ISSN 2395-8618

Use of Machine Learning Implementing Convolutional Neural Networks to Classify Export Markets for Persian Lime

Miguel Ángel Torres Medel, Carlos J. Rivero, Roberto Ángel Meléndez-Armenta, and Irahan Otoniel José Guzmán

***Abstract*—Mexico is the world's second largest producer of Persian lime; however, most producers of this citrus fruit cannot export on their own because they do not know the standards required to do so. As a result of this situation, they have to resort to intermediaries to carry out the fruit selection process, which causes them to lose a certain percentage of their profits. Therefore, the objective of this research is to generate a Machine Learning model with the implementation of convolutional neural networks to identify the different characteristics of the Persian lime and thus know to which market it should be exported. In this sense, the results obtained show that the model achieved an accuracy of 97% in its classification of the types of export markets established.**

***Index Terms*—Intelligent system, machine learning, Persian lime.**

I. Introduction

Mexico is among one of the first countries with more lemon production worldwide being the second place with 2.6 million tons each year, since it is distinguished by the good quality and diversity in terms of types, such as the Persian or seedless lemon (Citrus latifolia), another type of lemon is the Mexican, green or bitter (Citrus aurantifolia) and finally the yellow or Italian lemon (Citrus lemon), to mention a few, the type of lemon which will be focused the research work will be the Persian lemon [1].

Due to the fact that the quality of the mentioned fruit is high, the processes through which it passes are quite demanding, and they are usually in an empirical and traditional way, the mentioned process analyses the quality of the lemon since it can vary in sizes (measure 110 to measure 250), texture and colour, since when doing this analysis the lemons are going to be discriminated and to know to which market they can be exported.

Lemon producers, when making their respective cuts, have to go through an intermediary so that the fruit can be exported in an efficient way, based on the quality standards of the lemon. Because of this, the farmer loses a large percentage of profit, as he cannot make the correct selection to detect and decide which market is the most convenient to export the lemon to [2].

To address the above problem, it is proposed the development of a mobile application that performs the identification of the characteristics such as color, texture and size. In order to know to which market it is more convenient to export the citrus fruit and in such a way that it is a useful tool for the Persian lemon market, helping to reduce time and resources.

II. State of Art

Paper [3] addresses the issue of classification of some lemon species by implementing deep learning techniques. For the study they used approximately 2000 computer images belonging to 3 lemon species and implemented convolutional neural network algorithms; the results of the application of such technique could make the final model reach an accuracy of 99.48%, which shows that the model is feasible.

In the work [4], it is addressed the issue of detecting defects in sour lemons by applying artificial intelligence techniques. For their study, they used a dataset containing computer images of sour lemons, where they were divided into two groups that were healthy and damaged and implemented a Convolutional Neural Network and data augmentation, these were compared with other methods such as Local Binary Pattern, HOG, K-nearest neighbor, an Artificial Neural Network and a Support Vector Machine. The accuracy of the CNN was 100% for the detection of imperfections in citrus lemons.

The work [5] addresses the issue of detecting damage caused by the citrus leafminer in Mexican lemons; for the study they used digital images, which were taken with various digital cameras and in JPG/JPEG format, as this type of image gave better results. A KNN classifier was used to classify the images, which provided two classes, one for healthy leaves and the other for those with leafminer damage. The classifier gave good results when classifying, but it was recommended to take the photos in a controlled space, as light can affect the process.

In the paper [6], it is presented the issue of detection of physiological disorders of granulation and endoxerosis in oranges and lemons; for the study, computerized photographs of oranges and lemons with high degrees of expected defects were used and radiographs were simulated to take into account random fruit orientations. In order to make the detection of the mentioned images, naïve Bayes and kNN were used. With these methods, labor-intensive destructive sampling is avoided and this allows the non-destructive inspection of all fruits; the classifier had good results, for oranges it classifies 95. 7% efficiency and 93.6% for lemons.

The issue of disease detection in citrus fruits is discussed in [7]. For the research they used computerized images of citrus fruits; to enhance the detection they used an artificial

Manuscript received 18/01/2021, accepted for publication on 13/06/2021.

The autors are with the División de Posgrado e Investigación, Tecnológico Nacional de México, Instituto Tecnológico Superior de Misantla, Mexico (212t0025, 212t0027, ramelendeza, iojoseg}@itsm.edu.mx).

