Facial expression recognition system for images focus control

Jorge Gomez-Rojas
Engineering faculty
Universidad del Magdalena, Santa Marta, Magdalena, Colombia.
Email - jgomez@unimagdalena.edu.co

Rafael Linero-Ramos
Engineering faculty
Universidad del Magdalena, Santa Marta, Magdalena, Colombia.
Email – rlineror@unimagdalena.edu.co

Ronald Martínez-Abuabara
Engineering faculty
Universidad del Magdalena, Santa Marta, Magdalena, Colombia.
Email – rmartinez@unimagdalena.edu.co

Yesica Beltran-Gomez
Engineering faculty
Universidad del Magdalena, Santa Marta, Magdalena, Colombia.
Email – ybeltrang@unimagdalena.edu.co

Abstract- Human ocular refractive errors (myopia, hyperopia, presbyopia and astigmatism) are the most common visual problems in the world, represents 42% of the causes of visual impairment. This paper presents the design and implementation in Matlab of a system of classification of facial expressions to control the focus of images. It uses a cascading object detector that uses the Viola-Jones algorithm to train two classifiers, a machine of vector support and a decision tree. Through the implementation of classification systems, it was possible to recognize facial expressions with percentages of yield greater than 91%.

Keywords – Classification Systems, Object Detector, Image Processing, Matlab®.

I. INTRODUCTION

Visual impairment is related to a problem that affects visual acuity, visual field, eye motility, color vision or depth, affecting a person's ability to see (Dirección General de Educación Especial, 2016).

Visual acuity is the ability to see in detail the things that surround us at a certain distance (Torres, 2006). This ability depends on the normal and coordinated functioning between the eyes and the brain. There are approximately 1300 million visually impaired people, of whom 36 million are blind and 405 million have low vision (Bourne et al., 2017). There are approximately 285 million visually impaired people, of whom 39 million are blind and 246 million have low vision (World Health Organization, 2018). Visual impairment encompasses blindness while moderate and severe visual impairment is commonly regrouped under the term low vision.

The (World Health Organization, 2013) estimates that there are 153 million people with visual impairment worldwide due to uncorrected refractive errors. Within this category are myopia, hyperopia, astigmatism and presbyopia. Currently, these errors can be treated with corrective glasses, contact lenses, or refractive surgery.

Refractive errors such as farsightedness and presbyopia cause difficulty in clearly seeing nearby objects (World Health Organization, 2013). This naturally causes facial expressions to be made to try to focus and be able to read well at short distances.

Facial expressions are different gestural manifestations through which humans using the face express emotions, moods and are also emitted in response to some specific events (Hildebrandt, Recio, Sommer, & Wilhelm, 2014). It is possible that these manifestations are considered patterns.

In pattern recognition, a classification process consists of a cycle that includes data collection, selection and extraction of characteristics, choice of a classification model, training and evaluation of the classifier (Duda, Hart, & Stork, 2001).

Cascade object detectors are used to detect these facial expressions (Viola & Jones, 2001). Cascade object detectors are used to detect these facial expressions. These can detect categories of objects whose aspect ratio does not vary significantly. The object detection used uses the Viola-Jones algorithm. This algorithm slides a window over the image and using a cascading classifier decides if the window contains the object of interest.

Supervised learning algorithms such as vector support machines are used to perform pattern classification (Duda, Hart, & Stork, 2001). This method allows the construction of hyperplanes that separate the classes to be classified, finding the margin that maximizes the distance between the different classes.

Decision trees are also used in pattern classification techniques (Duda, Hart, & Stork, 2001). This method is based on a sequence of questions, where the next question depends on the previous one and the answers to those questions are yes / no or true / false, in order to verify if a value belongs to a certain set.

This work describes the design and implementation of a system developed in Matlab that detects the eye region using a cascading object detection algorithm. Additionally, characteristics of this detected region are extracted such as the luminosity of the image, the distance between the eyes and the eyebrows. Finally, two classifiers are designed, trained and evaluated; one based on a vector support machine and the other on a decision tree.

