition an entire set of images was presented in a random order to the human expert. The expert marked each image as either 'OK' (no defect visible) or 'not OK' (visible defect). These assessments were carried out twice. The assessments took place on different days to eliminate the effect of the experts remembering how they classified the images previously. Images with different classifications from both assessments were inspected again manually to determine the proper classification.

Deploying the AI model

The Intel Distribution of OpenVINO toolkit was used for both anomaly detection and defect classification when analyzing previously unseen burners. OpenVINO is a free software toolkit that accelerates the performance of deep learning inference from edge to cloud. OpenVINO supports execution across multiple processors and accelerators using a common API for C, C++, and Python. Supported processors and accelerators are CPUs, GPUs, field programmable gate arrays (FPGAs), and vision processing units (VPUs).

At the start of the project, there was no hardware specification chosen. In the final deployment, the hardware requirements would depend on the model performance, latency limitations, and number of cameras to be deployed. To provide flexibility, the project explored the option of using accelerators. The inference application can be deployed in a system with a combination of a host processor accelerated by Intel Iris Xe Graphics, Movidius Vision Processing Units, or Intel® FPGAs. OpenVINO works with whatever accelerators are available, so it's easy to switch between these accelerator options.

The OpenVINO model optimizer converted the YOLOv3 model based on the Darknet neural network framework into an intermediate representation (IR) format. The IR format is used by the OpenVINO Inference Engine, which can use models in different formats with various input and output formats.

The Signify platform was designed to work with Ubuntu 20.04 LTS or more recent versions.

Results

To help choose the final hardware specification, Signify measured the inference speed of the CPU, the integrated Intel Iris Xe Graphics processor, and the Intel Movidius Myriad X VPU.

The VPU is designed to work within a smaller power envelope than a CPU, and was able to carry out inference at a good speed of 8 frames per second (FPS). The Intel Core i7 processor could process 25 FPS using the CPU alone, but this was accelerated to 85 FPS using the integrated graphics processor².

The test showed that all three deployment options would be fast enough for Signify's requirements.

Lessons Learned and Recommendations

Signify found that not all defect types can be photographed properly using a single lighting condition. For most defects the back lighting works, but in some cases side lighting is best. It's a good idea to define which lighting condition is most suitable for each defect type. Separate datasets can be used for those defects that are best shown with back lighting and those best shown with side lighting. This saves time in preparing and classifying images, before the datasets are used for AI training.

The project discovered that the normal manufacturing variation between burners is large compared to the variation due to defects. The model sometimes activated on areas of the burner outside of the defective areas because the model had learned from other features on the burners that happened to coincide with the type of defect tube. For that reason, future projects should use more images from the various defects from multiple burners to provide a larger sample for training. Capturing more images would provide a more balanced dataset for training the neural network. Future work will look at using object detection rather than image classification to improve results.

A standard 2-megapixel camera is adequate, so Signify plans to use the lower resolution Axis P1375 Network Camera in its deployment. Lower resolution cameras of the same generation have better light sensitivity compared to higher resolutions. Light sensitivity is important because, given an amount of light, a more light-sensitive device will allow for faster image capture and thus a faster moving production line. That said, light is a controllable parameter in the setup so one could simply add light to allow for faster capture.

The testing was carried out using an 8th Gen Intel Core i7 processor, but thanks to improvements in Intel processors, Signify is now able to use an 11th Gen Intel® Core™ i5 processor with Intel® Graphics for its deployment.

The next steps require investigating the speed and accuracy of the computer vision model and then look at implementing the solution on the production line.

Conclusion

Signify's research found that computer vision can be used to streamline quality control on the production line. Before deployment, the model should be refined with more training data to improve the accuracy of defect detection. Signify can use the CPUs in its existing industrial PC hardware, including the integrated graphics processors, to achieve computer vision at the speed required.

Learn More

- Signify
- Axis Communications
- Intel® Distribution of OpenVINO™ toolkit
- Intel® Movidius™ Myriad™ X Vision Processing Unit
- Intel® Core™ i7 processors
- Intel® Iris® Xe graphics

Find the solution that is right for your organization. Contact your Intel representative or register at **Intel IT Center**.



Performance varies by use, configuration and other factors. Learn more at www.Intel.com/PerformanceIndex.

¹ 11 defective burners x 2 videos with different light settings x 512 frames = 11,264. 40 good burners x 2 x 512 = 40,960.

² Configurations: 11th Gen Intel® Core™ i7-1165G7, Intel® Iris® Xe Graphics, Intel® Movidius™ Myriad X. Software: YOLOv5s.

Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. See configuration disclosure for details. No product or component can be absolutely secure.

Intel does not control or audit third-party data. You should consult other sources to evaluate accuracy.

Intel technologies may require enabled hardware, software or service activation.

© Intel Corporation. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others.

1022/JW/CAT/PDF♻️PleaseRecycle 353487-001EN.