International Journal of Engineering Research and Technology. ISSN 0974-3154, Volume 14, Number 6 (2021), pp. 510-515
© International Research Publication House. http://www.irphouse.com

Approaches to Deep Learning-Based Apple Classification Scheme Selection

Edison N. Zuñiga[1], Sebastian Gordillo[2] and Fredy H. Martínez S.[3*]

[1]Facultad Tecnológica, Universidad Distrital Francisco José de Caldas, Colombia.

[2]Facultad Tecnológica, Universidad Distrital Francisco José de Caldas, Colombia.

[3]Associate Professor, Facultad Tecnológica, Universidad Distrital Francisco José de Caldas, Colombia.
ORCID: 0000-0002-7258-3909

**Corresponding Author*

Abstract

The apple as food has great nutritional importance given its high content of calcium, phosphorus, and stone minerals, as well as vitamin B and C, but it also has a strong impact on many developing economies for which it is an important economic livelihood. In many of these environments, fruit cultivation, as is the case with many other agricultural products, is carried out by small farming families without greater access to modern technology and resources to support their processes and increase their levels of quality and productivity. In this sense, this study proposes the development of an automated system for the classification of fruit from digital cameras. The objective is to identify an architecture based on deep learning capable of performing the categorization with high performance, to later recommend it for the design of an autonomous embedded system. In this sense, three convolutional architectures are evaluated using metrics that assess the generalization capacity of the models, and the one with greater possibilities of use on a real system is selected.

I. INTRODUCTION

Much has been said in recent years about fruit grading strategies as an application of intelligent autonomous systems as a solution to the low efficiency of traditional grading strategies in aspects such as quality, quantity, and production cost [1, 2, 3, 4, 5, 3, 6]. Convolutional neural networks are proposed as one of the most promising solutions since they allow the construction of systems that focus on the physical characteristics of the fruit, leaving aside elements of fruit-to-fruit variability, which allows their use in autonomous systems to improve the state in which the fruit is picked, and to be able to locate the centers where the fruit is selected, processed and distributed. These types of systems have been widely documented in the specialized literature [7, 8, 9]. Artificial neural networks are mainly used for data classification, pattern recognition, and behavioral predictions where generalization is essential, so they are optimal to perform in the classification of the state of fruits such as apple [7, 2], application in which could generate a positive and important impact on the marketing of this fruit [10, 11].

This research seeks to evaluate the behavior of some deep convolutional network schemes to identify some with adequate performance for the development of an automatic classification system, particularly for the apple market [12, 13]. In parallel investigations of the research group have identified some architectures with higher performance in this specific type of applications, so it is proposed the specific evaluation of these topologies with an own dataset adjusted to determine specific performance characteristics on this fruit [14].

No matter how good a categorization model is, if it lacks a complete and unbiased database, it will not be possible to achieve a viable result [6, 15, 4, 16, 4, 17, 18, 19, 20, 21, 22, 23, 1]. To build a model with efficient performance, a complete database is needed, which has as a foundation the categorization states correctly separated into groups of similar size. In terms of fruit processing in distribution systems, these categories should consider the handling states of the fruit, in particular, green, ripe, and extremely ripe or damaged fruit. Once the database is complete, it is possible to proceed to generate a classification model, an image pre-processing system, an embedded system capable of real-time processing, and an overall design that allows the system to be incorporated into existing fruit handling systems.

In many developing countries the processes related to the determination of the state and quality of fruit are in the hands of human operators, who can consider a large number of changes in the quality of the fruit by visual inspection, but it entails a high cost as well as restrictions in the marketing process. Consequently, this research proposes the development of an automatic system supported by deep networks, which in its first stage establishes the appropriate architecture for the learning model, about the dataset of the problem [19, 6]. Consequently, three convolutional neural network topologies are evaluated concerning their ability to classify apples into three categories defined according to their physical conditions, and the most suitable topology will be chosen for a fine design of an embedded automatic classification system [24].

Similar equipment is used in many fruit marketing systems for automatic sorting, and even as a strategy for identifying fruit rotation in supermarkets [2]. In addition, for many countries, the commercialization of fruit in local and international markets constitutes an important part of the economic income, so this type of technology has an important positive impact on a country's economy [19, 16].

