



# Segmentation Methods for Image Classification Using a Convolutional Neural Network on AR-Sandbox

Andres Ovidio Restrepo Rodriguez<sup>1</sup>(✉),  
Daniel Esteban Casas Mateus<sup>1</sup>, Paulo Alonso Gaona Garcia<sup>1</sup>,  
Adriana Gomez Acosta<sup>2</sup>, and Carlos Enrique Montenegro Marin<sup>1</sup>

<sup>1</sup> Universidad Distrital Francisco Jose de Caldas, Bogota, Colombia  
{aorestrepo, decasasm}@correo.udistrital.edu.co,  
{pagaonag, cemontenegrom}@udistrital.edu.co

<sup>2</sup> Fundación Universitaria San Mateo, Bogota, Colombia  
acgomez@funсанmateo.edu.co

**Abstract.** Fields such as early education and motor rehabilitation provide a space for their integration with augmented reality devices as AR-Sandbox, in order to provide support to these fields, generating a feedback of tasks carried out through the recognition of images based on convolutional neural networks. However, the nature of the AR-Sandbox generates a high noise level of the acquired images, for this reason the present study has as purpose the implementation and comparison of three segmentation methods (Canny Edge Detector, Color-space and Threshold) for the training and prediction phase of a convolutional neural network model previously established. When carrying out this study, it was obtained that the combined model with color-space segmentation presents an average percentage of 99% performance for the classification of vowels, described by the AUC of the ROC curve, this being the model with the best performance.

**Keywords:** Canny Edge Detector · Color-space · Convolutional Neural Networks · Image processing · Segmentation · Threshold · Vowels

## 1 Introduction

The AR-Sandbox is an augmented reality device, which, through the projection of digital topographic maps, allows to close the gap between 2D and 3D visualization, improving the thought and spatial modeling skills. This device has a large field of action, providing support and revolutionizing the traditional paradigm of education in early childhood and in medicine such as rehabilitation through motor therapy [1], for this, it is important that the AR-Sandbox provide feedback of the activity carried out.

The image recognition on the AR-Sandbox presents a high degree of difficulty, since when capturing the image from this device a high level of noise is presented, due to the multiple pigmentations according to the level of the sand, exposed by the nature

of this augmented reality tool. The AR-Sandbox makes use of a Kinect and a projector to make colorful representations in 3D on a sandbox [2].

This article presents preliminary results for model implementation results of a convolution neural network (CNN), as an image recognition technique on the AR-Sandbox, evaluating the application of three segmentation methods to the training dataset and the test dataset, for the training and prediction phase, in order to determine which segmentation method best fits the characteristics of the AR-Sandbox.

## 2 Background

Previous image processing besides image recognition through convolutional neural networks have had a large number of applications, which its goal is to optimize pattern recognition [3]. From different computer vision techniques, convolutional neural networks have been in a continuous growth since their creation [4].

Image segmentation methods are used to optimize the analysis and interpretation of images through a process depending on the segments of each image, there are a variety of algorithms and methods with advantages and disadvantages depending on the features of the image, these methods are classified depending on pixel intensity value, color, texture, etc. [5]. Authors like [6] use image segmentation in order to get information for underwater images, which are difficult to get and analyze due to environmental conditions, for these reasons, canny edge detector algorithm is used because it gave localization and response with extreme accuracy and it has a superior capability under noise conditions.

Finally, in previous studies, [7] where the purpose of this research was to establish a model of a CNN for the classification of geometric figures by optimizing hyperparameters using random search, evaluating the impact of the implementation of a previous phase of color-space segmentation to a set of tests captured from the AR-Sandbox, authors find that using the proposed method, an average decrease of 39.45% to a function of loss and an increase of 14.83% on average in the percentage of correct answers is presented.

Based on this panorama, the present study intends to carry out the recognition of images by the means of deep learning techniques, such as convolutional neural networks and image processing by three segmentation methods, with the purpose of determining the performance variation of a convolutional neural network, applying these segmentation methods the test and training datasets. In the next section, the selected study scenario is presented.

## 3 General Approach

Figure 1, presents the general approach of this study, which has as its starting point a base model of a convolutional neural network. The study starts from previous results of [7] that had the purpose of creating contributions towards image recognition based on immersive techniques, through deep learning using CNN, hyperparameters optimization and image processing. From this base model, CNN is trained, using a training

dataset composed of a series of vowels (A, E, I, O, U), which will be previously put through image processing, through Python and OpenCV, using three segmentation methods, Canny Edge Detector [6], color-space segmentation [7] and Threshold [8]. When obtaining the trained model, we proceed to make the model evaluation, making use of a test dataset, which was acquired in its entirety through the AR-Sandbox. The same segmentation method applied in the training phase is applied previously to the test dataset, for the evaluation phase of the model the following evaluation items are defined: Confusion Matrix, and the ROC Curves.