ISSN 2395-8618

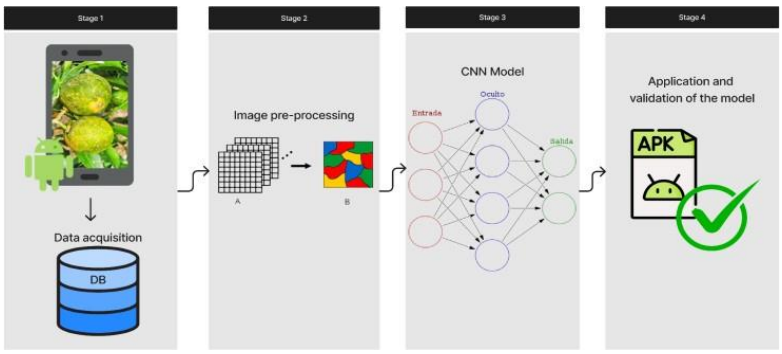

Fig. 1. Proposed methodology (Source: Own elaboration)

TABLE I
CLASSIFICATION OF LEMONS (SOURCE: OWN ELABORATION)

Classification	Train	Test	Valid	Size and Format
International persian lemon	x25	x15	x10	4000x 3000 PNG
National persian lemon	x25	x15	x10	4000x 3000 PNG
Persian lemon waste	x25	x15	x10	4000x 3000 PNG

intelligence technique called deep neural network (DNN), this network detects specific areas of the disease with its level of severity, it classifies 4 levels of severity, which are high, medium, low and healthy; the model predicts the low level of severity with 99% accuracy, the high level with 98%, the medium level with 97% accuracy and the healthy level with 96%.

The issue of cherry fruit packing methods in order to reduce waste and increase exportability and marketability is discussed in [8]. Computerized cherry images were used; computer vision techniques were applied, in this case a Convolutional Neural Network was used to detect the appearance of the cherries and provide an efficient system for classification; the Convolutional Neural Network performed well, showing an accuracy of 99.4%.

In [9], the authors applied eight techniques to classify Mexican lemons by their appearance; they used 913 computer images of lemons, which were divided into two classes: defective and healthy, in this case 8 deep learning models were used to make a comparison of which one performed better; the model that yielded the best result was 92% accuracy.

III. MATERIALS AND METHODS

The study objective of the present research work is to implement a convolutional neural network [10] to be able to assign based on the size, texture and shades of green, to the target market, based on the quality standards that each market presents. A pre-trained network of Keras was used, which was the VGG16 architecture which was selected for its pre-entrainment, which has been the winner of the ImageNet competition [8], and which adapts to the objectives and problems of the present work.

A. Dataset

The total of the multi-class sets were acquired from 150 images taken in a controlled space of lemon orchards specialized in citrus export, which will be our study in order to

ISSN 2395-8618

Fig. 2. VGG16 Architecture

TABLE II
VGG16 PRE-ENTRANT NETWORK ARCHITECTURE

Layer		Feature maps	Size
Input	Image	1	224x224x3
1	2xConv	64	224x224x64
	Agr. Max	64	112x112x64
3	2xConv	128	112x112x128
	Agr. Max	128	56x56x128
5	2xConv	256	56x56x256
	Agr. Max	256	28x28x256
7	3xConv	512	28x28x512
	Agr. Max	512	14x14x512
10	3xConv	512	14x14x512
	Agr. Max	512	7x7x512
13	FC	-	25088
14	FC	-	4096
15	FC	-	4096
Output	FC	-	1000

determine their size, color, etc. The preprocessing consisted of removing the background of these lemons so as not to interfere with the classification of the same, also the noise was omitted and background that had the captured images. Then the images were divided into 50 images to classify export lemons, which is equivalent to 33% of the total, within the percentage were divided into 25 images for training (training), 15 for test (test) and finally 10 to validate (validate) that the model works properly (See Table 1). For domestic consumption, with a percentage of 33% was used to classify correctly within the model, finally for the class "waste", are the lemons which are already in the "waste" class, which are the lemons which are already in the "waste" class. Finally, for the "waste" class, it is lemons, which are no longer of any relevance in the export market.

The images have characteristics such as texture, color, size, which we tried to adhere to the quality standards, so the images were taken from our own elaboration, visiting a company dedicated to citrus fruits such as CITREX located in Martínez de la Torre, in the state of Veracruz, which was discussed with the supervisor in order to extract the images used in this research.