II. METHODOLOGY

For the implementation of this system, an acquisition stage and a database elaboration stage were carried out for two specific facial expressions, definition of classes and feature vectors. Finally, the selection, design, training and evaluation of the classifiers was developed.

2.1 . Facial Expressions Database

Two specific facial expressions shown in Figures 1-4 were acquired. For this, a video input object was built, which represents the connection between Matlab and a camera as an image acquisition device. The database was made up of 100 images, with the faces of 2 different subjects, with 50 images of expression 1 and 50 images of expression 2.

2.2 Classes and feature vectors

The two facial expressions in the database define the classes of the system. The features extracted from the facial expression images were the luminosity in the eye region and the distance between the eyes and the eyebrows, in pixels.

Features are extracted using Matlab's vision.CascadeObjectDetector function. A cascading object detector enables detection of the eyes on the face, cropping this region of the image. Class 1 of the system is defined as a facial expression in which luminosity decreases in the eye area. The distance between the eyes and the eyebrows also decreases with respect to a neutral facial expression, as can be seen in Figures 1 and 2. Class 2 of the system is defined as a facial expression in which the luminosity increases in the eye area and the distance between the eyes and the eyebrows also increases with respect to a neutral facial expression, as can be seen in Figures 3 and 4.

2.3 Classification

To carry out the facial expression classification process, a vector support machine was implemented using Matlab's fitcsvm function. This uses by default a linear kernel and, from the observations made, results in the support vectors. This classifier was evaluated by means of cross validation with the crossval function using 90% of samples for training and 10% for evaluation. It must be taken into account that this 10% goes through all the observations until all of them are part of the evaluation process.

A decision tree was implemented setting the luminosity in the eye region in such a way that it allows a separation of the two classes. The chosen value is 1500831. On the other hand, the distance between the eyes and the eyebrows in the image are values less than or equal to 35 pixels for class 1 or greater than or equal to 60 pixels for class 2. This classifier was also evaluated through cross validation using 90% of samples for training and 10% for evaluation. Figure 5 shows the implemented decision tree.