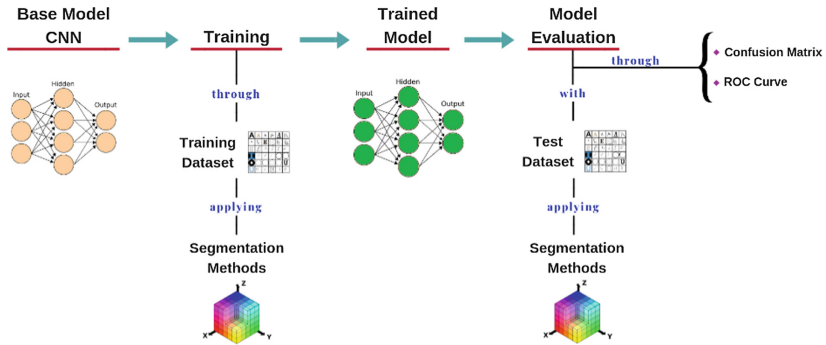


Fig. 1. General approach

### 4 CNN Base Model

In order to carry out the recognition and classification of vowels from the AR-Sandbox, a model of a convolutional neural network was selected [7], where the definition of a base model of a CNN was carried out, in order to perform the hyperparameters optimization by random search, defining a dictionary of hyperparameters. The model selected was for the recognition of geometric figures acquired from the AR-Sandbox. The selected model is presented in Fig. 2.

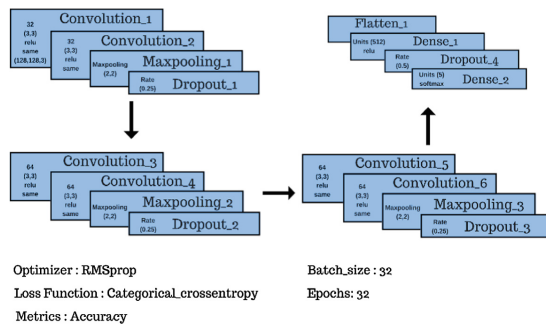


Fig. 2. CNN model

## 5 Image Processing

As a first step, the acquisition of images was done through the AR-Sandbox, which aims to support early education and rehabilitation in motor therapy, in order to change the traditional paradigm of exercises. Next, the segmentation methods and their process are presented.

### 5.1 Canny Edge Detector

As a first step, we proceed to select the original image, later the image has to be resized, in order to apply Gaussian filter and sobel procedure, finally, Non maximum suppression and hysteresis are applied to image [9]. Figure 3 presents the process.

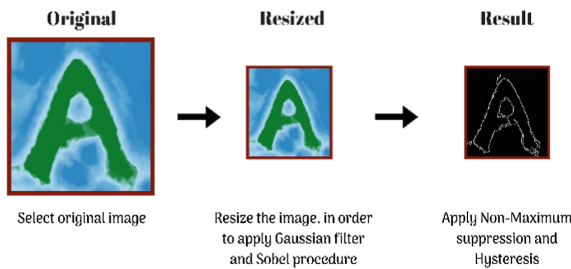


Fig. 3. Canny Edge Detector segmentation

### 5.2 Color-Space

As a first instance, the target color and a range of close colors are selected, in order to separate the colors in the image, therefore, the values within these parameters are painted of white color and the rest of a black color, this is called mask, which provides a contrast in the image. The white part of the mask will take the green color while the rest of the image is still black. Figure 4 presents the process.

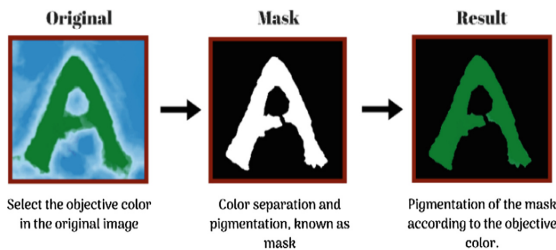


Fig. 4. Color-space segmentation

### 5.3 Threshold

As a first step, the original selected image must be passed to grayscale, later you have to apply the threshold function using Otsu binarization [10]. Figure 5 presents the process.

