AI APPLIED TO FIBER OPTIC METROLOGY

Data-Pixel, a Seikoh-Giken company

ABSTRACT

Automated cleanliness inspection of optical fiber endface is a critical and challenging vision task that can benefit from deep-learning enabled microscopes. This new technology revolutionizes the traditional inspection methods. It offers unmatched solutions to meet the production quality requirements and constraints of the fiber optic industry, whether they are covered by standardization rules or not.

Document Date: April 2021
Author: Julien Maille

Address Information

DATA-PIXEL,
27 rue Saturne, ZAC Altaïs
74650 Chavanod, France
Phone: +33 4 50 67 39 80
Web: https://www.data-pixel.com

THE IMPORTANCE OF CONNECTOR END-FACE CLEANING

Networks are under constant stress with the incredible growth in demand for bandwidth, recently accentuated by the increase in teleworking (total international bandwidth tripled since 2016, according to Telegeography[1], +35% in 2020). In that context, the quality of connections is a critical requirement for the performance of optical communication networks. It remains the leading cause of fiber related downtime and failures in data centers or telecom applications.

Cleanliness of end face terminations has a direct impact on connectors performances. Installing a compromised connector will contaminate or even damage the mating connector.

With a core size of only 9 microns for single mode fibers, even a microscopic scratch, debris or any other contaminations can degrade the connection by blocking the light beam or creating airgaps that prevent physical contact. Low loss (IL) and low return-loss (RL) connectors do not tolerate such contamination and won't achieve the expected performance if not clean.

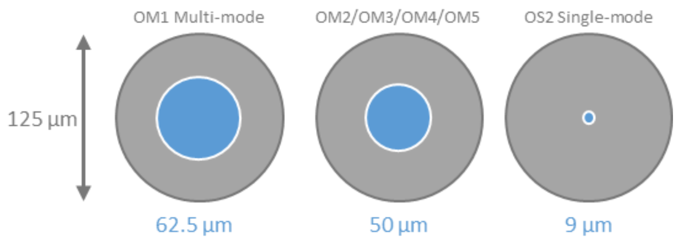


Figure 1 Optical fiber core diameters

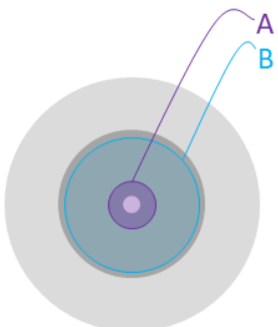
Industrial standards for cleanliness

Study by the International Electronics Manufacturing Initiative (iNEMI) has shown that contamination significantly increases insertion loss (up to 10 times), decreases return loss (up to 3 times), and increases Bit Error Rate Test (BERT) results (2-10 times)[2]. The analysis showed that scratches applied to the fiber mode-field diameter resulted in an increase of up to 25% of RL.

The International Electrotechnical Commission (IEC) provides a standard defining the areas of focus and their failure criteria. Non removable defects are categorized into 2 groups: scratches and defects (non-linear features). Depending on their width and location, a specific number of defects inside each group is allowed.

Compliance with the limit values defined by IEC 61300-3-35 Ed.2 guarantees the level of performance.

Acceptance criteria for single mode connectors			
Zone	**Region**	**Scratches**	**Defects**
A: Fiber core	0 – 25 µm	None	None
B: Cladding	25 – 115 µm	No limit ≤ 3 µm None > 3 µm	No limit ≤ 2 µm 5 from 2 - 5 µm None > 5 µm



While the IEC provides guidelines that aim to eliminate human subjectivity, counting, categorizing, and determining the size of each feature still leaves room for human error and inconsistency across manufacturing processes.

[1] https://blog.telegeography.com/internet-traffic-and-capacity-in-covid-adjusted-terms
[2] http://thor.inemi.org/webdownload/newsroom/Presentations/OMI/3.42BerdinskikhPaper.pdf

A DIFFICULT TASK FOR 'TRADITIONAL' INDUSTRIAL VISION

Large inspection area — small features

The challenge of endface inspection lies in the compromises faced by equipment manufacturers: microscopes need to be affordable, easy to use and keep a small form factor. At the same time, their optical system needs to capture a 300x300µm surface (single fiber connectors) or even a 3000x500µm (MT-24 multifiber) while maintaining a resolution small enough so that features of 1µm or less are not only visible but also reliably detected.

