

Image Classification Methods Applied in Immersive Environments for Fine Motor Skills Training in Early Education

Andrés Ovidio Restrepo Rodríguez¹, Maddyzeth Ariza Riaño¹, Paulo Alonso Gaona-García¹, Carlos Enrique Montenegro-Marín¹, Íñigo Sarria² *

¹ Universidad Distrital Francisco José de Caldas, Bogotá D.C., Colombia

² Universidad Internacional de La Rioja, Logroño (Spain)

Received 15 September 2019 | Accepted 16 October 2019 | Published 30 October 2019



ABSTRACT

Fine motor skills allow to carry out the execution of crucial tasks in people's daily lives, increasing their independence and self-esteem. Among the alternatives for working these skills, immersive environments are found providing a set of elements arranged to have a haptic experience through gestural control devices. However, generally, these environments do not have a mechanism for evaluation and feedback of the exercise performed, which does not easily identify the objective's fulfillment. For this reason, this study aims to carry out a comparison of image recognition methods such as Convolutional Neural Network (CNN), K-Nearest Neighbor (K-NN), Support Vector Machine (SVM) and Decision Tree (DT), for the purpose of performing an evaluation and feedback of exercises. The assessment of the techniques is carried out using images captured from an immersive environment, calculating metrics such as confusion matrix, cross validation and classification report. As a result of this process, it was obtained that the CNN model has a better supported performance in 82.5% accuracy, showing an increase of 23.5% compared to SVM, 30% compared to K-NN and 25% compared to DT. Finally, it is concluded that in order to implement a method of evaluation and feedback in an immersive environment for academic training in the first school years, a low margin of error must be taken in the percentage of successes of the image recognition technique implemented, to ensure the proper development of these skills considering their great importance in childhood.

KEYWORDS

Immersive Environment, Augmented Reality, Image Recognition Techniques, Support Vector Machine, K-nearest Neighbor, Decision Tree, Convolutional Neural Network.

DOI: 10.9781/ijimai.2019.10.004

I. INTRODUCTION

MUCH of the cognitive activities, such as writing, reading or speaking, also include the use of fine motor skills, involving hand movement, hand-eye coordination, sound production control skills, etc.; for this reason, it can be said that there is a relationship between cognitive development and the fine motor [1]. In older children and adults, the development of fine motor skills implies learning the set of complex and specific sequences of movement necessary to perform the activities of daily life, traditional writing and using the keyboard, leisure activities and vocational tasks [2]. In studies such as [1], it is determined that attention, fine motor skills and general knowledge help to predict the percentages that students will subsequently obtain in math, reading and science, so these indicators are essential for preparation in early education. In addition, and according to the above,

in [3] the authors perform the identification of brain wave patterns of people with dyslexia during the writing process, stating that writing skills are a relevant academic indicator. Therefore, since there is a need for the use of fine motor skills in cognitive activities, the constant motivation in school classrooms of fine motor practice is important.

Technology represents a great importance nowadays and there has been a great growth in the implementation of new mechanisms and means that allow to involve technology in education, as a way to increase motivation in learning. This is how Augmented Reality (AR) permits the creation of different immersive scenarios through the use of virtual elements on a real environment, facilitating the acquisition of knowledge, meaningful learning and user interest [4]-[5].

There are numerous references that have focused on the creation of AR environments in education and their influence on motivation, as well as promoting motor activities; studies such as [6]-[7] present a 3D navigation structure through the use of mobile devices, based on a Knowledge Organization System (KOS) using school subjects about animals, in a population of children between 10 and 12 years, in order to analyze aspects of learning and usability, through a set of questions related to the system; the results show that, in the test carried out, most participants successfully identified the animals, their relationship with the habitat and the category of the animal, according to the classification defined in the

* Corresponding author.

E-mail addresses: A. O. Restrepo Rodríguez (aorestrepo@correo.udistrital.edu.co), M. Ariza Riaño (marizar@correo.udistrital.edu.co), Paulo Alonso Gaona-García (pagaonag@udistrital.edu.co), Carlos Enrique Montenegro-Marín, (cemontenegrom@udistrital.edu.co), Íñigo Sarria (inigo.sarria@unir.net).

KOS. Authors such as [8] propose a virtual environment based on 3D navigation for gestures with two hands, offering control on the speed of navigation, being easy and intuitive; when evaluating the technique with a set of participants, the study finds its feasibility making it easy to learn and use, in addition to reaching high rates of accuracy, performance and prediction speed. In research such as [9], the planning, assignment and use of digital learning objects with augmented reality is proposed to support the teaching-learning process of higher education students, taking into account the students' learning style, resulting in a better academic performance, in addition to making the learning process more dynamic and attractive to students. For their part, the authors of [10] develop an augmented reality platform to simulate a 3D book versus a traditional book for chemistry learning in high school students, given the low popularity of the subject, resulting in those students who used the platform had better results in their semiannual exams, compared to those who used the book traditionally. In [11] a method of learning in Augmented Reality was carried out through the use of mobile devices, using multidimensional concept maps, and having a standard group and an experimental group for testing; as a result, it is evident that the experimental group performed significantly better than the other group and a greater motivation is concluded during the development of the activity. In a study carried out by [12], a mobile platform is exposed through the use of augmented reality, for descriptive geometry learning; results show a great increase in positive feelings on the part of the students, through the proposed method that is characterized by the use of virtual models of the figures.

