

## Device to evaluate cleanliness of fiber optic connectors using image processing and neural networks

Victor Fernandez, Javier Chavez, Guillermo Kemper

School of Electronic Engineering, Faculty of Engineering, Universidad Peruana de Ciencias Aplicadas, Lima, Peru

### Article Info

#### Article history:

Received Aug 19, 2020

Revised Dec 18, 2020

Accepted Jan 13, 2021

#### Keywords:

Cleanliness evaluation

Fiber optic connectors

Image processing

Neural networks

### ABSTRACT

This work proposes a portable, handheld electronic device, which measures the cleanliness in fiber optic connectors via digital image processing and artificial neural networks. Its purpose is to reduce the evaluation subjectivity in visual inspection done by human experts. Although devices with this purpose already exist, they tend to be cost-prohibitive and do not take advantage of neither image processing nor artificial intelligence to improve their results. The device consists of an optical microscope for fiber optic connector analysis, a digital camera adapter, a reduced-board computer, an image processing algorithm, a neural network algorithm and an LCD screen for equipment operation and results visualization. The image processing algorithm applies grayscale histogram equalization, Gaussian filtering, Canny filtering, Hough transform, region of interest segmentation and obtaining radiometric descriptors as inputs to the neural network. Validation consisted of comparing the results by the proposed device with those obtained by agreeing human experts via visual inspection. Results yield an average Cohen's Kappa of 0.926, which implies a very satisfactory performance by the proposed device.

*This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.*



### Corresponding Author:

Guillermo Kemper

School of Electronic Engineering, Faculty of Engineering

Universidad Peruana de Ciencias Aplicadas

Lima 15023, Perú

Email: guillermo.kemper@upc.pe

## 1. INTRODUCTION

Throughout the ages, humanity has faced the engineering challenge of transporting large quantities of information from one place to another, which has been faced through diverse means. During the 30s, the introduction of the coaxial and multipair cable technologies enabled transporting signals during many years. Later on, the increasing telecommunications demand and digitalization trends resulted in the development of a new technology which took off during the 60 s, known as fiber optics [1]. Meanwhile, other data transport technologies appeared such as microwaves and satellite communications, which use atmospheric properties for data transfer. However, each technology has its own operating conditions. For example, microwaves are severely affected by weather conditions and have limited capacity depending on the requirement [2]. On the other hand, the fiber optic communications requires that the connectors used to join the cables are completely clean and free of any kind of dirt to ensure the quality of the services. Otherwise, there will be a negative impact on the transmitted signal, which will impair the performance of the link [3, 4].

With this in mind, industry has developed technology both for identifying dirty connectors and for cleaning them. IEC 61300-3-35:2015 is the standard that sets the quality requirements for fiber optic connector terminations [5]. That is how the main problem to solve is develop a device that include strong

image algorithms to evaluate optical fiber connectors with high precision and low cost. In the scientific literature there are some proposals to solve the problem, Duffy *et al.* study in [6] automatized solutions to evaluate and clean fiber optic connectors. Rehman and Mozaffar [7] use wavelets to evaluate the dirtiness of connectors, while Filipenko in [8] use the interference method to analyze them. Commercially, companies such as VIAVI solutions [9] and EXFO [10] manufacture and commercialize inspection and evaluation equipment to verify the state of the fiber optic connector. Proposals in the scientific literature do not use image processing and neural networks, which can improve the method's precision. Likewise, the proposal by Duffy *et al.* uses a programmable logic controller (PLC), which is not the best alternative for a portable solution. On the other hand, commercial solutions imply making large investments which may not be viable for small and medium-sized telecommunications networks companies.

In the face of all these problems and drawbacks, the present work proposes a low-cost portable equipment capable of successfully evaluating and identifying the dirtiness of fiber optics connectors using an optical microscope, a single-board computer, open-source tools, image processing algorithms and neural networks. The device has a near-perfect assertiveness with evaluation results showing a Cohen's Kappa agreement index of 0.926. Furthermore, the equipment is well-suited to fieldwork requirements due to its portability.

## 2. DESCRIPTION OF THE PROPOSED DEVICE

Figure 1 shows the block diagram of the proposed algorithm. The monocular of the microscope has an adapter fixed to it, which safely holds the digital camera that takes the RGB image of the fiber optic connector. The following paragraphs describe the image processing steps and algorithms employed to achieve the desired results.

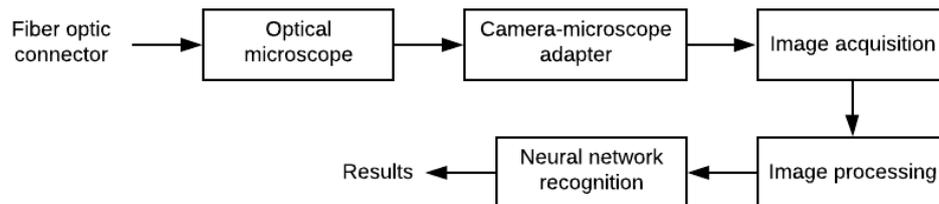


Figure 1. Proposed method flowchart

### 2.1. Optical microscope

It is composed of a manual device which allows visualization of the connector's ferrule and the contact area. Figure 2 shows the composing parts of a manual microscope. The eyepiece permits visualization of the fiber through an optical zoom up to 400X. The focus control handles the manual focus to improve image definition. The light source entrance helps the user note if the illumination LED is on or off.

### 2.2. Camera-microscope adapter and prototype device

The camera-microscope adapter consists of a piece, 3D printed in black PLA. This design is exclusive to the proposed solution and helps in creating a dark enclosure to have a stable, constant illumination when imaging takes place. Figure 3 shows a schematic of the adapter. The top piece is a removable lid that covers the digital camera, while the rectangular hole is the entrance for an HDMI connector to the camera. Figure 4 shows the camera installed in the rectangular area of the adapter see Figures 4(a) and 4(b). Figure 4(c) shows the HDMI cable connected to the camera. Figure 4(d) shows a picture of the bottom view, in which the camera lens is observed.