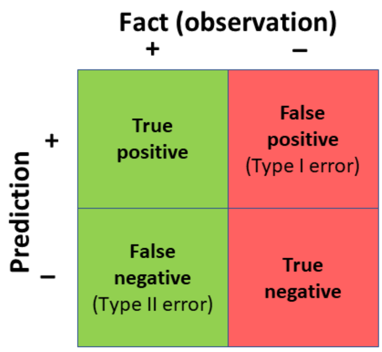
Typical high-resolution microscopes on the market will offer a 10x to 20x optical magnification, providing images with a pixel size in the range of half a micron. The software will then feed this image to a specialized algorithm in charge of locating the fiber, detecting, and classifying the defects. Defects will greatly vary in term of shapes, color, and contrast. Shallow scratches only visible to the eye of the trained expert, needs to be consistently detected by the vision software.

High sensitivity — even higher variability

These low contrast features, which sometimes differ by only 1 or 2 grey levels (on a scale of 255 for 8bit images), require software engineers to design algorithms that can separate weak signal from noise, and trigger detection with a very high sensitivity.

On the other hand, feature extraction cannot be too specialized, because contamination created in production and on the field comes in various colors and shapes. Different polishing processes will result in variable reflectivity, and microscopes will produce images with variable contrast.

Inspection algorithms are pushed to their limits in such a way that they can report false positive or false alarm, i.e., defects that do not exist or belong to acceptable features. This can happen, for instance, with bend insensitive fibers that show white rings around the core, bright spots due to internal reflection or white boundaries that can often be seen around big defects.



Confusion matrix

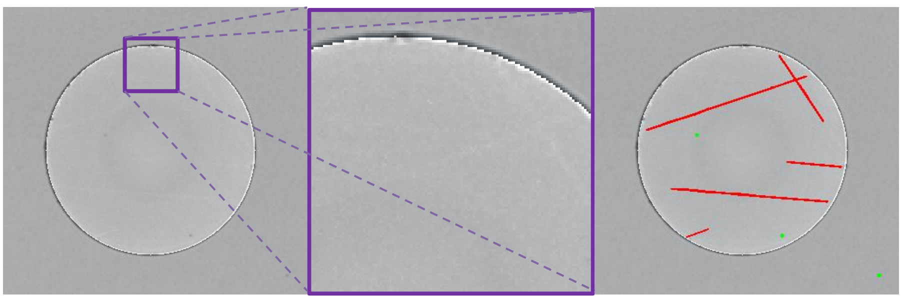
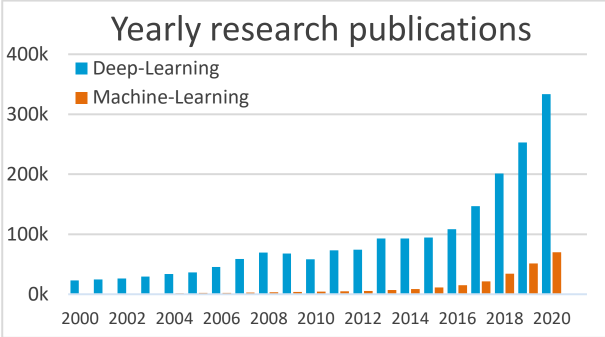


Figure 2 Example of faint scratches.
From left to right: raw image, close-up, and detection result

DEEP-LEARNING TECHNOLOGY



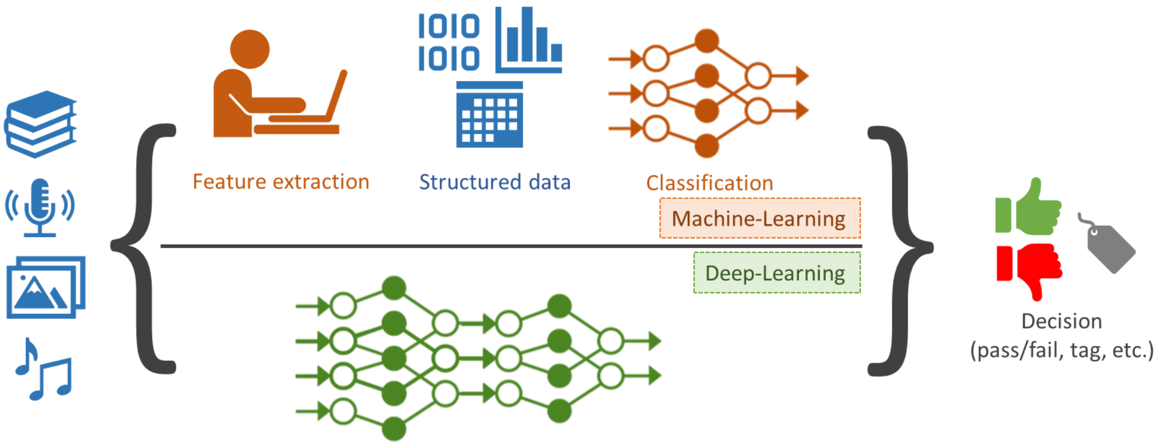
Deep-learning has been increasingly popular[3] and successful in many fields, enhancing the pictures taken on our phones, guiding self-autonomous cars or helping medical diagnosis. Now the same class of algorithm is replacing manufacturing process like quality inspection where automated judgement is required.