II. PROBLEM FORMULATION

The apple industry and its worldwide consumption have made apple quality standards extremely demanding. Therefore, harvesting and handling during commercialization require highly reliable and efficient grading processes to identify at least the three basic stages in which the fruit can be (green, ripe, and rotten). For this reason, industries rely on human personnel to grade these fruits. However, this strategy involves a high cost and a lower level of production compared to automatic systems which, based on neural models, can replicate the function of a human operator, guaranteeing not only the quality of the product at each stage but also it's level of production, which facilitates the projection and investment (Fig. 1).

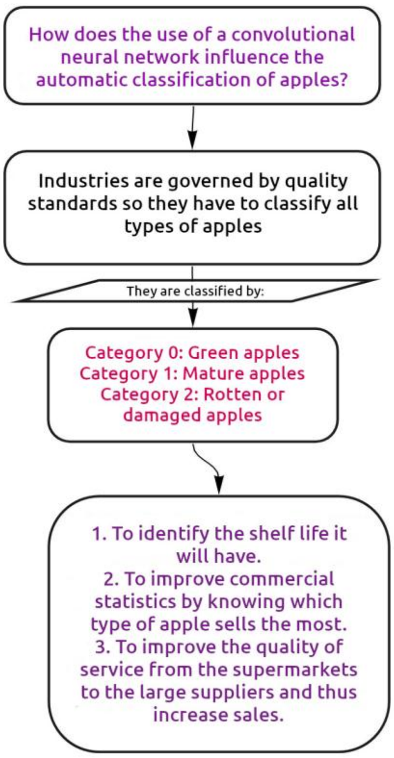

Fig. 1. Fruit selection features

The performance of this type of system is strongly linked to its learning capacity, so the dataset used for its training is fundamental. In this sense, the automatic classification system must be tuned to a specific set of images that represent the reality of the problem, and the optimal selection of the categorization architecture according to its specific performance must be guaranteed. The solution approach starts from the definition of a proper dataset according to the system performance characteristics, the minimum categories contemplated include damaged, mature and green apples, being able to consider other aspects such as size and shape in the future. For this dataset, the most adequate neural architecture for its implementation in an embedded system must be selected according to its performance, which implies training and evaluating it with the same optimizers and metrics.

III. METHODOLOGY

Convolutional neural networks correspond to receptive fields of neurons in the visual cortex of a biological brain. In the performance evaluation, we have selected three of these topologies: NasNet (Neural Architecture Search Network), ResNet (Residual Neural Network), and DenseNet (Dense Convolutional Network).

NasNet convolutional networks, like the other two models, are deep networks specialized in image identification and recognition. Although they correspond to architectures that require supervised training, in many cases, it is useful to use pre-trained architectures when there is not enough data to train them. This model allows a small dataset to be used to search for a structural building block that can then be transferred to a larger dataset.

A ResNet system is characterized by replicating the pyramidal cell architecture in the cerebral cortex. For this, the model uses forward jumps over some of the layers to send input information forward as input to other layers, where this information is added to the traditional connection (Fig. 2).

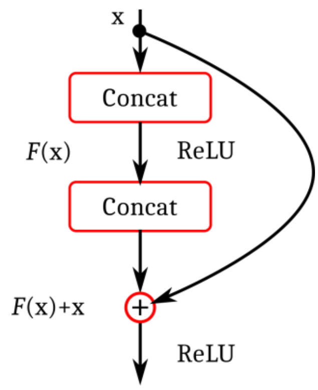

Fig. 2. Residual Neural Network (ResNet) architecture block

DenseNet is an extension of the residual ResNet model that adds residual connections not only to the module output but to all consecutive ones (Fig. 3). However, unlike ResNet, it does not combine features by summing but by concatenating. Fewer parameters and high accuracy are achieved compared to ResNet and pre-activation ResNet (the result, however, depends very much on the specific case understudy). Since each layer receives feature maps from all previous layers, the network can be thinner and more compact, the error signal can be easily propagated to previous layers more directly. In DenseNet, the classifier uses features of all complexity levels

International Journal of Engineering Research and Technology. ISSN 0974-3154, Volume 14, Number 6 (2021), pp. 510-515
© International Research Publication House. http://www.irphouse.com

and tends to give smoother decision boundaries, as well as reducing the leakage gradient problem. They also improve feature propagation and reduce the number of parameters within the network. One of the main improvements that DenseNet presents is that it manages to shorten the connections between layers near the input and output, thus increasing the density of the network.