B. CNN Model

The paper entitled Very Deep Convolutional Networks for Large-Scale Image Recognition [10] proposed a Convolutional Neural Network model, which achieved a test accuracy of 92.7% and ranked among the top five in ImageNet [11], which is a dataset of approximately 14 million images that are classified into 1000 different classes. This model was one of the most famous models submitted to ILSVRC-2014 [12] which shows an improvement over AlexNet by replacing large kernel-size filters (11 and 5 in the first and second convolutional layers) with multiple 3x3 kernel-size filters one after the other. The VGG16 model occupied an NVIDIA Titan Black GPU.

The VGG16 architecture is shown (Figure 2) in which the input to the first convolution layer is a 224 x 224 image. This figure passes through a set of convolutional (conv) layers, where the filters were occupied with a receptive field of a small size: 3x3 which allows capturing the notion of up/down, left/right and center.

IV. RESULTS

A. Convolutional Neural Network Architecture

In this research work, we have studied the pre-integrated model of Keras called VGG16, this architecture aims to have an optimal classification to be able to assign labels to the captured images, in National, International and Waste Consumption, with the help of a Machine Learning technique such as Convolutional Neural Networks (CNN).

Comparing an automatic system capable of analyzing an image of different lemons, with empirical knowledge, which can be wrong, this system can have an answer in a matter of seconds, about which market can be assigned a set of lemons.

ISSN 2395-8618

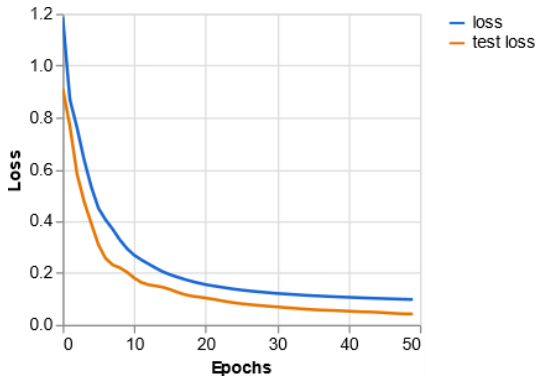

Fig 3. Epochs

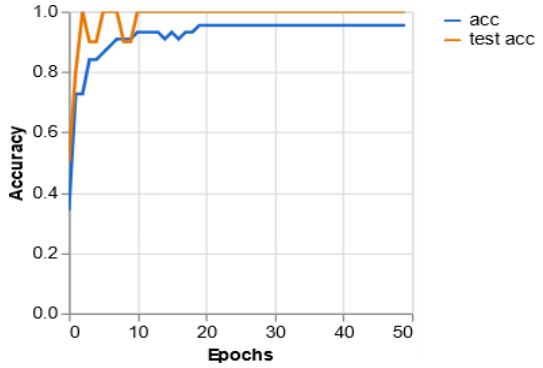

Fig 4. Accuracy

Fig. 5. Classification accuracy

Since, being programmed with a Convolutional Neural Network, gives us the advantage that these layers have an order in which the Persian lime image is processed, the most used layers are:

(i) Convolutional layers, which have n filters dedicated to generate n feature maps.

(ii) Subsampling layers, which each feature map is usually subsampled by a maximum value-based clustering operation called MaxPooling, which helps to progressively reduce the spatial size of the representation and the number of parameters to be trained.

(iii) Dense layers, which are layers with respectively connected neurons.

Figure 4 shows how the architecture proposed by Simonyan [10] is composed. The first column shows the layer number and the type of operation, the second column indicates the number of feature maps. Finally, we have the size, which indicates the number of output features of each of the layers. The purpose of using such a model is to classify the input image into one of several classes according to the training dataset.

The results of the convolutional neural network were satisfactory, only training 50-100 epochs (Figure 3), which avoided overtraining and a low percentage.

The level of accuracy was 97% (Figure 4) which was reflected in the mobile application developed for the Android Operating System, which can be observed that as the times pass the system stabilizes until an optimal result is obtained.

B. Results of the Mobile Application

Within the development of the mobile application and implementing the proposed architecture, we have as a result a user-friendly and functional application.