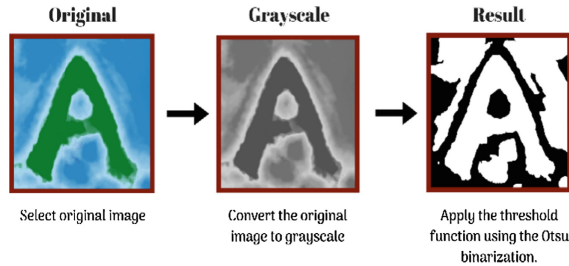


Fig. 5. Threshold segmentation

### 5.4 Similarity Indexes

In order to obtain the probability of similarity between the original image and the segmented image using the methods mentioned above, we proceed to calculate the coefficients of Jaccard and Sørensen-Dice, making use of Matlab, where the individual coefficient of each image is obtained in order to obtain the average value of each of the coefficients for each segmentation method. The obtained coefficients are presented in Table 1.

Table 1. Similarity indexes

Method	Index	
	Jaccard	Sorensen-Dice
Canny	0.3239	0.4356
Color-space	0.7143	0.7801
Threshold	0.4869	0.6244

## 6 Results

In this section the results are shown, in the first part the results show the ROC curves of the used test dataset and finally the models are compared by success rate obtained from the confusion matrices of each segmentation method.

To acquire the images and make the test dataset of images which are made in the AR-Sandbox, it was necessary to convene a group of ten people who made a total of 15 images each, which are 3 for each class.

### 6.1 ROC Curves

Figure 6A shows the ROC curves for the dataset with color-space, in Fig. 6B the ROC curve for the dataset with Canny Edge Detector is presented, in Fig. 6C the curve for the dataset with Threshold and Fig. 6D the curve without a segmentation method. Each of these figures has seven curves which are areas under the curve of the micro and macro level of the ROC curve, in addition the curves corresponding to each vowel represented by a number from 1 to 5 in alphabetical order.

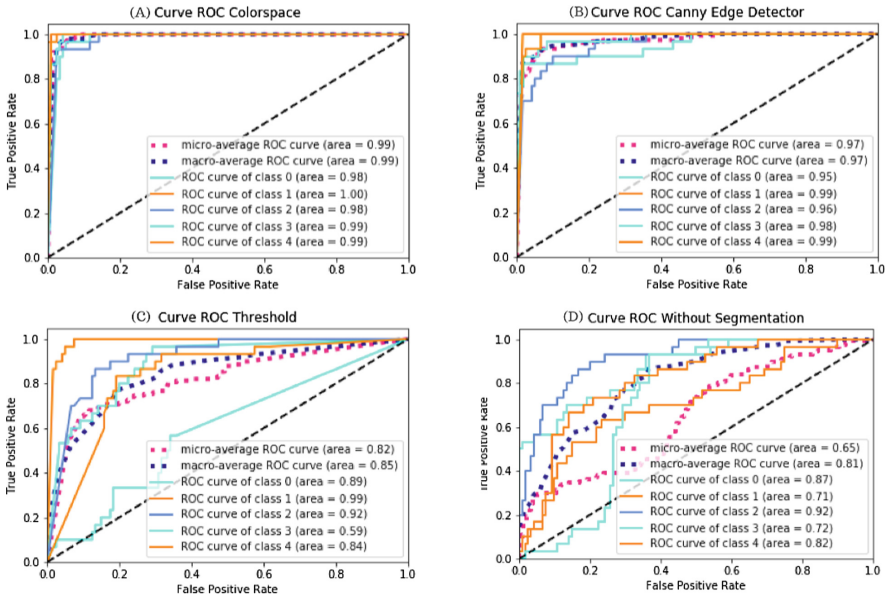


Fig. 6. ROC curves.

### 6.2 Confusion Matrix

With the help of the Python sk-learn library an algorithm is implemented for the creation of confusion matrices that will show the success rate of each method with respect to the test dataset. Table 2, shows the percentage of success rate of each segmentation method.

Table 2. Success rate segmentation methods

	Method			
	Without segmentation	Canny	Color-space	Threshold
Success Rate	57.33%	87.33%	92.67%	58.00%

## 7 Analysis of Results

Regarding the confusion matrix, the highlighted segmentation method is color-space with 92.67% having a difference of 5.34% with Canny Edge Detector, 34.67% with Threshold and 57.33% without segmentation method.