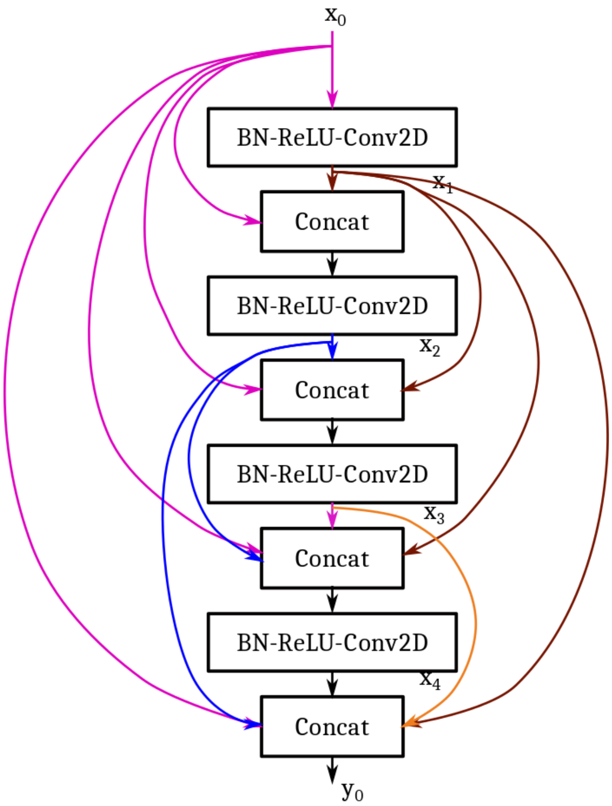

Fig. 3. Densely Connected Convolutional Network (DenseNet) architecture

The functional performance evaluation is performed with these three architectures. In each case, the same dataset of 225 images was used in each of the three categories. The three pieces of training were developed over 65 epochs and used the same loss functions (Categorical Cross Entropy) and the same optimizer (Stochastic Gradient Descendent - SGD). The dataset was randomly separated in all cases into two groups, 70% for training and 30% for validation (these groups were always randomly mixed before separation).

The performance of these models was evaluated at two points in time, observing the behavior during training, and after training from metrics with the validation data. During training, at each epoch, the accuracy and loss values were calculated. This was done both for the data used in training and for the unknown data in the validation group. This information was used to determine the generalization capacity of each model, as well as its performance under more demanding training. After training, the validation data were used to construct confusion matrices for each model. From them, we were also able to establish the metrics Precision (for each category, and model average), Recall, and the cumulative value of these F1-score.

IV. RESULTS AND DISCUSSION

Fig. 4 shows the accuracy behavior during training for each of the three models evaluated. For each epoch, the model accuracy was calculated for both training and validation data. In all three cases, the initial value was very low, but grew over the epochs to almost 100%, at least for the training data (red curve). The validation data behaved differently, for the NasNet model the accuracy did not rise above 40%, which already shows a mediocre performance of this architecture. The ResNet model reached a final value of 82%, and the DenseNet model 90%. The latter two showed good performance in generalization without overfitting.