The adapter must comply with some minimum requirements to be compatible with the monocular in the microscope. Per the manufacturer, the minimum focal distance is of 3.6 mm. Then, the designed adapter will achieve a focal distance of 8 mm and be able to take images of the full fiber optic for subsequent analysis. Figure 5 shows a picture of the adapter, the microscope, the ensemble and the labeled parts. Likewise, the Figure 5(d) shows the complete box that contains the processor, the touch screen and the program that will allow the evaluation. The complete prototype has characteristics:

- Easy to connect as seen in the Figure 5. The equipment allows connection through the HDMI cable and the camera sends the image to the processor. All the data processed in the single board computer will be displayed on the touch screen in real time.

- The portability of the solution is based on its ease of transport: lightweight and easy to handle.
- The prototype uses an AC-DC power adapter. However, it has the option to deploy an internal battery.

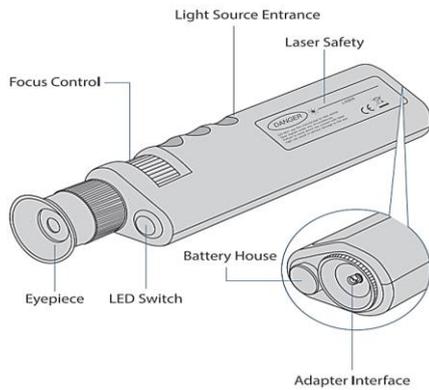


Figure 2. Optical microscope description  
Source: microscope quick start guide

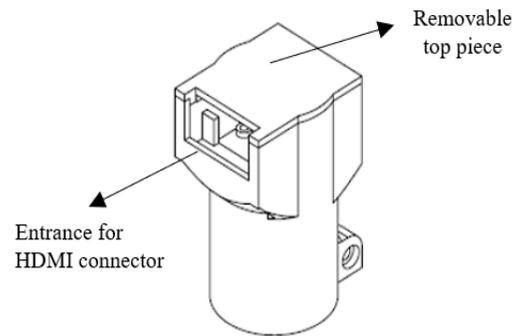


Figure 3. Adapter schematic diagram

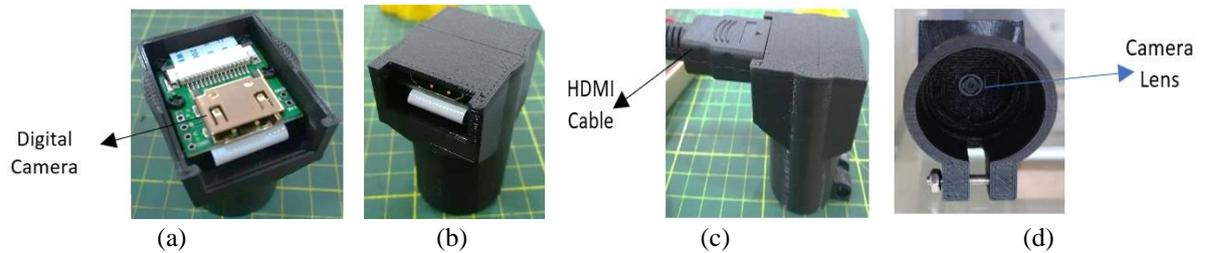


Figure 4. Camera-microscope adapter: (a) Camera installed, (b) Adapter with the top piece, (c) HDMI cable connected, (d) Bottom view

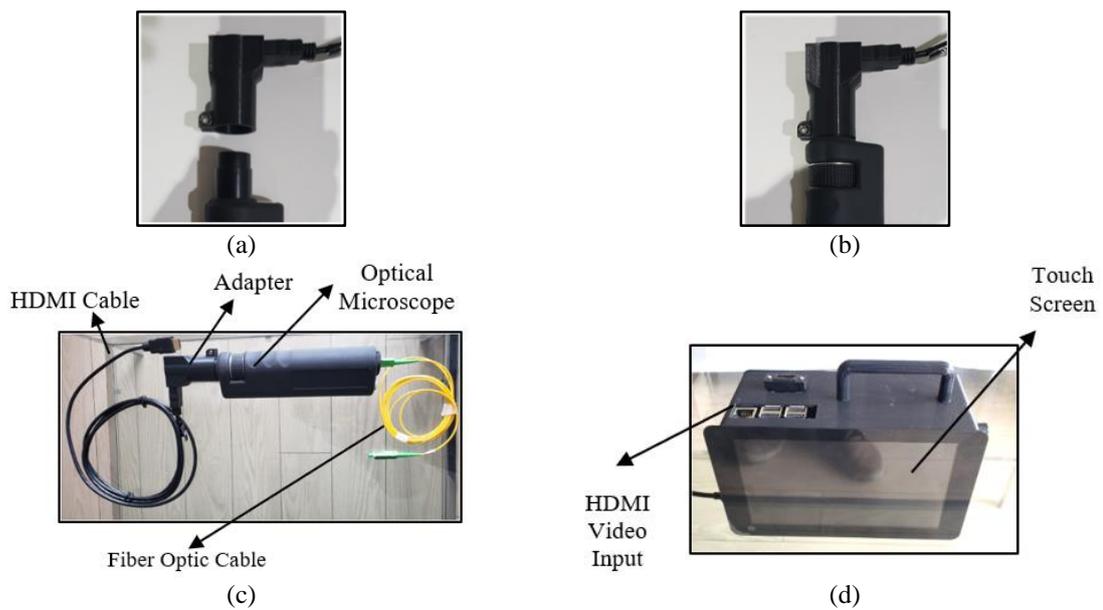


Figure 5. Adapter and prototype equipment: (a) Both elements separated, (b) Ensemble, (c) Parts and components for image acquisition, (d) Prototype box

### 2.3. Image acquisition

A 25x24x9 mm, 8 Megapixel digital camera with a Sony IMX219 sensor images the fibers. To allow the user real-time image visualization, the camera configuration was set to RGB (true color, 24 bits). Because of the small sizes of the dirt objects, the chosen image resolution was of 3280x2464 pixels. This translated into a spatial resolution of 0.0625  $\mu\text{m}^2$  per pixel, an appropriate size to adequately evaluate the objects. The default illumination intensity level was used since it is supplied by the microscope and kept stable by the adapter's enclosure. Figure 6(a) shows an example of an image acquired with the ensemble.