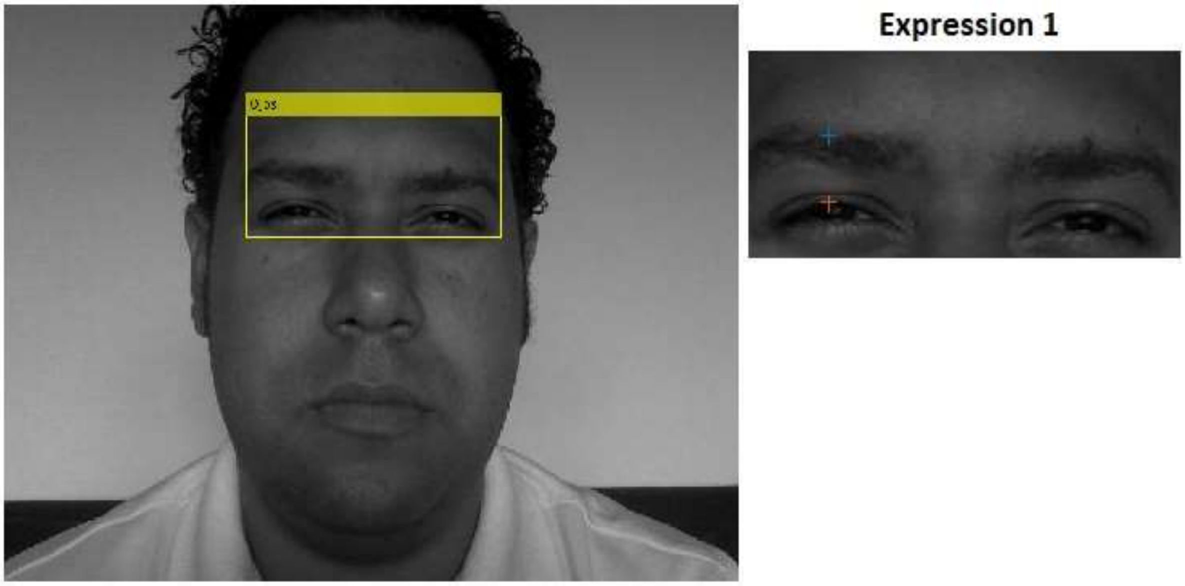

Figure 1. Detection of eyes in expression 1 - Subject 1

Source: Authors

Figure 2. Detection of eyes in expression 1 - Subject 2

Source: Authors

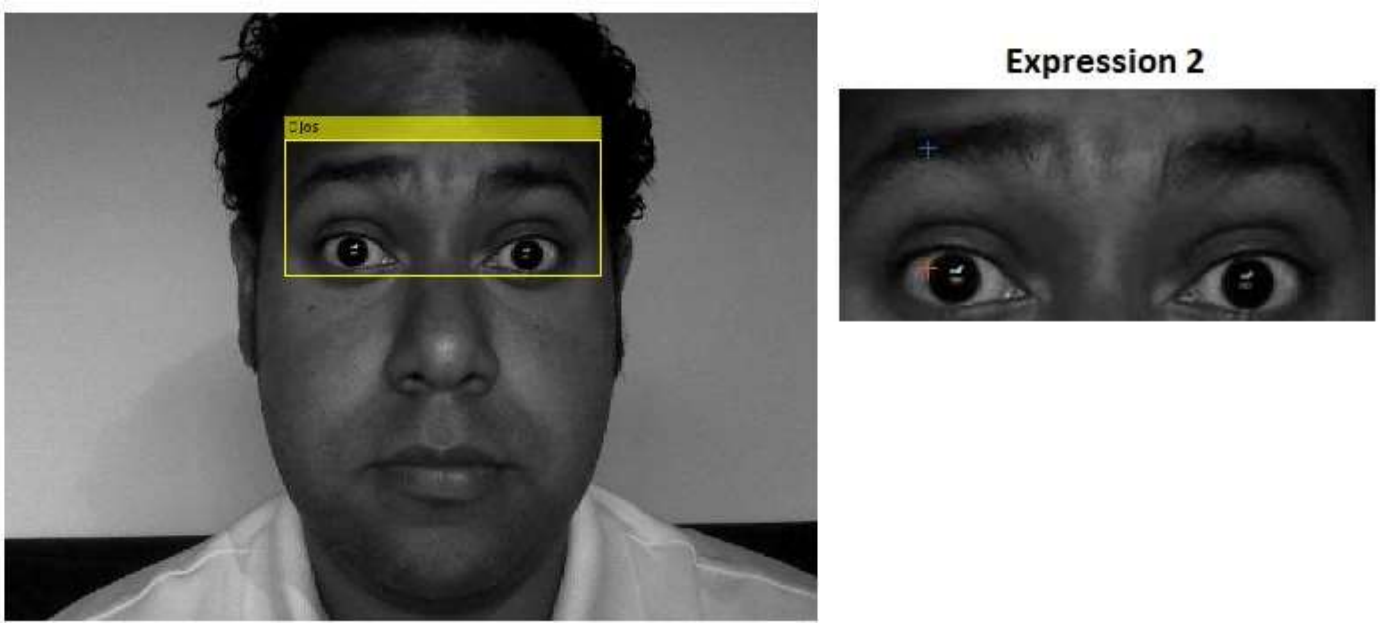

Figure 3. Detection of eyes in expression 2 - Subject 1

Source: Authors

Figure 4. Detection of eyes in expression 2 - Subject 2

Source: Authors

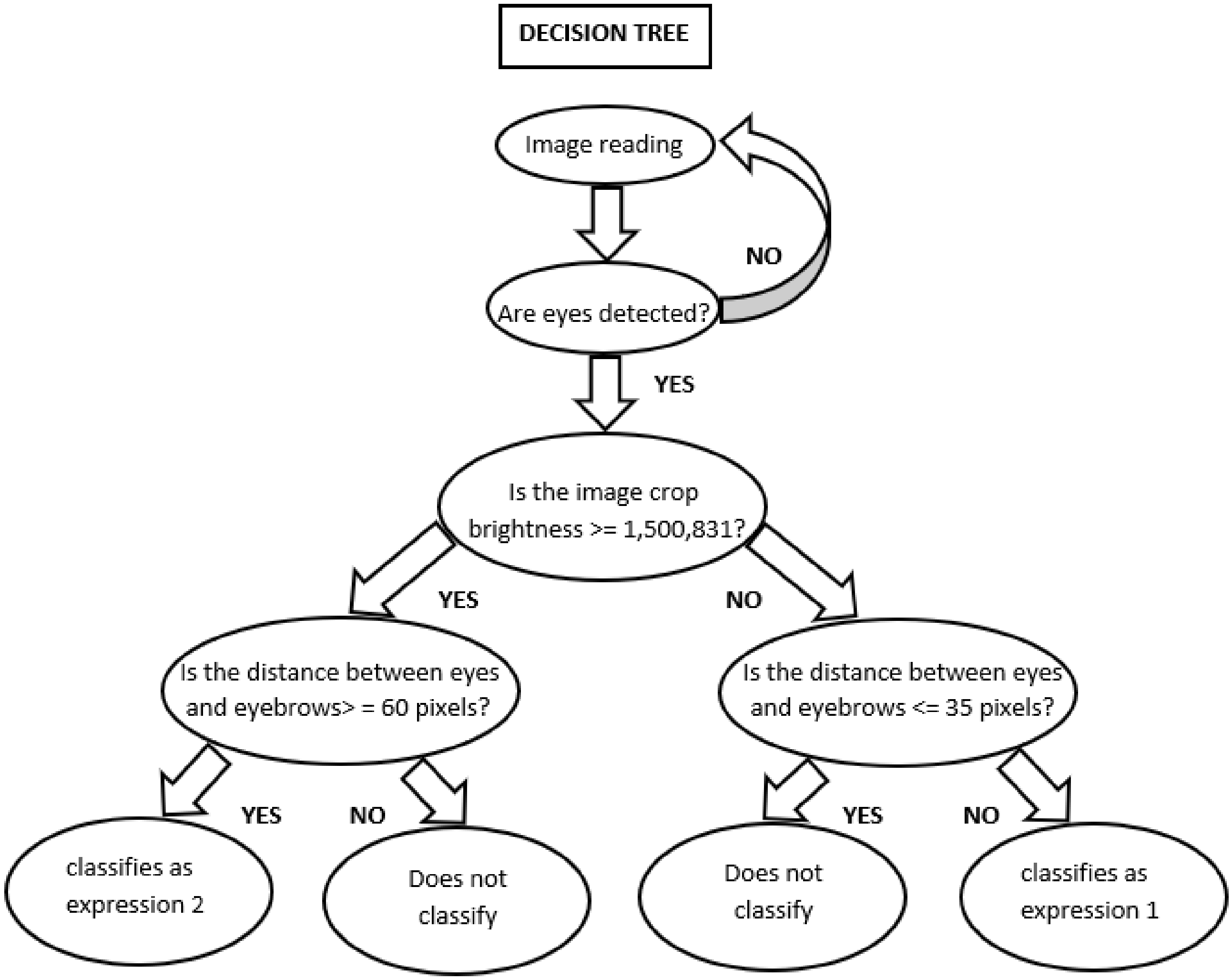

Figure 5. Decision tree implemented

Source: Authors

By generalizing the implemented classifier, a decision tree is designed and implemented to adapt to different lighting conditions. This characteristic may change according to the environment in which the tests are carried out. Figure 6 shows the second implemented decision tree.

2.4 . Experiments

Experiments designed to evaluate the classification system consist of conducting tests in real time. Test subjects are placed in front of the computer on which the system is implemented in Matlab. Tests were performed on 10 subjects who initially presented a system neutral facial expression and then used expressions 1 and 2 to control a consistent image. In these tests, the system was evaluated by acquiring 100 images for each subject, for a total of 1,000 evaluated images.

The evaluations of the system were carried out in a laboratory with dimensions of 5.2 meters x 3.5 meters with a height of 2.2 meters. In addition, 12 bulbs of 2,500 lumens each illuminate the laboratory in the absence of natural light. In the environment a temperature color of 6500 Kelvin is constant during all tests.

