Image Classification Applied to the Detection of Leather Defects for Smart Manufacturing

Alberto Ochoa-Zezatti[1], Oliverio Cruz-Mejía[2](✉), Jose Mejia[1], and Hazael Ceron-Monroy[3]

[1] Universidad Autónoma de Ciudad Juárez, Ciudad Juárez, México
{alberto.ochoa,jose.mejia}@uacj.mx
[2] Universidad Autónoma del Estado de México, Toluca, México
ocruzm@uaemex.mx
[3] Universidad Anáhuac México, México City, México
hazael.ceron@anahuac.mx

Abstract. In the shoe production workshops, animal leather is used as the main raw material. Generally, an operator manually checks the surface of the leather, making sure that it does not present defects that compromise the quality of the final product.

This type of inspection is subject to human error and uncontrollable factors, which represents an opportunity for the automation of the process through a system of artificial vision.

A data set was developed consisting of images of animal leather, in good coordination and with defects.

The digitized samples were subjected to image processing using OpenCV and Scikit-Learn, and then used in a convolutional neural network interfacing, using TensorFlow's Keras library in Python.

Finally, the trained model is capable of classifying new images into two possible groups: "Defective Leather" and "Defect-free Leather".

The trained model offers 80% predictive accuracy and 85% reliability. Although the result can be considered satisfactory, it is expected to raise the mentioned percentage with a more robust data set than the one used for the project.

Keywords: Image classification · Artificial vision · Convolutional neural network · Smart manufacturing · Footwear industry · Keras · Tensorflow

1 Introduction

The Mexican footwear production industry has faced different challenges in recent years, derived from the competitiveness that has resulted from the opening of borders and the introduction of manufactured products in countries with the necessary resources to produce greater volumes with reduced production costs [1]. Therefore the need to develop and implement technologies focused on the optimization and automation of processes arises, which pose a better competitive scenario for Mexican companies.

J. A. Marmolejo-Saucedo et al. (Eds.): COMPSE 2020, LNICST 359, pp. 52–61, 2021.
https://doi.org/10.1007/978-3-030-69839-3_4

The relevance of this industry in Mexico is undeniable, since it represents 1.7% of the total manufacturing industries, contributing 2.4% of the total employment in this area. Likewise, the municipality of Leon, Guanajuato contributes with 57.8% of the total production value, being the municipality with the greatest participation. This makes it an ideal candidate for the application of technology-based projects focused on process optimization [2].

It is equally important to define the global situation of Mexico in footwear production, since in 2019 it reached the ninth place in the list of main producers, with a contribution of 268 million pairs during this year. The list is headed by China, with a contribution of 13,478 million pairs in the same period of time, according to the portal Statista [3].

In a shoe production workshop, animal leather is the main raw material. The quality of the final product is directly related to the quality of the leather used. This is susceptible to present diverse defects in its surface, among which we can find fissures, wrinles, scars and holes.

These defects will cause important quality failures in the final product, so it will be necessary to reprocess the product or directly discard it.

Defects are easily detected with proper lighting and training. Inspection of the leather surface is commonly performed by an operator, who checks and validates the condition of the material. Operator involvement adds complications related to low productivity, fatigue and subjectivity to the procedure [4].

It is possible to automate the inspection process by using an artificial vision system [5] that consists of processing and classifying images of animal leather, determining whether they have surface defects. For this purpose, it is necessary to develop an image classification system based on neuronal networks, which have been used before for the detection of defects in several areas, such as in the 3D impregnation process [6], in the automotive sector [7], in the agro-food industry [8] and, like this work, in the footwear industry [9].

Throughout this paper we present the process carried out for an image processing and classification system whose objective is to detect superficial defects in the leather used in the production of footwear and which can be easily integrated into an automated artificial vision system that informs by means of an alarm when anomalies are found.

1.1 Objective

Train a convolutional neural network which is capable of classifying an image (photoraph) of the leather surface into two possible categories: “with defects” or “without defects.

2 Fundamentals

Image classification based on neural networks has proven to be a powerful tool in the area of quality control, allowing an artificial vision system to be able to distinguish defective pieces, automating the inspection process with positive results in the area of footwear, specifically in animal leather [9, 10].