Machine learning or deep learning?

Unlike traditional software where a vision expert is programming all the rules, machine learning (ML) is enabling a machine to learn from data: by feeding the algorithm data and providing a feedback loop, it will adjust iteratively to improve its accuracy.

Deep learning (DL) is a subset of ML algorithms capable of mimicking the actions of the human brain through neural networks, hence the term Artificial Intelligence. The main technical difference lies in the fact that:

- ML algorithms can process quantitative and structured data: numerical values.
- DL algorithms can process unstructured data such as sound, text, or images.



In the following sections, we will explain how software can extract meaningful information from images using convolution filters, the building blocks of any vision toolbox. We will then see that defect segmentation can be performed with deep learning techniques leveraging on those convolution filters.

Feature extraction is the action of reducing the complexity of the data by using domain knowledge. In our field, it means writing a software algorithm that replicates expert techniques to pre-process data and make patterns more visible to vision algorithm. With deep-learning, feature extraction is performed by the model itself: the algorithm will be trained to output the key elements that will determine the prediction we want it to make. In computer vision this is typically the edge extraction, shape recognition, pattern detection, etc.

During training the first layers of the neural network will naturally perform this task. The drawback is that it will requires many layers (hence the "deep") for the network to be powerful. Having literally millions of parameters that

[3] Number of publications extracted from https://app.dimensions.ai and compiled in April 2021.

need to be optimized, deeper networks are harder to train and more computationally expensive, both when trained and used for prediction.

Convolution & image analysis

Convolution filters have been a tool of reference in image processing for decades. With images represented as 2D matrix of numbers (one value for each pixel), these filters, or kernels, are small matrix of weights that slides over the image and perform a local weighted sum, as illustrated in Figure 3.

A basic kernel of 3x3 weight can blur or sharpen an image, perform edge detection, noise reduction, shape recognition, etc. Kernels are usually handcrafted by software developers, tailored to fit a specific problem. They can be stacked to solve more complex tasks, but it quickly becomes a nightmare to fine tune.

Convolutional Neural Networks (CNN) came from the assumption that these kernels can be learnt automatically, allowing programmers to stack them at a much higher scale.

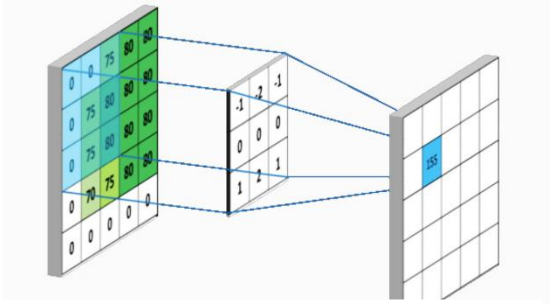


Figure 3 Matrix representation of a convolution

CNNs are made up of neurons with learnable weights. Each of them is a small kernel filter which receives a small input matrix, performs a weighted sum over these numbers, pass it through an activation function and returns an output. The first layers of the network operate on a local neighborhood (because of their small 3x3 pixel size) but since these layers feed other layers of kernel filters, the deeper ones can construct a global representation of the full picture by assembling representations of each small parts in a pyramidal way.

Semantic segmentation

CNNs have become widely popular in the last ten years and showed outstanding performances in 3 different fields:

- Classification: 1 image → 1 predicted class
- Detection: classification with bounding boxes
- Segmentation: 1 pixel → 1 predicted class

At a low level, endface inspection consists in locating objects in an image i.e., marking each pixel as a defect, a scratch, or the background. This field of computer vision is called semantic segmentation[4] and recent academic research has shown great contributions with very popular deep convolutional neural networks being published and becoming *de facto* the building blocks of all modern deep networks dedicated to object segmentation.