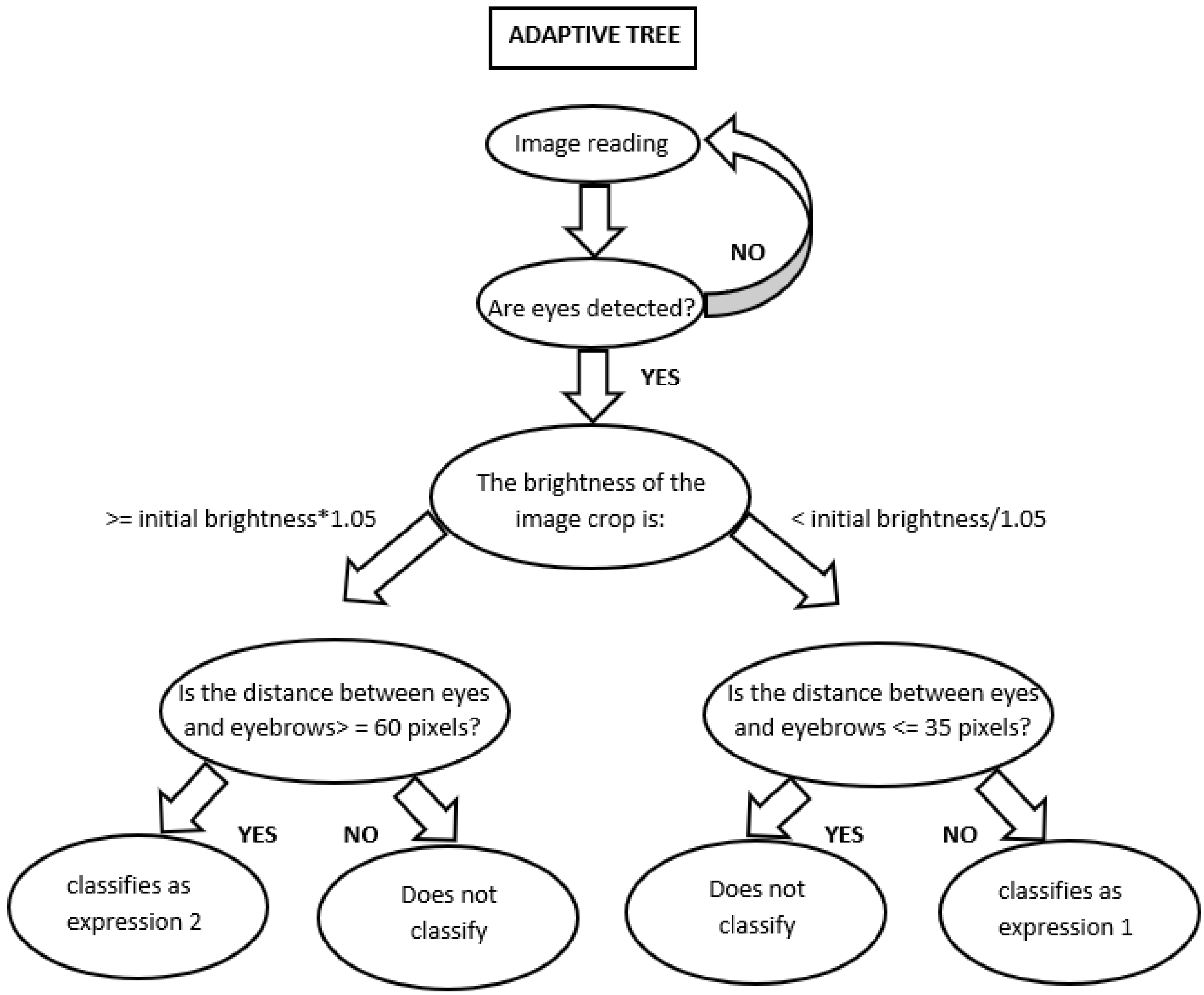

Figure 6. Adaptive decision tree implemented

Source: Authors

III. RESULTS

By extracting the characteristics of classes expression 1 and expression 2, two vectors are obtained. The luminosity vector and the distance vector are graphed and can be seen in Figures 7 and 8. The results obtained by the vector support machine are: Support vectors to classify the images where the eyes are located with respect to luminosity were values entre 921567 and 1500831. To classify the images with respect to the distance between the eyes and the eyebrows were values of 35 and 60 pixels.