For an image classification system, it is convenient to train a convolutional neural network. These are a type of neuronal network that works with two-dimensional matrices, so it is frequently applied in the area of artificial vision [11]. In Fig. 1, it is illustrated the operation of such a network.

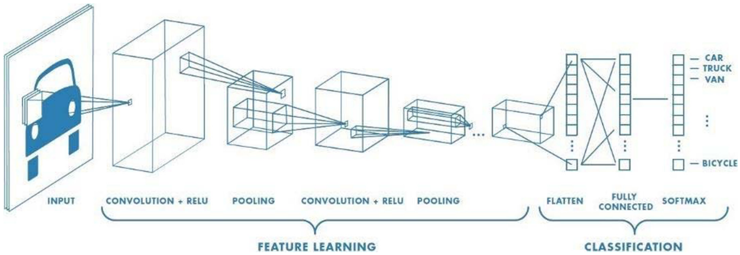

Fig. 1. Illustrated convolutional neural network. Example of vehicle image classification. Source: Towards Data Science.

The construction of the network will be done using the Keras library, which belongs to Tensorflow. This library facilitates the user's process of training deep learning models, as it contains a variety of previously validated "Deep learning frameworks" [12]. One of the advantages of this library is its ability to build robust models from relatively small datasets (less than 1000 samples per class).

Some useful features of Keras are the following:

- **Fit_generator**. It performs neural network training. It depends on the next parameters: epoch, steps per epoch, batches and batch size. Although they are modifiable, the most relevant are not those mentioned, but the learning rate and the error and validation training, which define the quality of the model.
- **ImageDataGenerator**. This module generates more images using the original input images. To achieve this, it performs zooming, re-cutting, brightness alteration and other modifications in order to expand our original dataset and provide robustness to the model. All image transformation paramters are controllable, allowing to set a ratio in which the module will randomly modify the image. For example, you can configure the rotation range within which the image will be randomly rotated (see Fig. 2).

Fig. 2. This is what our data augmentation strategy looks like. Source: The Keras Blog.

Finally, you must set the number of classes in which you want to classify the image in keras. Each class must have three sets of images: training, validation and test. It is important that these 3 sets are mutually exclusive. In the training set are the data that will train the model. The validation set is used to prevent overfitting. The test set is used only to verify the validity of the model.

In the industrial field, the image classification is intended to be incorporated into a complete machine vision system, which contains the following components:

a) Object
b) Image sensor (camera)
c) Lighting
d) Image processing module
e) Decision module (software)

The list of components of such a system may vary depending on the source consulted. This is an adaptation without omission of other proposed systems [13].

Figure 3 successfully illustrates an example of a machine vision system.

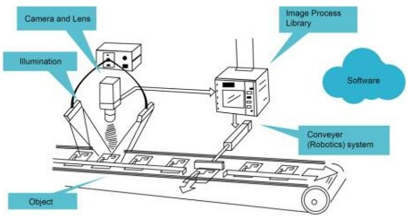

Fig. 3. Diagram of a machine vision system. Source: Innomiles International.

Image processing consists of 3 stages: smoothed, edge detection, enhancement.

Smoothed. At this stage the "Median Blur" filter from the OpenCV library is used. This filter is classified as a smoothed filter and its purpose is to reduce the noise in the images, which is useful to facilitate the conservation and edge detection [14]. The filter acts on the entire image and replaces the value of each pixel with the average value of the surrounding pixels within a defined radius. The only parameter that can be manipulated is the radius of the filter.

Edge Detection. The combination of the Sobel and Laplace operators, both of which are designed for edge detection, has been used before and has proven to be efficient and compatible with each other [15], allowing sharp changes in the original image to be successfully highlighted. Using a combination of two different edge detection filters is an unpopular technique, although very detailed results can be achieved.

One of the recognized edge detection operators is the so-called "Sobel Filter", which produces an image that emphasizes the edges of the original. The use of this filter has become widespread in image processing and machine vision [16].

The Laplace Operator is the second operator in the combination.

Enhancement. The last stage of digital image processing is an image enhancement using the Pillow library, available in Python. The "Contour" filter is used, which gives us a negative of the input image and improves the contours of the input image.