As evidenced in Fig. 6, the model that stands out with respect to the AUC, making the average between micro average and macro average, is the Color-space model with a 0.99 performance, followed by Canny Edge Detector with a 0.97, Threshold with 0.835 and finally the model without segmentation method with 0.73. In addition, it can be seen that the model with color-space segmentation increases the precision value by 0.26. Table 3 presents a detailed analysis by vowel.

**Table 3.** Specific analysis: ROC curve

Method		Vowel				
		A	E	I	O	U
Canny	AUC	0.95	0.99	0.96	0.98	0.99
Color-Space		0.98	1.00	0.98	0.99	0.99
Threshold		0.89	0.99	0.92	0.59	0.84
Without segmentation		0.87	0.71	0.92	0.72	0.82

## 8 Conclusions

From the analysis presented and taking as criteria the ROC curve and confusion matrix, it was determined that the segmentation method with the highest performance for the recognition of vowels on CNN is Color-space, given the multiple pigmentation characteristics of the AR-Sandbox, the choice of a target color represents a greater percentage of correct answers in its prediction phase. By applying the same segmentation method in the training and prediction phase, positive contribution is made in the identification and extraction of characteristics for the classification of images, as mentioned by the authors in [7].

On the other hand, it is determined that the similarity indexes are not decisive for the choice of a segmentation method for the recognition of images by CNN in an AR-Sandbox, since as can be seen the Canny Edge Detector method presents the lowest coefficients of similarity contradicting the results of confusion matrices and ROC curves that show a high performance of the CNN trained with this segmentation method.

From a model with a high level of precision for objects identification from AR-Sandbox, one can begin to expand the spectrum of functionality and use this augmented reality device in early education and rehabilitation.

According to the above, the combination of multimedia, deep learning together with image processing provides an opportunity to perform a multidisciplinary work in parallel with traditional paradigms in the fields of education and medicine.

## References

1. Woods, T., Reed, S., Hsi, S., John, W., Woods, M.: Pilot study using the augmented reality sandbox to teach topographic maps and surficial processes in introductory geology labs. *J. Geosci. Educ.* **64**(3), 199–214 (2016)
2. Giorgis, S., Mahlen, N., Anne, K.: Instructor-led approach to integrating an augmented reality sandbox into a large-enrollment introductory geoscience course for nonmajors produces no gains. *J. Geosci. Educ.* **65**(3), 283–291 (2017)
3. Hijazi, S., Kumar, R., Rowen, C.: Using convolutional neural networks for image recognition. Cadence (2015). [https://ip.cadence.com/uploads/901/cnn\\_wp-pdf](https://ip.cadence.com/uploads/901/cnn_wp-pdf). Accessed 05 Febrero 2019
4. Neha, S., Vibhor, J., Anju, M.: An analysis of convolutional neural networks for image classification. *Procedia Comput. Sci.* **132**, 337–384 (2018)
5. Kaur, D., Kaur, Y.: Various image segmentation techniques: a review. *Int. J. Comput. Sci. Mob. Comput.* **3**(5), 809–814 (2014)
6. Alex Raj, S.M., Jose, C., Supriya, M.H.: Hardware realization of canny edge detection algorithm for underwater image segmentation using field programmable gate arrays. *J. Eng. Sci. Technol.* **12**(9), 2536–2550 (2017)
7. Restrepo Rodriguez, A.O., Casas Mateus, D.E., Gaona Garcia, P.A., Montenegro Marin, C. E., Gonzalez Crespo, R.: Hyperparameter optimization for image recognition over an AR-Sandbox based on convolutional neural networks applying a previous phase of segmentation by color-space. *Symmetry* **10**(12), 743 (2018)
8. Rui, W., et al.: Threshold segmentation algorithm for automatic extraction of cerebral vessels from brain magnetic resonance angiography images. *J. Neurosci. Meth.* **241**, 30–36 (2015)
9. OpenCV, Canny Edge Detector. [https://docs.opencv.org/2.4/doc/tutorials/imgproc/imgtrans/canny\\_detector/canny\\_detector.html](https://docs.opencv.org/2.4/doc/tutorials/imgproc/imgtrans/canny_detector/canny_detector.html). Accessed 2 Feb 2019
10. Mordvintsev, A., Abid, K.: Image Segmentation with Watershed Algorithm, OpenCV, (2013). [https://opencv-python-tutroals.readthedocs.io/en/latest/py\\_tutorials/py\\_imgproc/py\\_watershed/py\\_watershed.html](https://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_imgproc/py_watershed/py_watershed.html). Accessed 2 Feb 2019