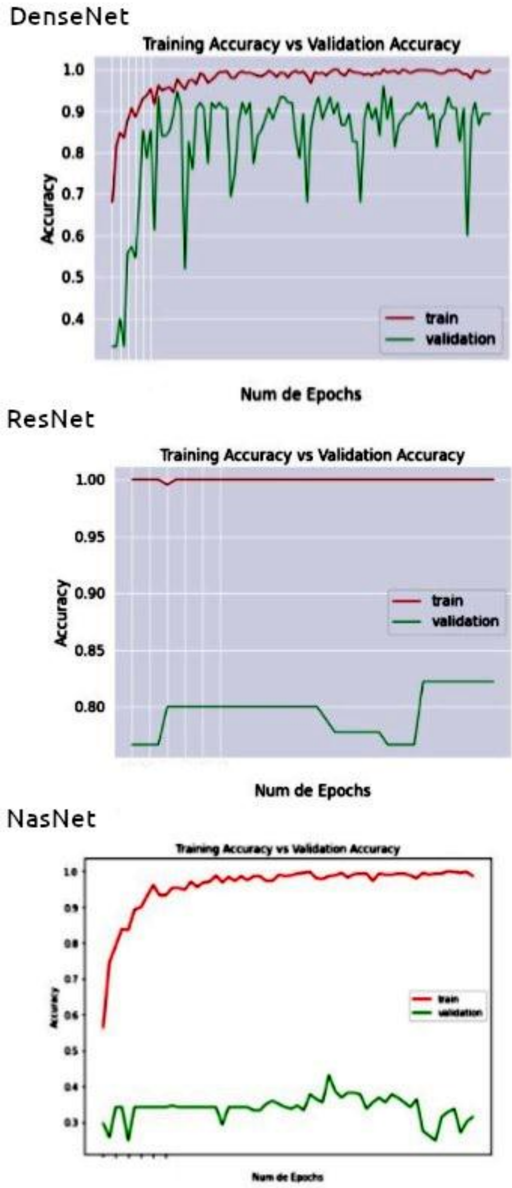

Fig. 4. Accuracy of each of the models evaluated throughout the training process. The red curve corresponds to the training data and the green curve to the validation data

The behavior of the loss function in Fig. 5 follows these same results. This function was also calculated throughout the training, again for training data and validation data. The ResNet model again shows mediocre behavior for the validation data (green curve) with values that do not go below 0.7. All three models show good behavior with training data (red curve), but only the DenseNet and NasNet models also show good values for validation data.

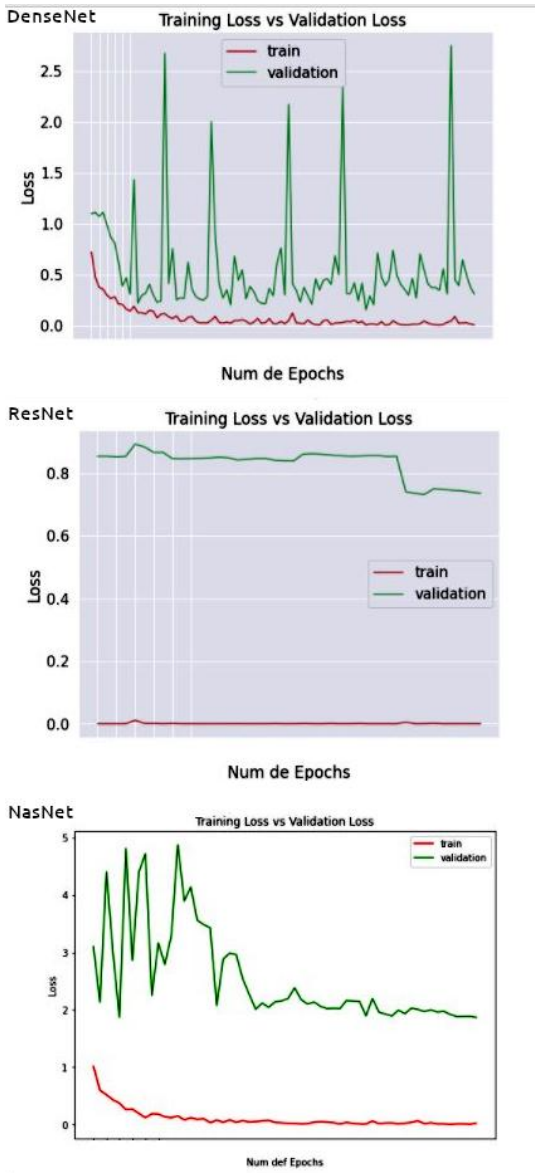

Fig. 5. Loss of each of the models evaluated throughout the training process. The red curve corresponds to the training data and the green curve to the validation data

The confusion matrix of the classification models allows observing graphically the capacity of the model to categorize each data within the correct category (Fig. 6). In the color scale used, the largest number of elements is represented by light colors and the smallest number of data by dark colors. Therefore, the best performance is obtained by the model whose diagonal contains lightboxes, while the boxes above and below the diagonal remain dark. This type of behavior can be observed in the DenseNet model, followed by the ResNet model. The NasNet model has recognition problems and places most of the data in the second category. Even so, under the diagonal of the matrix for the DenseNet model, there are squares that are not completely dark, meaning that elements are being classified in this column that do not belong to it. The fine analysis of this behavior is observed in the metrics derived from these confusion matrices.