This image processing is sufficient to be used in the training of the neural network, since the result is a binary image that highlights only the relevant contours (edges of the piece and defects within it).

3 Dataset

In order to obtain a sample of significant size for later analysis, a footwear production workshop located in the city of León, Guanajuato, was contacted.

The workshop was asked for a wide variety of leather samples, composed of elements in good condition and elements with defects.

The sample consists of 92 different sheets of leather, among which the defects are predominant, which will be described below.

3.1 Defects

It is important to define what is considered a defect and what is not, in addition to the possible repercussions of these. Figure 4 shows two examples of leather in good condition, suitable for use in the production of footwear.

Fig. 4. Samples of leather without defects.

This leather has the characteristic of not containing wrinkles, folds, scars or holes on its surface. The absence of these defects allows its proper handling in the production process of footwear.

In Fig. 5, we find examples of the different defects present in the sample. In this case, no more types of defects were present, which presents an opportunity to expand the sample in order to obtain a more robust model.

It is important to remember that the presence of these defects may or may not be harmful depending on the type of product being manufactured. For the footwear industry in particular, the absence of these defects over a wide area is of great importance, since the leather will be subjected to processes of tension that the material will probably not

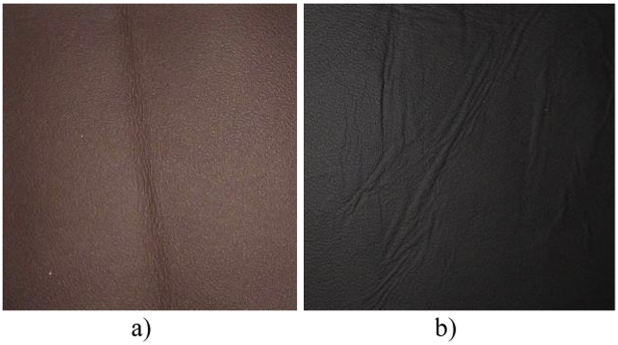

Fig. 5. Common defects in leather: a) Scars b) Creases.

be able to withstand if any of the defects are present. It may happen that the matte leather will resist the whole procedure, although this is even worse, since the final product will have a poor quality and a life span notably less than the expected.

3.2 Obtaining Samples

As mentioned, the leather samples were provided by a shoe workshop located in the city of Leon, Guanajuato. The digitalization process consisted in taking pictures using a Samsung S5KGM1 sensor with 48-megapixel resolution.

The capture of photographs was done in a clean space, so that the sensor captured only the leather sheet. An LED lamp (or any other direct light source) is needed to continuously illuminate the leather samples, so that the picture is clear and easy to process.

4 Procedure

Once the dataset is built, the next step is the digital processing of images.

For this, the OpenCV and scikit-learn libraries, both available in Python, were used. In comparison to other researches, a combination of different filters was used, so that they could highlight the information related to the defects and at the same time reduce the information that is not useful for the analysis of the images.

With the processed images the next step is to build a convolutional neural network.

For this purpose, the Keras library, which belongs to Tensorflow.

4.1 Image Processing

Figure 6 shows an image processing results, allowing to buy between the original and the processed image, as well as the sample of leather with defects and without defects.

The Median Blur filter, available in the OpenCV library, was used. The only modifiable parameter is the radius of pixels used for averaging. The larger this radius is, the