### 2.4. Image processing and descriptors

The following section details the steps taken to analyze the acquired image and obtain the three selected image descriptors to train the neuronal network and determine the level of cleanliness or dirt in the regions of interest.

#### 2.4.1. Image cropping

The image acquired in step 2.3 is cropped, taking into account the fiber core and cladding areas (area of interest for processing). The resulting image has a smaller size and its digitized primary components of light are expressed as  $I_R(x, y)$ ,  $I_G(x, y)$  and  $I_B(x, y)$ . The values of this components are 8 bit-integers ([0,255]) according with true color format.

#### 2.4.2. Grayscale conversion

In order to reduce computational load without diminishing the algorithm's performance, the image from Figure 6(b) is converted into an 8-bit grayscale image, according to (1) [11].

$$I_1(x, y) = 0.299I_R(x, y) + 0.587I_G(x, y) + 0.114I_B(x, y) \quad (1)$$

where  $I_R(x, y)$ ,  $I_G(x, y)$  and  $I_B(x, y)$  are the primary components (red, green and blue) of the cropped image, and  $I_1(x, y)$  is the grayscale image as shown in Figure 6(c).

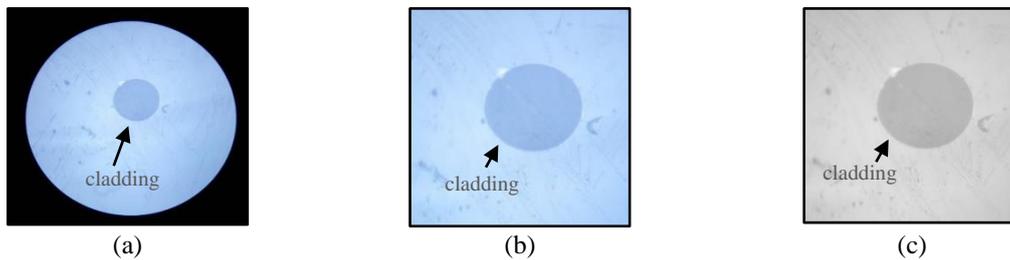


Figure 6. Example results of the grayscale conversion: (a) Acquired image example, (b) Image with primary components  $I_R(x, y)$ ,  $I_G(x, y)$  and  $I_B(x, y)$ , (c) Image  $I_1(x, y)$

#### 2.4.3. Histogram equalization

Image  $I_1(x, y)$  as shown in Figure 7(a) goes through a histogram equalization process [12]. The resulting image is labeled  $I_2(x, y)$  Figure 7(b). Histogram equalization showed better results than enhancement via histogram stretching, because it improved the later segmentation between the cladding region and the rest of the image.

#### 2.4.4. Gaussian filtering

Image  $I_2(x, y)$  is filtered with a 5x5 Gaussian mask [13] to eliminate undesired edges and contrast regions. Such mask is defined by (2) [14].

$$h_1(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} \quad (2)$$

In this case we consider  $\sigma = 1$  to obtain satisfactory results. The filtering process is expressed through convolution between  $I_2(x, y)$  and  $h_1(x, y)$  mask:

$$I_3(x, y) = I_2(x, y) * h_1(x, y) \quad (3)$$

Figure 7(c) illustrates the input image  $I_2(x, y)$  and the output image  $I_3(x, y)$ .

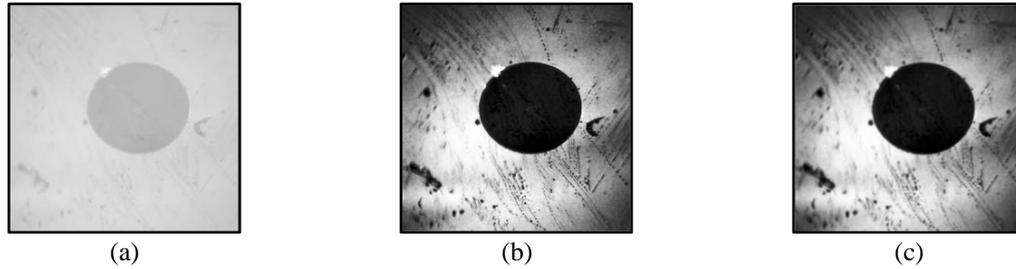


Figure 7. Example results of the histogram equalization and Gaussian filtering: (a) Image  $I_1(x, y)$ , (b) Image  $I_2(x, y)$ , (c) Image  $I_3(x, y)$

#### 2.4.5. Image thresholding

Image  $I_3(x, y)$  is thresholded in order to obtain the cladding segmentation mask, according to (4):

$$I_4(x, y) = \begin{cases} 255, & I_3(x, y) > \mu_1 \\ 0, & \text{any other case} \end{cases} \quad (4)$$

Any pixel with a higher value than the threshold  $\mu_1 = 104$  will be set to 255. This threshold was chosen because it yielded the best results for different types of fibers and levels of cleanliness. Figure 8(a) illustrates the output of the thresholded image,  $I_4(x, y)$ .

#### 2.4.6. Canny filtering

Canny filtering is applied to image  $I_4(x, y)$  to obtain the circle of the fiber cladding. The Canny filtering involve: smoothed using a Gaussian kernel, obtaining of edge strength from Sobel filtering, calculation and quantization of the edge direction, nonmaximum suppression and thresholding with hysteresis [15]. In this case, two hysteresis thresholds ( $\mu_{MAX} = 110$  and  $\mu_{MIN} = 50$ ) are set to obtain the shapes and borders of the objects in the thresholding. The result of this filtering is  $I_5(x, y)$ , shown in Figure 8(b). The chosen hysteresis thresholds enabled reducing false contours for different types of fibers and levels of cleanliness.