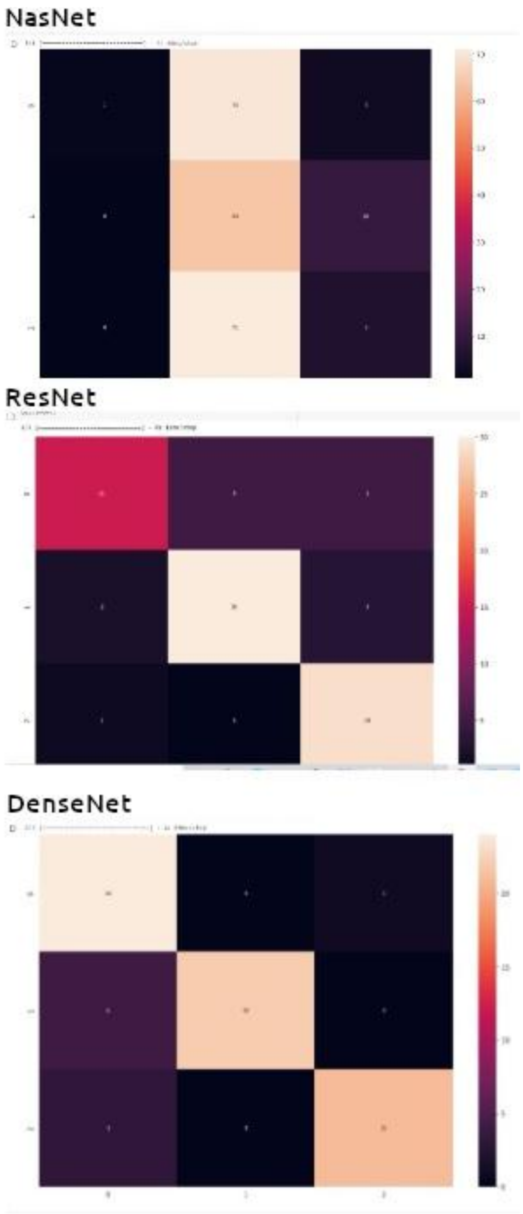

Fig. 6. Confusion matrix of each of the models calculated with the validation data. The color scale assigns light colors to the highest number of elements, and dark colors to the lowest number

From the data in the boxes of the confusion matrix of each of the models, it is possible to calculate the precision and recall metrics and with them the f1-score metric. These metrics show in detail the categorization ability of each of the models, either to place each data in the correct column, the ability to place only the correct data in the correct categories, and the average behavior per model. The results for the three models are shown in Fig. 7. The best average performance for the three metrics was obtained by the DenseNet architecture, with values close to 90%. This model was followed by the ResNet architecture with values in the order of 81% and last place the NasNet model with 54% in accuracy, 32% in the recall, and 20% in f1-score.

The individual values per category show that the NasNet model managed to place all the corresponding elements in the first category, but also placed there elements that did not correspond to the category. The ResNet and DenseNet models performed better in terms of the correct categorization of elements.

NasNet

	precision	recall	f1-score
0	1.00	0.01	0.03
1	0.31	0.86	0.46
2	0.32	0.08	0.12
accuracy			0.32
macro avg	0.54	0.32	0.20
weighted avg	0.54	0.32	0.20

ResNet

	precision	recall	f1-score
0	0.83	0.60	0.70
1	0.86	0.86	0.86
2	0.78	0.97	0.87
accuracy			0.82
macro avg	0.82	0.81	0.81
weighted avg	0.83	0.82	0.82

DenseNet

	precision	recall	f1-score
0	0.77	0.96	0.86
1	1.00	0.85	0.92
2	0.95	0.88	0.91
accuracy			0.89
macro avg	0.91	0.89	0.90
weighted avg	0.91	0.89	0.90

Fig. 7. Metrics for each of the models are calculated from the confusion matrix. Average values and values for each category are shown

In all the tools used to evaluate the performance of the classification models, the best performance was achieved by the DenseNet model, followed by the ResNet architecture. These two models can be used for the development of the embedded system since their values indicate a high capacity for fruit identification. The NasNet model has a poor performance in all metrics, showing its inability to work with unknown images, i.e., it is not able to generalize in the identification of specific parameters that determine the state of the fruit.