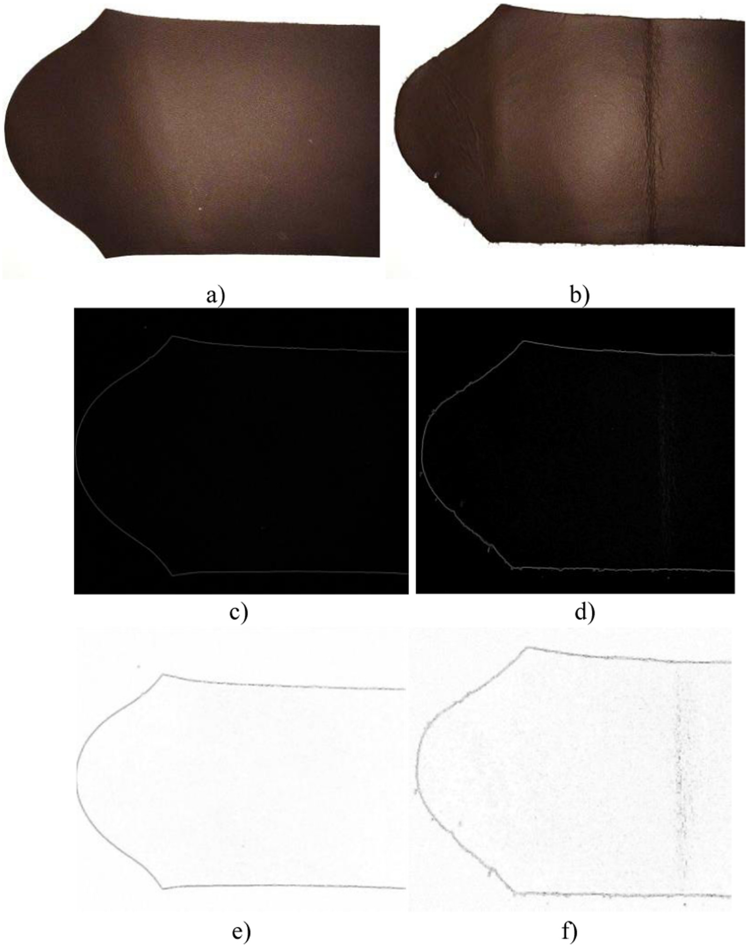

Fig. 6. Leather without defects: a) Original image, c) Edge Detection, e) Contour enhancement. Leather with defects: b) Original image, d) Edge Detection, f) Contour enhancement.

less sharp the image is. This filter allows us to get rid of noise and irrelevant information present in the image.

Later, we used the combination of operators "Sobel Filter" and "Laplace Operator", belonging to the Scikit-learn and OpenCV libraries, respectively. The benefits of this combination were described previously and the result can be seen in Fig. 6(c and d).

Finally, we applied the "Contour" filter from the Pillow library, which gives us a negative of the image and an enhancement in the contours, as shown in Fig. 6(e and f).

4.2 Neural Network Training

To develop the image classification system, a convolutional neural network was trained using TensorFlow's "Keras" library, whose operation and validity were developed in the "Fundamentals" section.

The network was configured so that the training consists of 20 epochs, with a learning rate of 0.0004 and a minimum accuracy of 75%. Raising any of these parameters (reducing in the case of the learning rate) offers a better-quality training, although this requires a computer with better performance than that available (especially on the GPU) and long waiting times, so it was decided to keep this.

5 Results

The Keras library provides us with information on the result of the training, which was 80% accurate. To measure the validity and check the reliability of the model built, the data set called test set is used, which consists of 20 images, 10 of them of leather in good condition and 10 of leather with defects, in order to check if the model is able to classify them correctly. It is important to clarify that the test images cannot be the same ones used to train the network. The results of this test are focused on Table 1.

Table 1. Results of model predictions.

Image description	Test samples	Correct predictions
Flawless leather	10	7
Flawed leather	10	10
Total	20	17

The model works well in detecting errors, as it successfully sorts out images showing defective leather parts. However, only 7 out of 10 images made an accurate prediction. This means that the system would potentially reject 30% of the samples that are in good condition. However, it would be able to identify an error 100% of the time. Overall, the system seems to make predictions with an 85% effectiveness.

6 Future Work

At least 3 next steps are easily identifiable after the completion of this paper.

Improvement. The first step is to improve the model. To do this, it is strictly necessary to extend the data set used, to perform a strict sampling of at least 2000 samples and with a greater variety of defects. In addition, the image processing can also be optimized compared to that used in other projects of a similar nature.

Artificial Vision System. The second step consists of adding the code to a complete artificial vision system. Since it is written in Python and can be quickly co-written in the cloud, a Raspberry Pi would be ideal to reduce implementation costs. It would also require the placement of a conveyor belt, direct lighting and RGB LED lights or speaker for an alarm to notify that an error was detected.

Implementation. The last step is the implementation of the system in a shoe production workshop, provided that the model has proved useful. A demonstration would be made to the managers to allow them to verify that the system works with reduced margins of error, so that they feel confident to automate the inspection process.