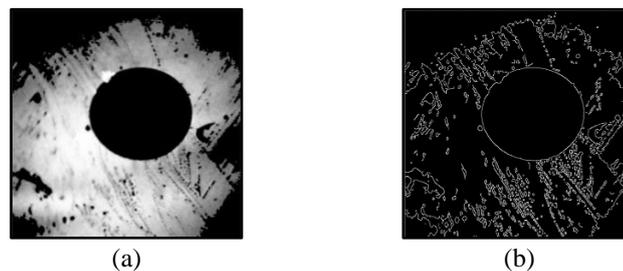


Figure 8. Example results of the thresholding and filtering image: (a) Image  $I_4(x, y)$ , (b) Image  $I_5(x, y)$

#### 2.4.7. Circle detection via Hough transform

The Hough transform for circle detection [16] is applied to image  $I_5(x, y)$ . In this case a 3-dimensional accumulator was used, storing the 2 coordinates of the center and the radius. In the implementation of Hough's transform, the parameter "dp" was set to 3, in order to resolve the search of a large circle with low distortion levels. Gradski [17] defines this parameter as the resolution of the accumulator image, which enables the creation of an accumulator with a lower resolution than the original image. Essentially, a larger "dp" value translates into a smaller accumulator matrix. The chosen value resulted in successfully detecting the circle which exactly belongs to the cladding, and greatly reduced the

false detection of circles. A larger “dp” value might have reduced the detected circle quantities to zero in some cases. The transform requires setting a minimum distance between the centers of the circles. If this distance is too small, false circles may be detected. If it is too large, circles may go undetected. It is known that the circle to be detected has a radius of 250 pixels, so this value serves as the starting point for the following computation.

If the value were chosen to be less than or equal to 250, most detected circles would be superimposed over each other. If the value were chosen to be 500 (the circle diameter), some circles tangent to the true circle would be detected. Figure 6 shows that the cladding is almost in the middle of the image, and the region of interest is not large, so lost circles outside of this region would not be relevant for this process. Thus, the distance between centers of circles is defined to be 1000, which reduced the quantity of possible circles in the selection process. The results of the transform are the circle radius  $r_c$  and the coordinates of the circle center  $(x_c, y_c)$ . With these coordinates and the circle radius  $r_c$ , the detected circles are drawn over the original image such that the regions of interest are defined. Figure 9 shows an example where results correspond to 3 case: Clean fiber as shown in Figure 9(a), dirty fiber as shown in Figure 9(b) and very dirty fiber as shown in Figure 9(c).

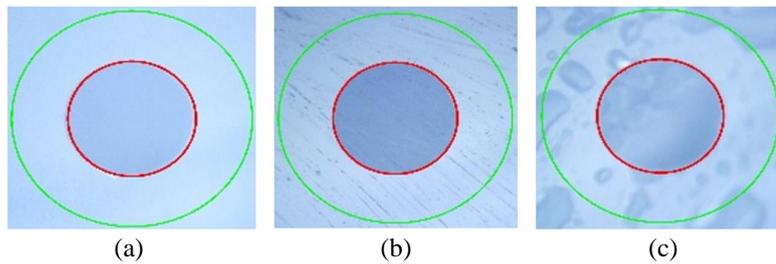


Figure 9. Example results of the transform: (a) Clean fiber, (b) Dirty fiber, (c) Very dirty fiber

#### 2.4.8. Regions of interest segmentation

The segmentation of the regions of interest takes the circle coordinates and radius from the previous step and generates a mask  $Z_1(x, y)$  the same size of  $I_1(x, y)$ , according to the (5).

$$Z_1(x, y) = \begin{cases} 1, & (x - x_c)^2 + (y - y_c)^2 \leq r_c^2 \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

where  $(x_c, y_c)$  corresponds to the center coordinates of the found circle and  $r_c$ , the radius of the found circle. Using the (6), the mask  $Z_1(x, y)$  segments the region of interest in  $I_1(x, y)$ :

$$I_6(x, y) = Z_1(x, y) \cdot I_1(x, y) \quad (6)$$

where  $I_6(x, y)$  is the output image after masking. Figure 10(a) shows an example of  $I_6(x, y)$ .

Since the fiber cleanliness must also be evaluated around the outer neighborhood of the cladding, the outer ring was also segmented. Thus, mask  $Z_3(x, y)$  was computed according to the (7).

$$Z_3(x, y) = Z_2(x, y) - Z_1(x, y) \quad (7)$$

where  $Z_2$  is defined according to (8).

$$Z_2(x, y) = \begin{cases} 1, & (x - x_c)^2 + (y - y_c)^2 \leq (2r_c)^2 \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

Finally, the segmentation of the grayscale pixels in the cladding exterior ring is computed according to (9).

$$I_7(x, y) = Z_3(x, y) \cdot I_1(x, y) \quad (9)$$

Figure 10(b) shows an example of  $I_7(x, y)$ .



Figure 10. Example results of the segmentation: (a) Image  $I_6(x, y)$  (segmented cladding), (b) Image  $I_7(x, y)$  (segmented outer ring)

#### 2.4.9. Obtaining descriptors

These steps evaluate the segmented regions  $I_6(x, y)$  and  $I_7(x, y)$  to compute the first two descriptors: cladding and ring variance. Furthermore, they also evaluate image  $I_1(x, y)$  to find the third descriptor: quantity of contours in the image. The first step consists on calculating the histograms of  $I_6(x, y)$  and  $I_7(x, y)$ , without considering pixels with zero intensity (black pixels), such that histograms only contain only gray pixels. Figures 11 and 12 show these histograms.

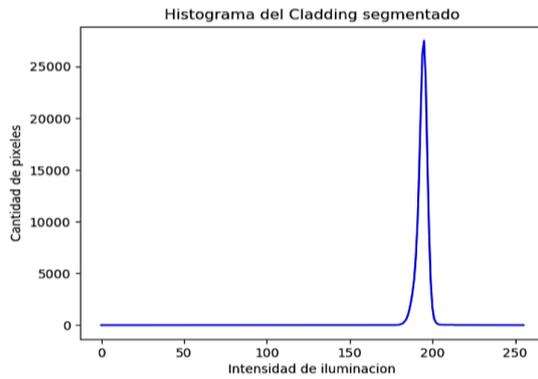


Figure 11. Segmented cladding histogram

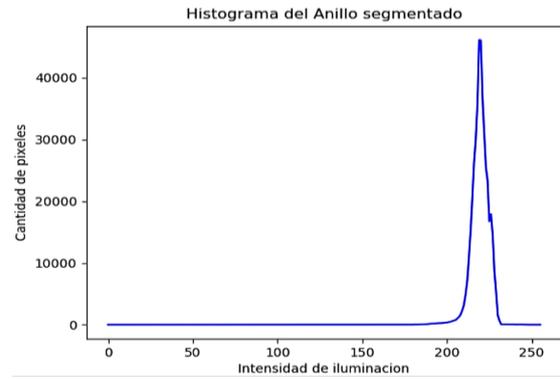


Figure 12. Segmented ring histogram

These histograms serve to calculate the variances [18] in each region (cladding or ring). These variances are defined as  $\sigma_{cladding}^2$  and  $\sigma_{ring}^2$  according to (10).

$$\sigma^2 = \frac{\sum_{i=1}^m (X_i - \bar{X})^2}{m-1} \quad (10)$$

where  $X_i$  represents a pixel value of a region,  $m$  the number of pixels of a region and  $\bar{X}$  the average value of a region (cladding or ring).