Many models have been released within the common deep-learning frameworks (like Tensorflow or Pytorch) and experimenting with a few of them is best advised before one decides to develop a custom CNN. VGG, ResNet, MobileNet are some of the popular models which have shown high accuracy in the ImageNet[5] recognition challenge, a popular benchmark.

[4] A Review on Deep Learning Techniques Applied to Semantic Segmentation - https://arxiv.org/pdf/1704.06857.pdf
[5] ImageNet dataset - https://www.image-net.org/challenges/LSVRC/

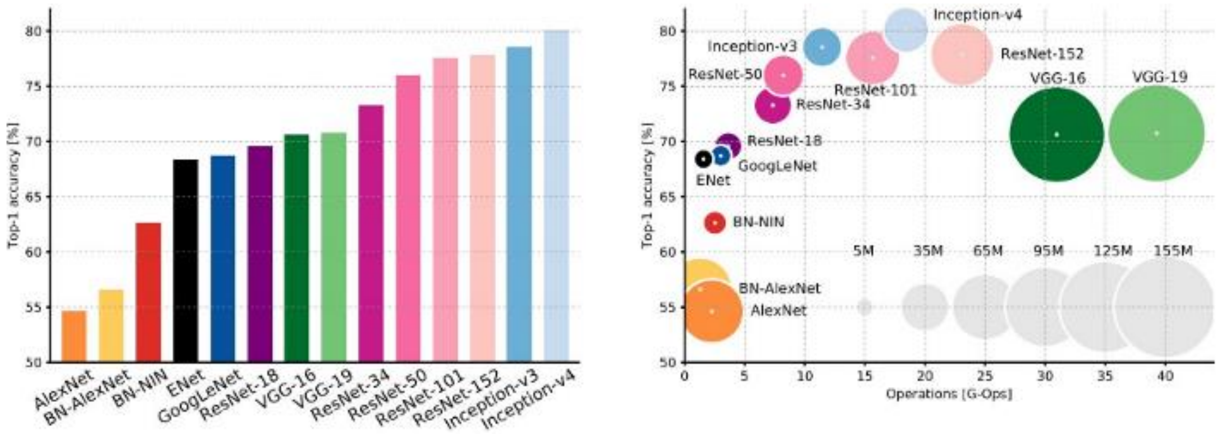


Figure 4 Comparison of popular CNNs accuracy and size

A model is only as good as its training dataset

One of the critical parts of any machine learning project is the available training dataset. Its size, content and quality will sometime determine which training techniques are possible. This step takes time, domain expertise and the right tools to collect, manage, sort, and label the pictures.

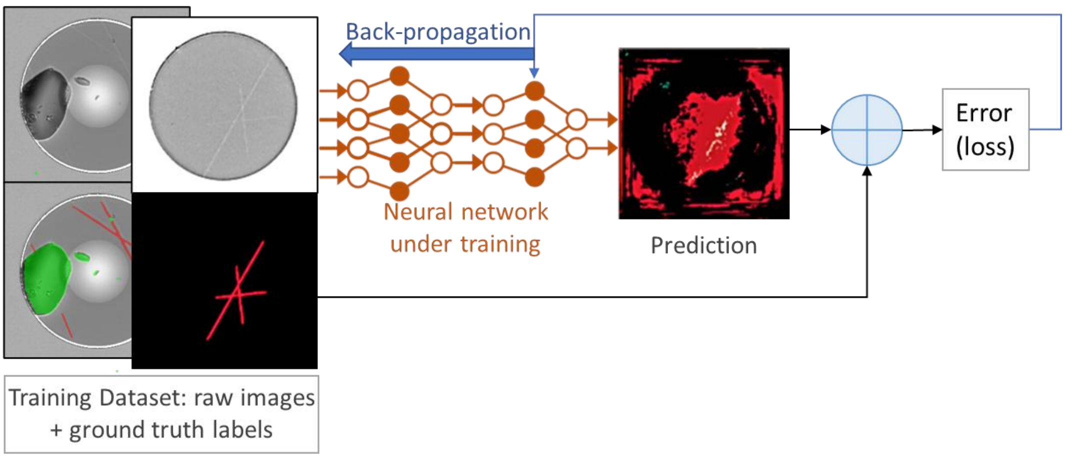
Our initial approach was to define 2 classes (scratch & defect) and gather a representative dataset of endface pictures for which several experts agreed and labelled the location and class of all anomalies (ground truth). Each pixel of each image of the dataset has been assigned a class or defined as a background pixel.