These results can be seen in Figure 9. High efficiency (100%) is obtained by evaluating the training of this implemented classifier through cross-validation. To do this, the Matlab kfoldLoss function was implemented.

The first implemented decision tree effectively classified each of the 100 images contained in the database with which we worked in the present system. Efficiency was evaluated by cross validation. The results obtained are shown in Figures 1 to 4.

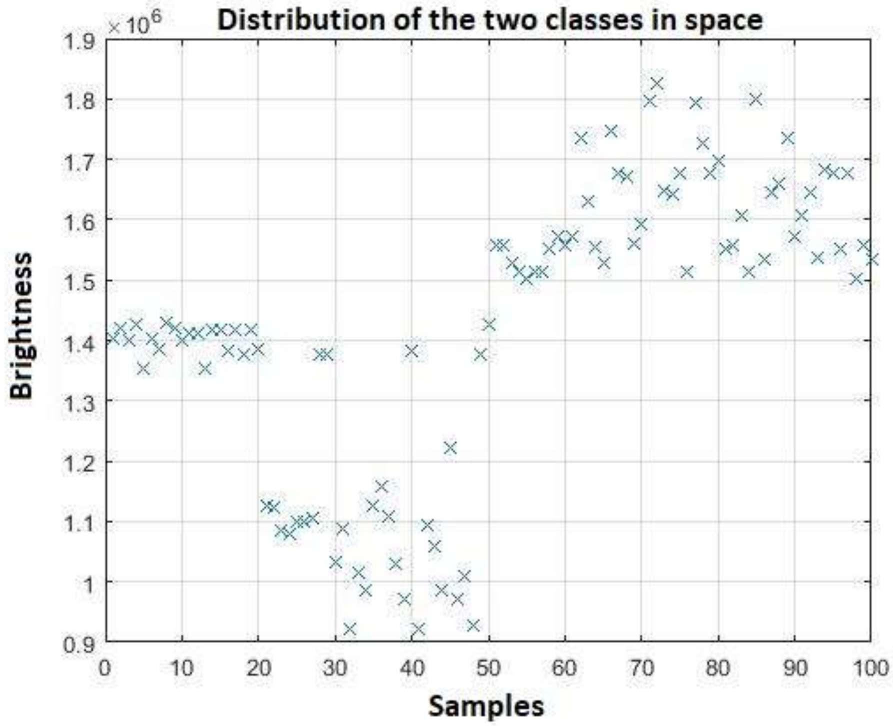

Figure 7. Distribution of the classes in the space according to the luminosity

Source: Authors using Matlab®

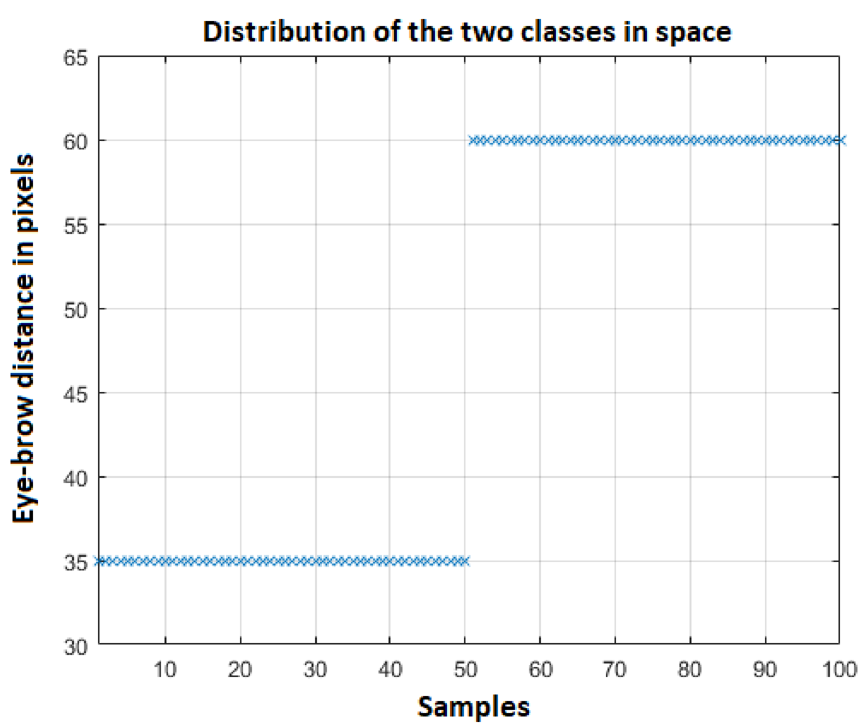

Figure 8. Distribution of the classes in the space according to the distance between the eyebrow

Source: Authors using Matlab®

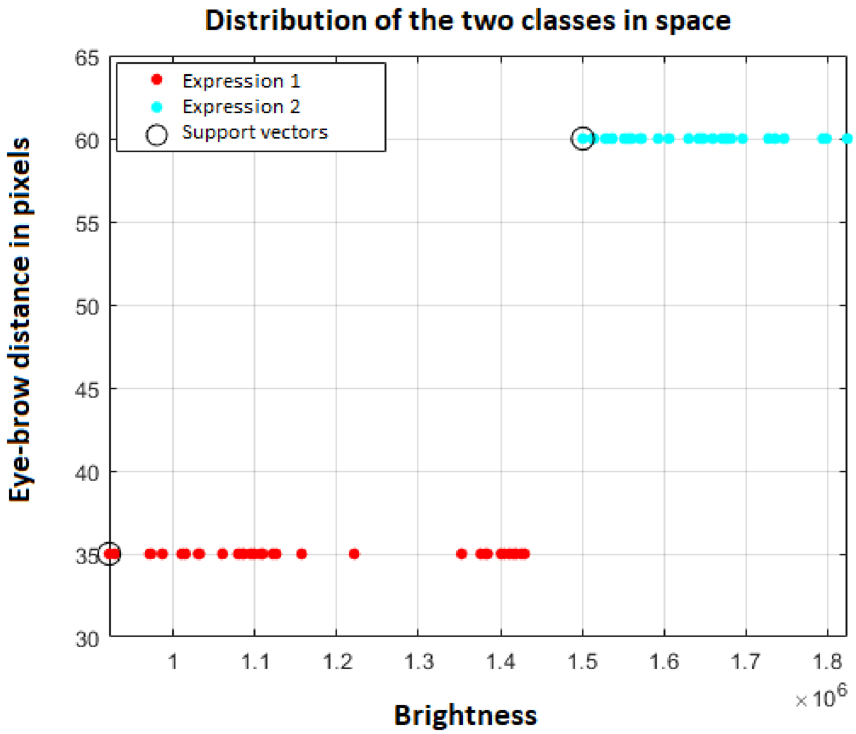

Figure 9. Distribution of classes in space and support vectors

Source: Authors using Matlab®

Finally, the evaluation of the second implemented decision tree was performed, obtaining a yield of 91.4%. This is 914 well-classified images of 1000. The results of one of the tests carried out are shown in Figures 10 to 13, whereby means of the two facial expressions carried out, the zoom of a displayed template is modified.

Table 1 presents a summary of the results obtained with the test subjects when using the facial expression classification system to control the focus of images and manage to read in detail on a computer screen.

The results of the experiments shown in Table 1, 90% of the test subjects indicated that it is possible to practice reading in a text located on the computer.

Subject	Samples	Zoom frequency			Identifies with sharpness
		2x	4x	6x	
1	10	80%	10%	10%	YES
2	10	10%	70%	20%	YES
3	10	0%	90%	10%	YES
4	10	10%	80%	10%	YES
5	10	90%	10%	0%	YES
6	10	80%	10%	10%	YES
7	10	80%	20%	0%	YES
8	10	20%	80%	0%	YES
9	10	80%	10%	10%	YES
10	10	90%	10%	0%	NOT

Table 1. Results of the experiments

Source: Authors

What is a refractive error?