The variances of each region are radiometric descriptors that allow to detect the dirt presence level. In this case a region with high level of dirt also will have high variance value. The third descriptor is the number of contours ( $NumCont$ ) identified in the  $I_1(x, y)$  image. This descriptor was chosen because the number of contours increases significantly when the dirt level is high. It helps to differentiate between different levels of dirt or cleanliness.

To obtain this descriptor, the steps 2.4.2, 2.4.3, 2.4.4, 2.4.5 and 2.4.6 are applied to  $I_1(x, y)$ . The difference is in the hysteresis thresholds used in the Canny filtering step:  $\mu_{MAX} = 200$  and  $\mu_{MIN} = 10$ . These hysteresis values attempt to increase the threshold range such that more borders will be found in the whole image, instead of only in the cladding region. This translates into searching for more borders and shapes that could correspond to dirtiness. Figure 13 shows the output image  $I_8(x, y)$ .

The next step consists of finding the quantity of contours in image  $I_8(x, y)$ . Figure 14 shows a flowchart to achieve this objective. Figure 14 shows that Suzuki's algorithm [19] analyzes image  $I_8(x, y)$  to then discard redundant points and obtain a list with the coordinates of all the relevant contours. Then, the integer variable  $NumCont$  is calculated from this list's length.

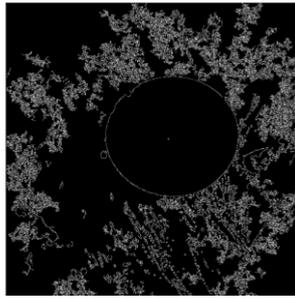


Figure 13. Image  $I_g(x, y)$

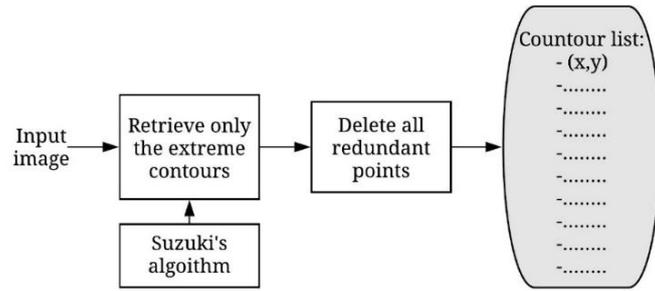


Figure 14. Flowchart for contour quantification

Figure 15 illustrate the basic steps of Suzuki’s algorithm. The algorithm scans the image from left to right, labels edges as outer or hole edges and establishes the hierarchy between the discovered pixels. This repeats for each image row, from top to bottom. The algorithm considers contours as continuous and smooth curves. Finally, the steps achieve the following 3 descriptors:  $\sigma_{Cladding}^2$ ,  $\sigma_{ring}^2$  and  $NumCont$ .

**2.5. Identification of the cleanliness level by neural networks**

A trained model of neural networks was used to identify the cleanliness level of the fiber optic connectors [20]. The model inputs consist on the three previously calculated descriptors, and the output is a 3-element array which indicates the cleanliness of the fiber optic connector. Figure 16 illustrates how the descriptors are input into the neural network and the model outputs the binary array as described above.

[0,0,1]: CLEAN state

[0,1,0]: DIRTY state

[1,0,0]: VERY DIRTY state

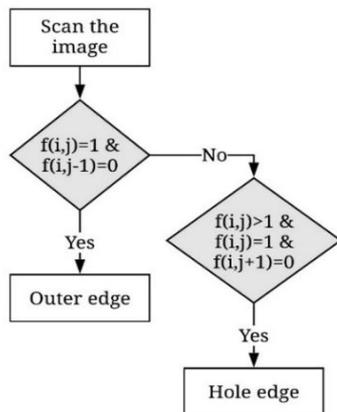


Figure 15. Suzuki’s algorithm basic steps

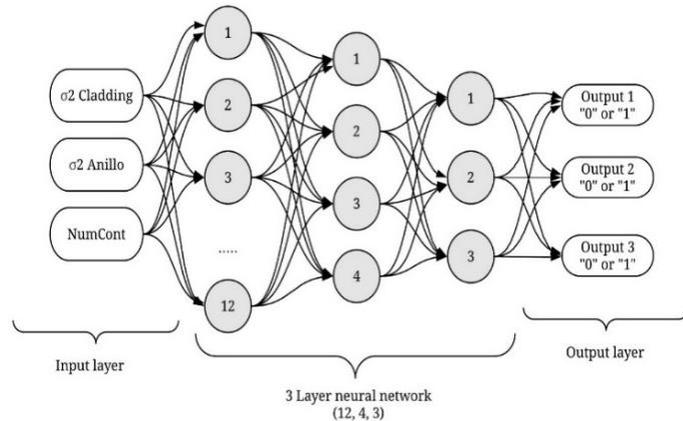


Figure 16. ANN diagram

The dataset consisted of 99 pictures of fiber optic connectors, evenly distributed into 3 groups of 33 photos for each possible output. The dataset was further divided into training and evaluation sets:

- Training set (70% of the total dataset)
- Evaluation set (30% of the total dataset)

The objective was to achieve a precision of at least 90%. To achieve this, the network was set according to the following characteristics:

- Network size: Sánchez [21] mentions that the importance of the hidden layer size to the neural network model fit has been stated many times in the scientific literature, but no conclusive results have been demonstrated. Nonetheless, there are plenty guiding criteria to start building said models. Some of the criteria considered in this work are described by Del Carpio [22]:
  - a) Quantity of hidden layers: more layers add complexity to the network and make the training slower. If data are linearly separable, hidden layers are not necessary, but in this case, data are not linearly separable.

b) Quantity of neurons in each hidden layer: these neurons are related to the input variables, such that this quantity must not be larger than twice the input size. If this is not enough to achieve acceptable results, more neurons are added to the output layer.