Training with labels is called supervised training. Another approach called unsupervised training only requires clean pictures, but it usually outputs imprecise heatmaps of defects and does not achieve the "pixel perfect" precision we need for endface testing. Once the training dataset is ready, the network can be trained.

Training workflow

Training a network is performed by running an optimization algorithm in charge of iteratively minimizing a cost (or loss) function: this function is a method of evaluating how well the algorithm models the dataset and predicts correct results. It is up to the data engineers to define the right loss function that will suit their use case.

The training process starts by feeding the network with an image from our training dataset, then comparing the predicted segmentation with the correct labels. The mismatch between the prediction and the ground truth allows computation of the loss and results in a correction that is sent back through the layers of the network.



This back propagation will adjust the networks weights of the convolution filters. With each iterations, the loss should decrease as the network converges, eventually stopping the feedback loop when predictions are getting very close to the ground truth. Models with millions of adjustable weights typically requires thousands of iterations until they give acceptable prediction, this can take severeal hours of computation, even on modern optimised computers.

Figure 5 Evolution of predictions of the network during training

Inference

Inference occurs when a previously trained model is fed with an image in order to generate a prediction. In our case, that is when our customers use it in production through Blink, our measurement software.

The model defined earlier outputs 2 maps (see Figure 6): one for scratches, another for defects. Each of them will allow the post-processing software routines to locate, measure and count defects according to the IEC pass/fail criteria.

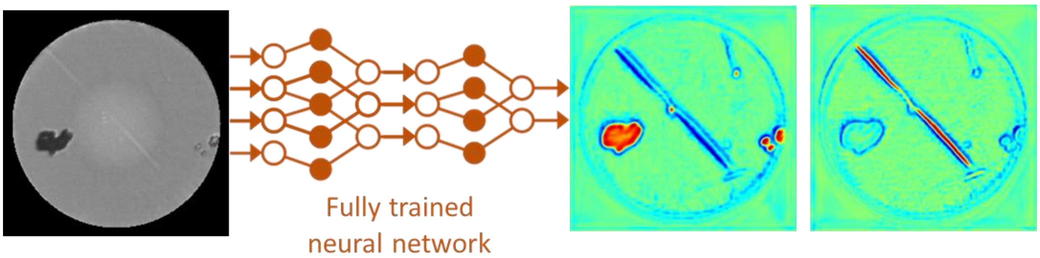


Figure 6 The network receives an image from the microscope and produce heatmaps for each class of defects