V. CONCLUSION

This paper evaluates the performance of three image classification models based on deep learning to identify the most suitable one for use in an embedded system for the automatic classification of apple fruit. Such a system is proposed as an alternative for automation in fruit growing and processing systems, as opposed to manual techniques carried out by operators, to guarantee quality and increase production. For the evaluation, three architectures identified as suitable for this application were selected based on their use in similar systems. The topologies used are NasNet (Neural Architecture Search Network), ResNet (Residual Neural Network), and DenseNet (Dense Convolutional Network). These three models were trained with a proprietary dataset under the same conditions, and evaluated with the same metrics, both during and after training, mainly by observing performance with a dataset not used in training. All three models were able to replicate the dynamics for the training data, but for the validation data, the NasNet model performed poorly, showing its inability to generalize the system dynamics. The best performance in all metrics was achieved by the DenseNet architecture, which is recommended for fine training of the system and its possible implementation on an embedded system. The ResNet architecture also showed good performance and is postulated as a second working option. The research continues with a fine-tuning of these two models, and their evaluation with larger datasets.

ACKNOWLEDGMENTS

This work was supported by the Universidad Distrital Francisco José de Caldas, specifically by the Technological Faculty. The views expressed in this paper are not necessarily endorsed by Universidad Distrital. The authors thank all the students and researchers of the research group ARMOS for their support in the development of this work.

REFERENCES

[1] A. Lerma, A. Santoyo, F. Villalobos, and P. Reyes, "Clasificación automática de anastomosis mediante redes neuronales convolucionales en video fetoscópico", Revista Iberoamericana para la Investigación y el Desarrollo, vol. 12, no. 1, p. 25, 2021, ISSN 2007-7467, doi:10.23913/ride.v11i22.856.

[2] C. Delfin, C. Gonzalez, and J. Olguin, "Clasificación de manzanas utilizando visión artificial y redes neuronales artificiales", Ingeniería y Región, vol. 20, no. 2, pp. 54–58, 2018, ISSN 1657-6985, doi:10.25054/22161325.1917.

[3] C. Holguin, J. Cortes, and J. Chaves, "Sistema automático de reconocimiento de frutas basado en visión por computador", Revista Chilena de Ingeniería, vol. 22, no. 4, pp. 504–516, 2014, ISSN 0018-926X, doi:10.25054/22161325.1917.

[4] J. Wen, E. Sutre, D. Mauricio, J. Samper, A. Routier, S. Bottani, D. Dormont, and S. Durrleman, "Convolutional neural networks for classification of alzheimer's disease: Overview and reproducible evaluation", Medical Image Analysis, vol. 63, no. 101694, pp. 1–9, 2020, ISSN 1361-8415, doi:10.1016/j.media.2020.101694.

[5] J. Barrera, "Clasificación de imágenes hiperespectrales de frutas y verduras mediante aprendizaje profundo", Instituto Tecnológico Metropolitano, , no. 4, pp. 1–94, 2019.

[6] S. Gonzalez, "Evaluación del uso de redes neuronales convolucionales para clasificación de cultivos mediante imagen multiespectral", Universidad de Valladolid, vol. 3, no. 4, p. 94, 2019, ISSN 0018-926X, doi:13246-020-00865-4.

[7] V. Gutierrez, "Comparación de arquitecturas de redes neuronales convolucionales para la clasificación de enfermedades en tomate", Universidad Autonoma de Zacatecas, no. 1, p. 23–32, 2019.

[8] Y. Zhang, S. Wang, G. Ji, and P. Phillips, "Fruit classification using computer vision and feedforward neural network", Journal of Food Engineering, vol. 6, no. 4, p. 11, 2014, ISSN 0018-926X, doi:13246-020-00865-4.