Another criterion is directly related to the final adjustment error and desired precision. These criteria guided the model building and simulation, which was done with TensorFlow (a free and open-source software library for dataflow and differentiable programming developed by Google). This simulation helped in defining the final network size: 3 hidden layers with 12, 4 and 3 neurons, which improve the loss during training and testing. Scikit-learn [23] explains each neuron in the hidden layer transforms the values from the previous layer with a weighted linear summation, followed by an activation function. The output layer receives the values from the last hidden layer and transforms them into output values.

- Alpha=0.06, for overfitting correction. This value contributes to keeping the precision above 90% by avoiding overfitting. Scikit-learn use this parameter as a regulator which penalizes overfitting by restricting the weight magnitude.
- ReLU (rectified linear unit) activation function for the hidden layer.
- Training algorithm ‘adam’. Scikit-learn defines ‘adam’ as a stochastic gradient-based optimizer, which can automatically adjust the amount to update parameters based on adaptive estimates of lower-order moments
- Maximum iterations=11000

With these hyperparameters, the model achieved a training precision of 97% and a testing precision of 92%. Figure 17 presents the training and testing algorithm flowchart. The first step is building the network model, defining neurons, layers, and iterations. Next, training takes place, where loss is analyzed at the end: if it is lower than 0.3, the model is tested, otherwise, training resumes. The 0.3 loss ensured that by the end of the training, precision was of at least 90%. Finally, if precision is higher than 91%, the model is saved, otherwise training resumes.

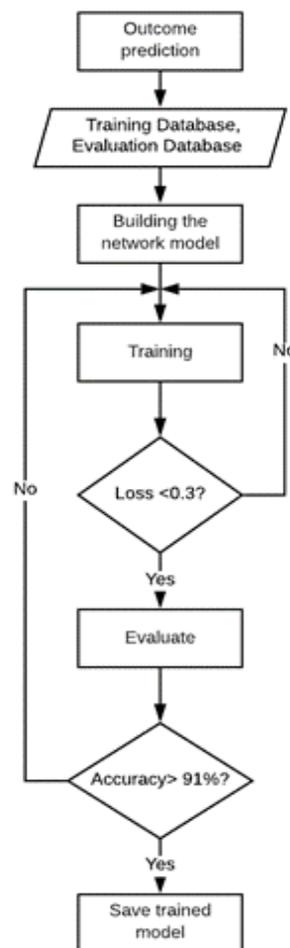


Figure 17. Neural network training flowchart

### 3. RESULTS AND DISCUSSION

First of all, as result, the device reaches the denomination of a low-cost device. Also, the novel proposal device permits a high-quality image acquisition in order to solve the issues registered in previous solutions and proposals [7]. In the other hand, to test the accuracy of the device, two human experts evaluated the cleanliness/dirtiness of 27 SC/UPC fiber optic connectors through survey-type formats. This evaluation was performed by visual inspection of images acquired with optical microscope. However, for validation, only cases where both experts agreed in their evaluation were considered.

Furthermore, the images of the same connectors were evaluated by the proposed algorithm (output of the neural network algorithm). Using the Cohen's Kappa Index [24] which is served to measure the agreement level according between two or more observers according to the values in Table 1 [25, 26], the agreement level according between experts and the device proposed are defined by the magnitude of the index K [27].

Table 1. Kappa Cohen index interpretation

Kappa Cohen Index Interpretation	Kappa index Range
Poor coincidence	$K \leq 0.20$
Fair agreement	$0.20 < K \leq 0.40$
Moderate agreement	$0.40 < K \leq 0.60$
Satisfactory agreement	$0.60 < K \leq 0.80$
Near-perfect agreement	$K > 0.80$

Figure 18 shows the flowchart to compute K. A first agreement evaluation of the 27 connectors resulted in 6 of them where human experts had differing results, due to subjective perception. After eliminating those 6 conflicting elements, the evaluation set was reduced to 21 connectors (i.e., the device was evaluated 21 times), which results are indicated in Table 2. Next, Table 3 shows the confusion matrix between expert evaluation and the proposed algorithm's evaluation. These results are shown as proportions in Table 4. The metric K [24] is computed according to (11).

$$K = \frac{p_o - p_c}{1 - p_c} \quad (11)$$

where:

$p_o$  = proportion of evaluated samples where judges agree

$p_c$  = proportion of evaluated samples expected at random

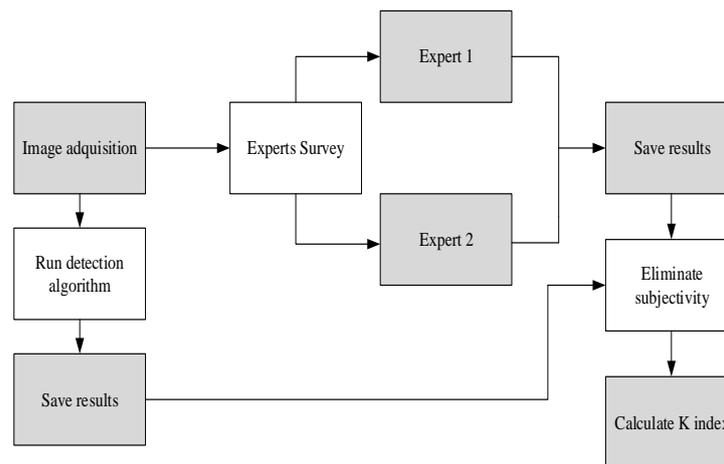


Figure 18. Procedure to compute K

Calculating (11) for the confusion matrix at Table 4 yields:

$$K = \frac{[0.190 + 0.429 + 0.333] - [(0.190 \times 0.238) + (0.429 \times 0.429) + (0.381 \times 0.333)]}{1 - [(0.190 \times 0.238) + (0.429 \times 0.429) + (0.381 \times 0.333)]} = 0.926$$

The agreement results from Table 5 is interpreted according to Table 1, such that:

$$K > 0.8 \rightarrow \text{Near perfect concordance}$$

Next, the index is calculated for each possible result: clean, dirty, and very dirty. Results are shown in Table 6. Note that the proposed method achieves accurate results when compared to human experts.