In this work we managed to classify three types of market where lemon producers can sell their product, this was achieved thanks to the implementation of a neural network, achieving the combination with a mobile application developed in Android Studio, the results of the aforementioned network had an accuracy of 97%. One of the points that resulted in more work was the creation of the dataset, as it had to be assembled from scratch and had to visit lemon orchards and packing houses to create it and thus the network had a better performance.

V. Conclusions and Future Work

As future work we propose the development of a multiplatform application for mobile devices, since it would allow reaching more devices and would have more reach to the users who need it, farmers or people in the field.

Another point to take into account is that the work thanks to the convolutional neural networks based on the quality standards of the Persian lemon, can know which market could be exported, so in future work could have more functions such as detection of diseases or pests, which would help a lot to know which lemon is the idea for such consumption.

The dataset which was worked on was obtained from packing houses located in Martínez de la Torre within the city of Veracruz, Mexico, which based on experts on the subject, commented on the quality characteristics that their company has, in this case called CITREX, which a point to be developed in the future, can be more types of lemon or the detection of diseases of different fruits, so that more than a mobile application is a tool of great help for a larger scope of farmers.

The Convolutional Neural Networks (CNN) were a very helpful tool for supervised learning, where we indicate the labels that were established in this work, so a point to improve would be the implementation and comparison with the different techniques of Machine Learning, within Artificial Intelligence.

References

[1] F. V. Pat, C. I. Caamal, A. F. Jerónimo and T. R. Mendoza, “Costos y Competitividad de la Producción del Limón Persa en el Municipio de Martínez de la Torre, Veracruz,” *Ciencias Sociales: Economía y Humanidades*, 2015.

[2] R. R. Ruiz, G. V. Vela and L. R. G. Moreno, “Exportación de cítricos mexicanos, alternativas para el mercado de

ISSN 2395-8618

exportación," *Horizontes de la Contaduría en las Ciencias Sociales*, vol. 3, no. 6, pp. 77–85, 2016.

[3] J. Y. AlZamily and S. S. A. Naser, "Lemon classification using deep learning," *International Journal of Academic Pedagogical Research (IJAPR)*, vol. 3, no. 12, pp. 16 –20, 2020.

[4] A. Jahanbakhshi, M. Momeny, M. Mahmoudi and Y. D. Zhang, "Classification of sour lemons based on apparent defects using stochastic pooling mechanism in deep convolutional neural networks," *Scientia Horticulturae* vol. 263, 2020. DOI:10.1016/j.scienta.2019.109133.

[5] J. L. Carranza Flores, "Detección automática del daño causado por el minador de los cítricos en la hoja de limón mexicano," Tecnológico Nacional de México, 2020.

[6] M. van Dael, S. Lebotsa, E. Herremans, P. Verboven, J. Sijbers, U. L. Opara and B. M. Nicolaï, "A segmentation and classification algorithm for online detection of internal disorders in citrus using X-ray radiographs," *Postharvest Biology and Technology*, vol. 112, 2016. DOI:10.1016/ j.postharvbio.2015.09.020.

[7] P. Dhiman, V. Kukreja, P. Manoharan, A. Kaur, M. M. Kamruzzaman, I. B. Dhaou and C. Iwendi, "A Novel Deep Learning Model for Detection of Severity Level of the Disease in Citrus Fruits," *Electronics*, vol. 11, no 3, pp. 495, 2022. DOI:10.3390/electronics11030495.

[8] M. Momeny, A. Jahanbakhshi, K. Jafarnezhad and Y. D. Zhang, "Accurate classification of cherry fruit using deep CNN based on hybrid pooling approach," *Postharvest Biology and Technology*, vol. 166, no. 1, pp. 111204, 2020. DOI:10.1016/j.postharvbio.2020.111204.

[9] A. Hernandez, F. J. Ornelas-Rodríguez, J. B. Hurtado-Ramos and J. J. González-Barbosa, "Accuracy comparison between deep learning models for Mexican lemon classification," *International Congress of Telematics and Computing*, pp. 62–73, 2021. DOI:10.1007/978-3-030-89586-0_5.

[10] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv*, vol. 1409-1556, 2014. DOI:10.48550/ arXiv.1409.1556.

[11] A. Krizhevsky, I. Sutskever and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 25, 2012.

[12] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma and L. Fei-Fei, "Imagenet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, 2015. DOI: 10.1007/s11263-015-0816-y.