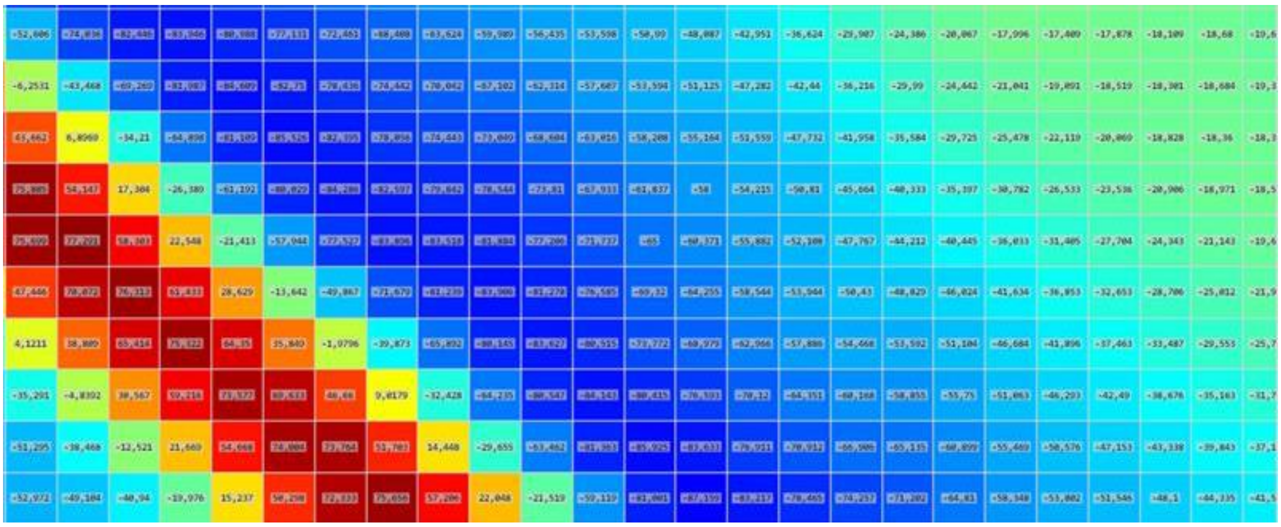
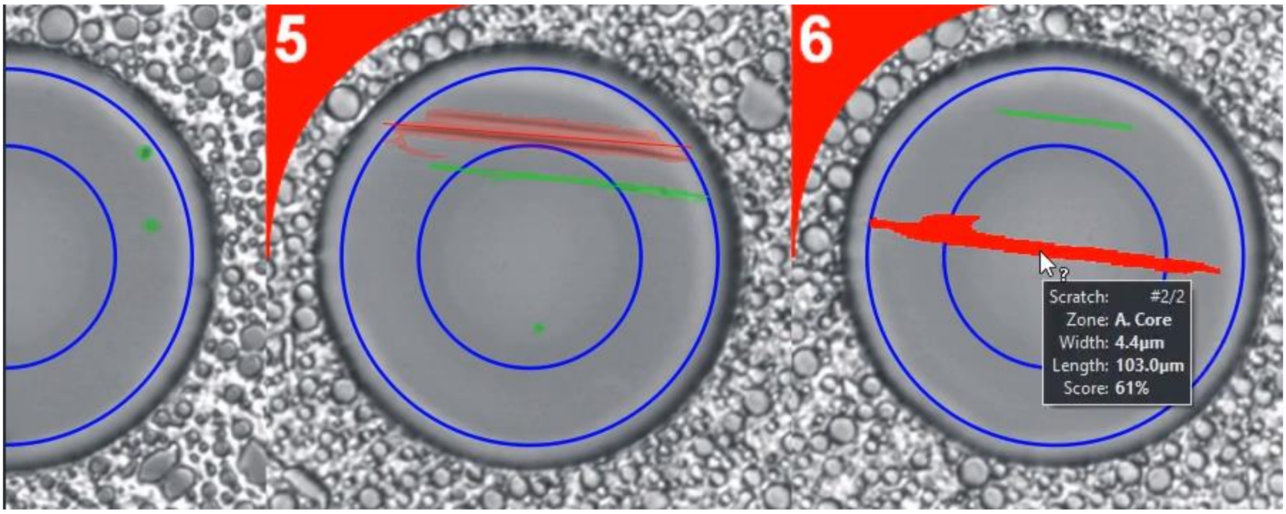


Figure 7 A zoom on the scratch heatmap shows the score of each pixel

As shown in the Figure 7, these maps are not "black or white" binary images, for each pixel a defect score is obtained, correlated to its probability of being a defect of given class, according to the way the network was trained. It means we can set a strict threshold (near zero) or relaxed (80% or higher) to segment our image. This allows the final user of the inspection software to adjust the level of sensitivity he would like to achieve for each product.

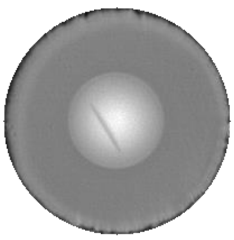
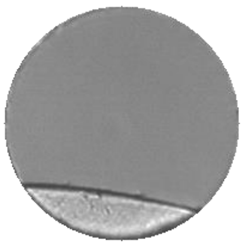
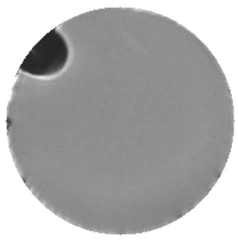
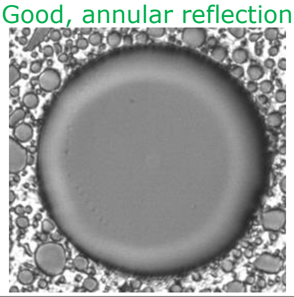
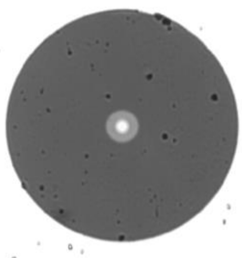
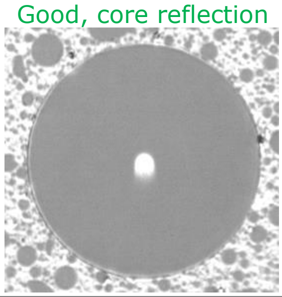
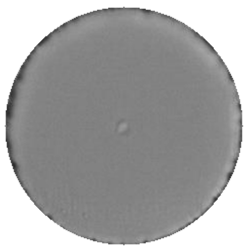