Table 2. Results to avoid subjectivity

# Image	ID Connector	Expert 1	Expert 2	Experts' coincidence	Prototype Result
1	Connector 95	Dirty	Dirty	Yes	Dirty
2	Connector 97	Dirty	Dirty	Yes	Dirty
3	Connector 63	Very Dirty	Very Dirty	Yes	Very Dirty
4	Connector 2	Clean	Clean	Yes	Clean
5	Connector 5	Clean	Clean	Yes	Clean
6	Connector 69	Very Dirty	Very Dirty	Yes	Very Dirty
7	Connector 62	Very Dirty	Very Dirty	Yes	Very Dirty
8	Connector 7	Clean	Clean	Yes	Clean
9	Connector 65	Very Dirty	Very Dirty	Yes	Very Dirty
10	Connector 1	Clean	Clean	Yes	Clean
11	Connector 67	Very Dirty	Very Dirty	Yes	Very Dirty
12	Connector 70	Very Dirty	Very Dirty	Yes	Very Dirty
13	Connector 96	Dirty	Dirty	Yes	Dirty
14	Connector 68	Very Dirty	Very Dirty	Yes	Very Dirty
15	Connector 100	Dirty	Dirty	Yes	Dirty
16	Connector 66	Very Dirty	Very Dirty	Yes	Very Dirty
17	Connector 98	Dirty	Dirty	Yes	Dirty
18	Connector 4	Dirty	Dirty	Yes	Clean
19	Connector 99	Dirty	Dirty	Yes	Dirty
20	Connector 101	Dirty	Dirty	Yes	Dirty
21	Connector 64	Very Dirty	Very Dirty	Yes	Very Dirty

Table 3. Confusion matrix

		Experts			
		Clean	Very Dirty	Dirty	Total
Prototype results	Clean	4	0	1	5
	Very Dirty	0	9	0	9
	Dirty	0	0	7	7
Total		4	9	8	21

Table 4. Confusion matrix of proportions

		Experts			
		Clean	Very Dirty	Dirty	Total
Prototype results	Clean	0.190	0.000	0.048	0.238
	Very Dirty	0.000	0.429	0.000	0.429
	Dirty	0.000	0.000	0.333	0.333
Total		0.190	0.429	0.381	1.000

Table 5. Index of agreement result

Index	Value
Agreement index(K)	0.926
# Valid cases	21

Table 6. Individual agreement index result

Index		Value
Agreement index(K)	Clean	0.859
	Dirty	0.529
	Very Dirty	1.000
# Valid cases		21

#### 4. CONCLUSION

The conclusions and final remarks are a low-cost device, capable of evaluating the cleanliness of fiber optic connectors with similar results than human experts, has been successfully developed. The microscope adapter achieved stable illumination and safely holds the digital camera. The image processing algorithm computes adequate variables, which reflects on the Cohen's Kappa index values being close to perfect agreement. Nonetheless, when analyzing results individually, the index for dirty connectors is 0.529, which is classified as Moderate Agreement. Although this can be considered an acceptable result, it will be necessary to enhance training such that the agreement value can be higher for this particular result.

The algorithm takes 3 seconds to output the image classification. The digital camera microscope adapter is practical and innovative. The proposed system can be implemented in companies and organizations that already have the optical microscope, further reducing the system cost and increasing the market acceptance. The algorithm might fail when analyzing fiber optic connectors with a broken or damaged ferrule. Since this case is not considered in the algorithm design, the result might not agree with the connector being clean, dirty, or very dirty. As opportunities for improvement, there is the development of a case that can withstand the inclemency of field work and complement the algorithm developed with instances that allow us to distinguish what type of dirt is present and perhaps carry out an automatic cleaning, using electronic tools.

## ACKNOWLEDGEMENTS

The authors would like to thanks “Universidad Peruana de Ciencias Aplicadas” for technical and economic support.