A refractive error is a very common eye disorder. It occurs when the eye cannot clearly focus the images from the outside world. The result of refractive errors is blurred vision, which is sometimes so severe that it causes visual impairment.

The four most common refractive errors are:

- myopia (nearsightedness): difficulty in seeing distant objects clearly;
- hyperopia (farsightedness): difficulty in seeing close objects clearly;
- astigmatism: distorted vision resulting from an irregularly curved cornea, the clear covering of the eyeball.
- presbyopia: which leads to difficulty in reading or seeing at arm's length, it is linked to ageing and occurs almost universally.

Figure 10. Test template presented in original size

Source: https://www.who.int/news-room/q-a-detail/what-is-a-refractive-error shown using Matlab®

Figure 11. Detection of expression 1

Source: Authors using Matlab®

What is a refractive error?

A refractive error is a very common eye disorder. It occurs when the eye cannot clearly focus the images from the outside world. The result of refractive errors is blurred vision, which is sometimes so severe that it causes visual impairment.

The four most common refractive errors are:

- myopia (nearsightedness): difficulty in seeing distant objects clearly;
- hyperopia (farsightedness): difficulty in seeing close objects clearly;
- astigmatism: distorted vision resulting from an irregularly curved cornea, the clear covering of the eyeball.
- presbyopia: which leads to difficulty in reading or seeing at arm's length, it is linked to ageing and occurs almost universally.

Figure 12. Test template increased in size by detection of expression 1

Source: https://www.who.int/news-room/q-a-detail/what-is-a-refractive-error shown using Matlab®

Figure 13. Detection of expression 2

Source: Authors using Matlab®

IV.CONCLUSION

By implementing classification systems such as vector support machines and decision trees, it was possible to recognize facial expressions with performance percentages greater than 91%.

The implemented facial expression classification system allows you to control the focus of images, improving their visualization and sharpness. In addition, it is possible to read in detail on a computer screen mitigating ocular refractive errors.

It was observed that when using cascading classification methods based on the Viola-Jones algorithm combined with classifiers such as vector support machines or decision trees, adjusted results are obtained in the recognition of facial expressions.

The characteristics extracted from the database of acquired images corresponding to luminosity in the eye region and the distance between the eyes and the eyebrows, allowed separable classes to be obtained, which guaranteed fit results at the time of the classification.

By using cascading classification methods based on the Viola-Jones algorithm combined with classifiers such as vector support machines or decision trees, the efficiency of results obtained in the recognition of facial expressions is increased.

REFERENCES

Bourne, R. R., Flaxman, S. R., Braithwaite, T., Cicinelli, M. V., Das, A., Jonas, J. B., ... & Naidoo, K. (2017). Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis. The Lancet Global Health, 5(9), e888-e897.

Dirección General de Educación Especial. (2016). Dirección General de Educación Especial - Secretaría de Educación del Estado de Veracruz. Retrieved Marzo 1, 2016, from Dirección General de Educación Especial - Secretaría de Educación del Estado de Veracruz: http://eespecial.sev.gob.mx/difusion/visual.php

Duda, R. O., Hart, P. E., & Stork, D. G. (2001). Pattern Classification. New York: John Wiley & Sons, Inc.

Hildebrandt, A., Recio, G., Sommer, W., & Wilhelm, O. (2014). Facial EMG Responses to Emotional Expressions Are Related to Emotion Perception Ability. PLOS, 1.

World Health Organization. (2013). Organización Mundial de la Salud. Retrieved Mayo 1, 2016, from http://www.who.int/blindness/AP2014_19_Spanish.pdf

World Health Organization. (2018). Ceguera y discapacidad visual. World Health Organization.

Torres, A. (2006). Atención al educando ciego o con deficiencias visuales. San José, Costa Rica: Universidad Estatal a Distancia.

Viola, P., & Jones, M. J. (2001.). Rapid Object Detection using a Boosted Cascade of Simple Features. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1, 511–518.