REPLACING THE HUMAN EYE

Retraining for special use cases

While traditional handcrafted kernel filter is challenging to readjust and overcome false positive or undetected defects, deep-learning CNNs can easily be retrained on a dataset featuring these new special cases. We have been successful learning new types of defect that are not defined by the IEC guidelines like unpolished or broken fiber, core cracks or fiber chips.

On the other hand, unexpected variation that can cause false positive were easily retrained including saturated bright cores, partially filled cladding and bend insensitive fibers cladding.

Core crack	Broken fiber	Unpolished	Chipped fiber
			
Good, annular reflection	Dirty bend insensitive	Good, core reflection	Good
			

Fast and repeatable measurements

Deep learning is known for being a computationally intensive task requiring fast computers. While our implementation will benefit from latest processors and graphic cards (GPUs), it will also run on any standard, recent computer.

Type	Fiber screening	+ Endface view
MT12	**4 sec.**	6 sec.
MT24	**5 sec.**	7 sec.

Figure 8 Total scanning time

Measurement speed does not show any degradation over standard image processing techniques: inspecting a typical MT-12 fibers connector takes less than 4 seconds[6], including autofocus and image acquisition. An MT-24 can be scanned and processed within 5 sec. This is substantially faster than visual screening by an operator who needs to carefully scroll through the fibers and assess them one by one, while making sure he doesn't miss one of them. With a typical screening time of 1sec/fiber, automated devices can bring a 3x speed improvement.

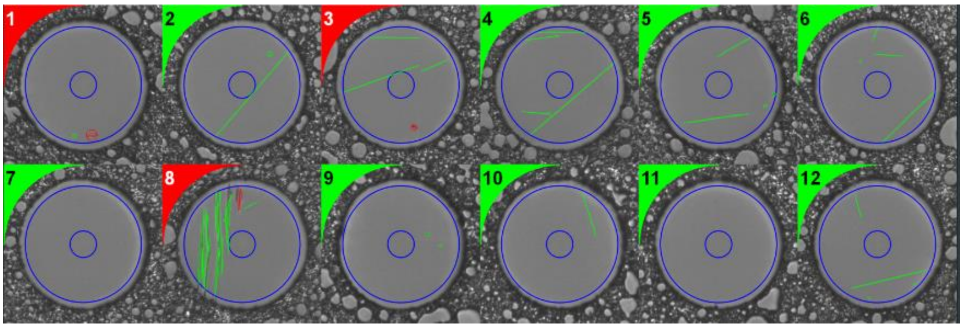


Figure 9 All in one view showing all inspected fibers simultaneously.

A fast measurement is only worth if it is correct and reproducible. Such a reproducibility study has been performed on several connectors selected for their representative defects. All connectors were measured 25 time with reinsertion. Our conclusions are as follows:

- For defects from 4µm to 10µm, the standard deviation of measurement is 0.2µm.
- For scratches & defects < 4µm, the standard deviation of measurement reaches 0.1µm.

Comparing this to human repeatability is tricky, partially because humans memorize and recognize defects when they see them multiple times. Although we are not aware of any published study, feedback from the industry shows it will be strongly affected by:

- Training level
- Illumination and focus adjustment
- Display-screen contrast and resolution

New tools for connectors manufacturers — upcoming applications

Because deep-learning vision algorithm are trained rather than programmed, any new vision challenge can be tackled by building a new dataset of pictures and their corresponding ground-truth labels. With this approach, Data-Pixel has developed automated inspection devices for the full endface of MT ferrules, observed at low magnification, a capability unparalleled on the market.

Due to the high variability of materials and polishing process, the pictures taken by microscope show a high variability in their texture, as shown in Figure 10.