## REFERENCES

- [1] Y. Roriguez, “Fibra Óptica,” Santa Fe: El Cid Editor, 2009.
- [2] E. H. Perez, “Introducción a las Telecomunicaciones Modernas,” Mexico DF: Limusa, 2004.
- [3] Chougule, Sharada V., and Mahesh S. Chavan, “Investigation of Connector Performance and Damage to Optical Fiber Connectors with Dust During High-Power Transmission,” in *Journal of Lightwave Technology*, vol. 36, no. 17, pp. 3586–3593, Sep. 2018.
- [4] J. Simatupang, “Theoretical Analysis of Backreflections in Bidirectional Wavelength Division Multiplexing-Passive Optical Networks,” *Bulletin of Electrical Engineering and Informatics (BEEI)*, vol. 2, no. 1, pp. 45–52, Mar. 2013.
- [5] M. Brown, “Achieving IEC Standards Compliance for Fiber Optic Connector Quality,” *VIAVI*, [Online]. Available: <https://www.viavisolutions.com/en-us/literature/achieving-iec-standards-compliance-fiber-optic-connector-quality-white-paper-en.pdf>
- [6] J. Duffy, D. Hashimoto, J. Sloan, and G. Bellegarde, “Development of an Automated Cleaning System for Multi-ferrule Fiber Optic Connectors,” in *Conference on Optical Fiber Communication/National Fiber Optic Engineers Conference*, San Diego, CA, 2008, pp. 1–5, doi: 10.1109/OFC.2008.4528037.
- [7] A. Rehman and W. B. Mozaffar, “Optical Fiber Connector Surface Defect Detection Using Wavelets,” in *2007 International Conference on Information and Emerging Technologies*, Karachi, 2007, pp. 1–6.
- [8] A. I. Filipenko, “Fiber connector end face analysis by interference method,” in *Proceedings of LFNM'2000. 2nd International Workshop on Laser and Fiber-Optical Networks Modeling (Cat. No.00EX419)*, Kharkiv, Ukraine, 2000, pp. 74–76, doi: 10.1109/LFNM.2000.854045.
- [9] “Fiber Inspection & Cleaning,” *VIAVI*, 2020. [Online]. Available: <https://www.viavisolutions.com/en-us/product-category/enterprise-networks/test-certification/fiber-inspection-cleaning>
- [10] “Fiber Inspection,” *EXFO*, 2020. [Online]. Available: <https://www.exfo.com/es/productos/pruebas-de-redes-de-campo/inspeccion-de-fibra/>
- [11] UIT-R, “Parámetros de codificación de televisión digital para estudios con formatos de imagen normal 4:3 y de pantalla ancha 16:9,” Recomendación UIT-R BT.601-7, 2011. [Online]. Available: [https://www.itu.int/dms\\_pubrec/itu-r/rec/bt/R-REC-BT.601-7-201103-1!!PDF-S.pdf](https://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.601-7-201103-1!!PDF-S.pdf)
- [12] C. Platero, “Apuntes de Visión Artificial,” Universidad Politécnica de Madrid, 2007, Art. no. 120.
- [13] J. García, “Estudio de Filtros Óptimos para la eliminación de ruido en imágenes digitales usando algoritmos de multirresolución no lineales,” Repositorio Digital de la Universidad Politécnica de Cartagena, pp. 3, 2009.
- [14] R. L. Cañar and R. Escandón, “Implementation of a prototype for capture and digital processing of thermal images acquired from a UAV,” *Enfoque UTE*, vol. 9, no. 1, pp. 1–11, Mar. 2018.
- [15] E. A. Sekehravani, E. Babulak, and M. Masoodi, “Implementing canny Edge detection algorithm for noisy image,” in *Bulletin of Electrical Engineering and Informatics*, vol. 9, no. 4, pp. 1404–1410, Aug. 2020.
- [16] H. Ohmaid *et al.*, “Iris segmentation using a new unsupervised neural approach” in *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 9, no. 1, pp. 58–64, Mar. 2020.
- [17] G. Bradski and A. Kaehler, “Learning OpenCV,” 2008. [Online]. Available: <https://www.bogotobogo.com/cplusplus/files/OReilly%20Learning%20OpenCV.pdf>
- [18] C. Salvatierra, “Fundamentos de procesamiento digital de imágenes de satélite,” 2004. [Online]. Available: [http://ffyl1.uncu.edu.ar/IMG/pdf/Notas\\_clase\\_Procesamiento\\_digital.pdf](http://ffyl1.uncu.edu.ar/IMG/pdf/Notas_clase_Procesamiento_digital.pdf)
- [19] S. Suzuki and K. Abe, “Topological Structural Analysis of Digitized Binary Images by Bordes Following,” *CVGIP*, vol. 30, no. 1, pp. 32–46, 1985.
- [20] S. Haykin, “Neural Networks: A Comprehensive Foundation,” *2nd edition USA*, Prentice Hall, 1999.
- [21] P. Sánchez and J. Velásquez, “El rol del algoritmo de entrenamiento en la selección de modelos de redes neuronales,” *Revista U.D.C.A Actualidad & Divulgación Científica*, vol. 14, no. 1, pp. 149–156, 2011.
- [22] C. Del Carpio, “Método de reconocimiento de patrones basado en procesamiento digital de imágenes y redes neuronales orientado al registro automático de evaluaciones académicas,” M.S. thesis, Universidad Nacional de Ingeniería, Facultad de Ingeniería Eléctrica y Electrónica, Lima, Perú, 2015.

- [23] Scikit-Learn, "Varying regularization in Multi-layer Perceptron," 2019. [Online]. Available: [https://scikit-learn.org/stable/auto\\_examples/neural\\_networks/plot\\_mlp\\_alpha.html](https://scikit-learn.org/stable/auto_examples/neural_networks/plot_mlp_alpha.html)
- [24] J. Cohen, "A Coefficient of Agreement for Nominal Scales," *Educational and Psychological Measurement*, vol. 20, no. 1, pp. 37–46, Apr. 1960.
- [25] S. M. Vieira, U. Kaymak, and J. M. C. Sousa, "Cohen's kappa coefficient as a performance measure for feature selection," in *International Conference on Fuzzy Systems*, Barcelona, 2010, pp. 1–8.
- [26] J. Lazar, J. Feng and H. Hochheiser, "Research Methods in Human-Computer Interaction," *2nd Edition, Morgan Kaufmann*, 2017.
- [27] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," *Biometrics*, vol. 33, no. 1, pp. 159–174, Mar. 1977.

## BIOGRAPHIES OF AUTHORS



**Victor Fernandez**, Electronic Engineer of the Peruvian University of Applied Sciences (UPC), he obtained a Diploma Project Management and Agile Methodologies from Telefonica Foundation. He has been part of telecommunications project for more than 4 years. He currently works on development of new technologies and IT applications.



**Javier Gustavo Chavez**, Electronic Engineer of the Peruvian University of Applied Sciences (UPC), he obtained a Diploma on Telecommunication Management in Graduate School of the Peruvian University of Applied Sciences (EPG-UPC). Optical DWDM Network Specialist. He has been part of optical fiber projects for 7 years. He currently works on optimizing and supporting the National Optical Fiber Network of Perú.



**Guillermo Kemper** received the B.S. degree in Electrical Engineering in 1994 from the Antenor Orrego Private University (UPAO), Trujillo, Peru. In years 1996 and 2001 he obtained a Master and a PhD degree in Electrical and Communications Engineering respectively. Both degrees from the State University of Campinas (UNICAMP), Brazil. He has taken part for more than 3 years in the research agreement between the CPqD-Telebras (Brazilian Telecommunications Company) and UNICAMP, developing and designing audio and video coders. Currently, he is working as a research professor at the Electrical Engineering Program of the Peruvian University of Applied Sciences (UPC) in Lima, Peru. In addition, he is a researcher at the National Institute of Research and Training in Telecommunications (INICTEL). His research interests include voice, audio, images and video processing, as well as digital communications and digital television.