[9] D. Fajardo, A. Toalombo, and M. Franklin, "Desarrollo de un clasificador de aceitunas de mesa por su variedad", Universidad de Guayaquil, , no. 1, p. 25, 2020.

[10] R. Moreno, G. López, C. Buritica, F. Aldana, and F. Martínez, Los recursos distribuidos de bioenergía en Colombia, Universidad Distrital Francisco José de Caldas, 2020, ISBN 978-958-787-258-3.

[11] W. Sánchez and V. Guerrero, "Ecology, construction, and innovation: An alert towards change", Tekhnê, vol. 15, no. 2, pp. 45–58, 2018, ISSN 1692-8407.

[12] E. Jacinto, F. Martínez, and H. Montiel, "Driver for visualization of graphics on vga screens using fpga's", International Journal of Applied Engineering Research, vol. 13, no. 20, pp. 14728–14732, 2018, ISSN 0973-4562.

[13] A. Rendón, "Design and evaluation of volume unit (vu) meter from operational amplifiers", Tekhnê, vol. 16, no. 2, pp. 31–40, 2019, ISSN 1692-8407.

[14] F. Martínez, C. Hern"andez, and C. Rendón, "Identifier of human emotions based on convolutional neural network for assistant robot", Telkomnika, vol. 18, no. 3, pp. 1499–1504, 2020, ISSN 1693-6930, doi:10.12928/telkomnika.v18i3.14777.

[15] M. Velasquez, "Implementación de redes neuronales convolucionadas para la clasificación de imágenes de sensores remotos en la amazonia", Universidad Distrital Francisco José de Caldas, vol. 3, no. 4, p. 94, 2019, ISSN 0018-926X.

[16] J. Suarez, "Modelo de aprendizaje automático para la predicción de la calidad del café", Universidad Distrital Francisco José De Caldas, p. 118, 2019.

[17] H. Chanampe, M. Vega, J. Molinari, G. Carrascosa, and A. Lorefice, "Modelo de redes neuronales convolucionales profundas para la clasificación de lesiones en ecografías mamarias", XXI Workshop de Investigadores en Ciencias de la Computación, vol. 3, p. 6, 2019, ISSN 0018-926X, doi:13246-020-00865-4.

[18] J. Vega, "Modelo de aprendizaje profundo red neuronal convolacional (cnn) para clasificación de calidad de acidos grasos por imagenes de semillas de helianthus annuus", Universidad Nacional de La Plata Facultad de Informática, vol. 3, no. 4, p. 94, 2019, doi:13246-020-00865-4.

[19] O. Cristobal, "Clasificación no destructiva de frutas utilizando inteligencia artificial: Revisión bibliográfica", Universidad Señor de Sipán, vol. 3, no. 4, p. 94, 2019, doi:13246-020-00865-4.

[20] F. Quiroga, "Medidas de invarianza y equivarianza a transformaciones en redes neuronales convolucionales. aplicaciones al reconocimiento de formas de mano.", Universidad Nacional de La Plata Facultad de Informática, vol. 3, no. 4, p. 94, 2013, ISSN 0018-926X.

[21] M. Lopez, "Identificación de sistemas no lineales con redes neuronales convolucionales", Centro de Investigacion y de Estudios Avanzados del Instituto Politecnico Nacional, vol. 3, no. 4, p. 94, 2017, ISSN 0018-926X, doi:13246-020-00865-4.

[22] G. Gutierrez and J. Ballesteros, "Identificación del estado de madurez de las frutas con redes neuronales artificiales, una revisión", Revista Ciencia y Agricultura, vol. 13, no. 1, pp. 117–132, 2016, ISSN 0122-8420.

[23] I. Apostolopoulos and T. Mpesiana, "Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks", Physical and Engineering Sciences in Medicine, vol. 3, no. 4, p. 94, 2020, ISSN 0018-926X, doi:13246-020-00865-4.

[24] G. Figueredo, "Desarrollo de un clasificador de aceitunas de mesa por su variedad", Universidad Pedagógica y Tecnológica de Colombia, , no. 1, p. 23–32, 2016.