7 Conclusion

Image processing and neural network training using Keras and Tensorlow with a small data set is a potentially useful tool in the area of quality control and machine vision.

Although the trained model presents weaknesses in the identification of parts without defects, the percentage is high enough to consider it a useful option, highlighting that it is obviously subject to improvement and making it clear that as further work the data set to be processed must be greatly expanded.

The processing of the images is also subject to improvement, different combinations of filters could offer better results.

The image processing, the convolutional neural network training and the prediction system were executed in the cloud using the Google Colab platform, which supports Python. Also, Python is the most widely used language in the Raspberry Pi ordinances, frequently used in the Internet of Things. These facts give a better perspective to systems like this, since the inclusion in a production line using Raspberry Pi and wireless network modules, as well as the ability to work in the cloud, imply a large area of implementation opportunity at low cost, provided that it is done with the necessary knowledge.

References

1. Pacheco Vega, R.: Historia de dos ciudades: un análisis comparativo de los distritos industriales del cuero y calzado en León y Guadalajara. In: Meeting of the Latin American Studies Association (2004)
2. INEGI: Estadísticas a propósito de la Industria del calzado, México (2014)
3. Footwear - worldwidel Statista Market Forecast. https://www.statista.com/out-look/11000000/100/footwear/worldwide
4. Perng, D., Liu, H., Chang, C.: Automated SMD LED inspection using machine vision. Int. J. Adv. Manuf. Technol. **57**, 1065–1077 (2011)
5. Aguilar-Torres, M., Argüelles-Cruz, A., Yánez-Márquez, C.: A real time artificial vision implementation for quality inspection of industrial products. In: 2008 Electronics, Robotics and Automotive Mechanics Conference, pp. 277–282 (2008)
6. Wu, M., Phoha, V.V., Moon, Y.B., Belman, A.K.: Detecting malicious defects in 3D Printing Process Using Machine Learning And Image Classification. In: ASME 2016 International Mechanical Engineering Congress and Exposition (2016)

7. Armesto, L., Tornero, J., Herraez, A., Asensio, J.: Inspection system based on artificial vision for paint defects detection on cars bodies. In: 2011 IEEE International Conference on Robotics and Automation, pp. 1–4. IEEE (2011)
8. Gnanavel, S., Manohar, S., Sridhar, K., Sokkanarayanan, S., Sathiyanarayanan, M.: Quality detection of fresh fruits and vegetables to improve horticulture and agro-industries. In: 2019 International Conference on contemporary Computing and Informatics (IC3I), pp. 268–272 (2019)
9. Bong, H., Truong, Q., Nguyen, H., Nguyen, M.: Vision-based inspection system for leather surface defect detection and classification. In: 2018 5th NAFOSTED Conference on Information and Computer Science (NICS), pp. 300–304 (2018)
10. Liong, S.T., et al.: Efficient neural network approaches for leather defect classification. arXiv preprint. (2019)
11. Simard, P., Steinkraus, D., Platt, J.: Best practices for convolutional neural networks applied to visual document analysis. In: Seventh International Conference on Document Analysis and Recognition, vol. 3 (2003)
12. Nandy, A., Biswas, M.: Reinforcement learning with keras, tensorflow, and chainerrl. In: Reinforcement Learning, pp. 129–153 (2017)
13. Liu, B., Wu, S., Zou, S.: Automatic detection technology of surface defects on plastic products based on machine vision. In: 2010 International Conference on Mechanic Automation and Control Engineering, pp. 2213–2216 (2010)
14. Perreault, S., Hébert, P.: Median filtering in constant time. IEEE Trans. Image Process. **16**(9), 2389–2394 (2007)
15. Gupta, S., Porwal, R.: Combining laplacian and sobel gradient for greater sharpening. ICTACT J. Image Video Process. **06**, 1239–1243 (2016)
16. Jin-Yu, Z., Yan, C., Xian-Xiang, H.: Edge detection of images based on improved Sobel operator and genetic algorithms. In: 2009 International Conference on Image Analysis and Signal Processing, pp. 31–35 (2009)
17. Simard, P., Steinkraus, D., Platt, J.: Best practices for convolutional neural networks applied to visual document analysis. In: Seventh International Conference on Document Analysis and Recognition, 2003, vol. 3 (2003)