[6] Benchmarks performed on an Intel i9 processor and a Nvidia GeForce 2070 graphic card

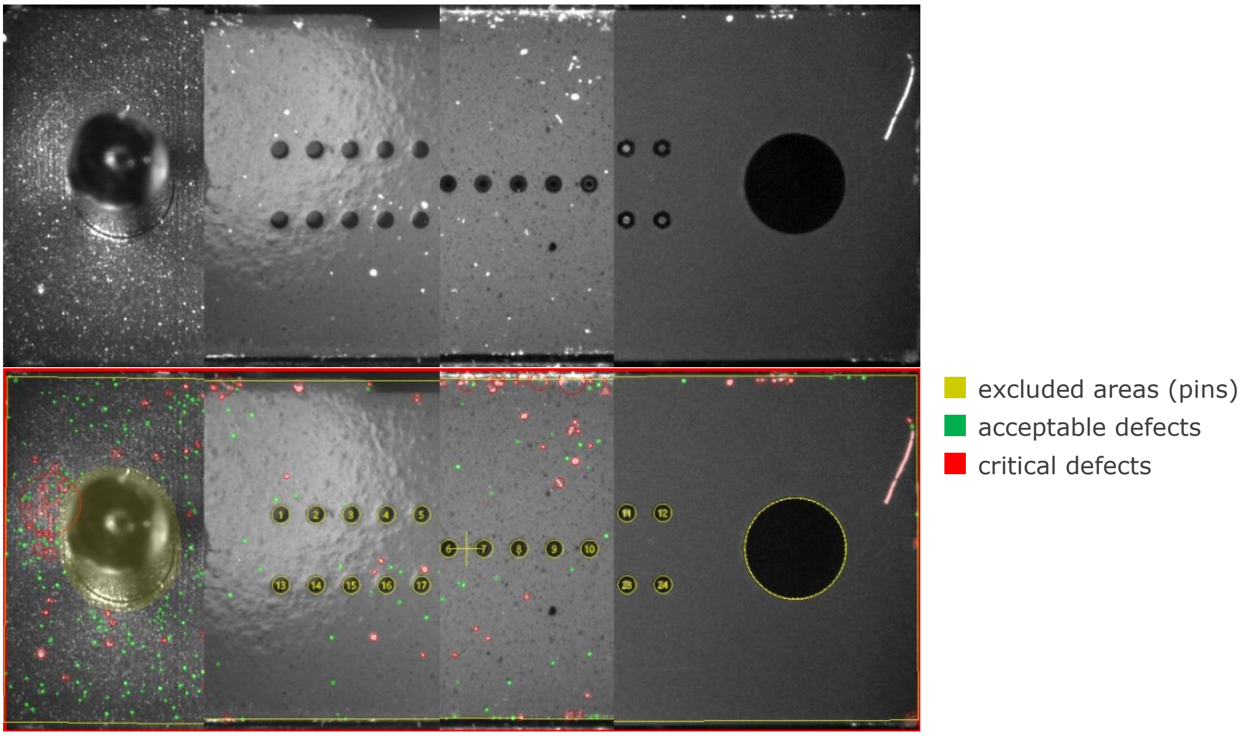


Figure 10 Stitched images of 4 MT ferrules with variable aspect

With the help of defect scores generated by the deep-learning model, customers can easily adjust the sensitivity of the inspection software. They can easily set a threshold on the size and surface of the defect and have the software automatically accept or reject.

Such a feature is an absolute requirement for vision tasks that are still not standardized but for which inspection is already performed at several steps on the production line or on the field.

Figure 11 Data-Pixel DScope EFI

CONCLUSION

All the advantages of deep-learning automated inspection show that automatic fiber end-face testing is now the most successful way to certify compliance with the IEC requirements. This significantly improved reliability is achieved without sacrificing speed and allows handling of new images or defect types through re-training.

Being trained with examples instead of being programed with rigid feature definitions, deep-learning vision enables inspection not only on standardized and well-defined defects/products but also on customer's own acceptance rules: it just requires them to provide pictures of good and bad connectors.

Within the world of optical interfaces, this new generation of software makes it possible to automate vision on large core fibers with varying cladding size and aspect, multicore fibers or crystal-photonic fibers, connectors with micro lens, etc.

The learning capacity of these AI based software is a continuous ongoing process without any limitation. We have seen in this paper the tremendous benefit it already brings and how it will continue to satisfy new quality requirement of the industry related to fiber optic endface inspection.