Likewise, there are studies related to AR in early education, as well as the influence of technology in increasing motivation and interest during the learning process in children. For their part in [13] they develop an interactive-educational game for learning concepts such as colors, geometric shapes and mathematics through augmented reality, for children between 4 and 7 years old, obtaining as a result a pleasant and fun teaching environment, as well as the increased interest and commitment on the part of the students, who were also evaluated by their parents. In [14] content on animals is implemented for children aged 4 to 5 years, based on augmented reality; as a result, the authors present an increase in the effectiveness of learning and an increase in student activity, as augmented reality is the promoter of active behavior and development of communication skills in children.

Images classification methods in artificial intelligence aim to improve the feature and pattern extraction learning, and their application can be seen in a number of contexts, including education [15]. Methods such as Support Vector Machine (SVM) that is based on the principle of margin maximization [16], K-Nearest Neighbor (K-NN) whose purpose is to find the nearest neighbor between nodes [17], Convolutional Neural Network (CNN) that it is a neural network characterized by using the convolution operation [18] and Decision Tree (DT) defined as a simple representation of a finite set of classes [19], are considered the most commonly used and known traditional approaches [20] [21]. Studies like [22] present the application of a convolutional neural network model for plants recognition as of an images dataset; as a result, 86.2% effectiveness was seen with the CNN model implemented. In the field of education, references such as [15] worked with a classification method and the use of augmented reality by presenting a CNN model for geometric figures classification captured from an AR-Sandbox, with the aim of supporting early education and fine-motor therapy in children. The authors found that the proposed method shows a decrease of 39.45% based on loss and an increase of 14.83% of correct answers, concluding that there is an increase in performance with the selected CNN model.

Taking into account the above, this study aims to apply the classification methods SVM, CNN, K-NN and DT with images captured from the immersive environment presented in [23], a study that addresses an environment based on augmented reality in support of training in

children from 4 to 6 years old, through 3D scenarios where children reinforce the learning of vowels and their training. The application of the methods is done in order to provide feedback on the activities carried out by the children and these will be compared using a confusion matrix, classification report and cross validation as metrics to measure their performance, given that, traditionally, machine learning-based approaches are assessed in a cross validation scenario that validates the classification model by assessing how the result will generalize to an independent dataset [24]. Additionally, the representation of a confusion matrix helps in analyzing the belongingness of each class and makes a clear view of the classification accuracy of each class showing the correct classification and misclassification of each class against the belongingness of each target class [25], in addition, usually classification report is used to check the quality of classification algorithm predictions. The rest of this article is organized as follows: in Section II there are Materials and methods where the general scenario is presented, which is divided into the immersive environment and the development of activities feedback with the application of images classification methods; Section III presents results and discussion; finally, conclusions and future work are presented.

II. MATERIALS AND METHODS

Fig. 1 shows the general scenario with which it will work, structured with two main components: *Immersive environment* and *Activities feedback*. The Immersive environment component is based on an Augmented Reality prototype where a series of 3D scenarios are presented with which it interacts through a gesture control device and whose objective is the practice of vowels for the support of fine-motor therapy in children from 4 to 6 years old. The Activities feedback component complements the first by determining if the child's exercise corresponds to the expected result; this is achieved with the application of the classification methods CNN, SVM, K-NN and DT, taking as a starting point the images obtained with the immersive environment. Finally, in order to measure the performance of the methods, a comparison is made between them by means of confusion matrix, classification report and cross validation as performance metrics, considering that they are techniques that have been addressed in studies such as [20] [21] [26], presenting a comparison between the methods when recognizing images in fields such as the identification of diseases in plants, numbers and elements of the daily life using data sets such as MNIST and CIFAR 10 and finally, identification of types of rice grains.

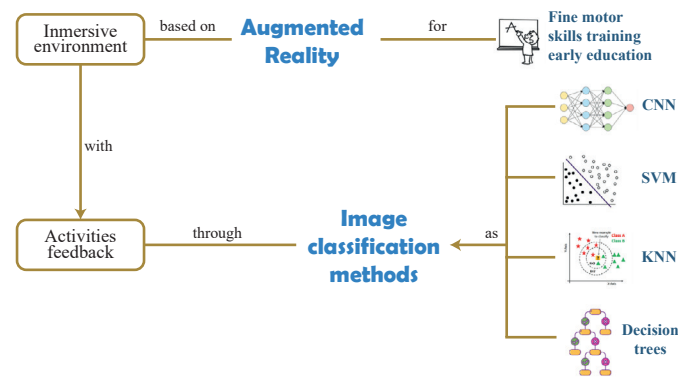


Fig. 1. General scenario. Source: Authors

A. Immersive Environment

The immersive environment used is presented in [23], where the authors developed a set of scenarios based on Unity 3D. There are five modules, one for each vowel and each one consists of scenarios

where multimedia elements, such as videos, related words and images, are presented for the learning of the vowel, and a training scenario for its practice. The interaction with the scenarios is done through the Leap Motion gesture control device which detects and tracks the position of the participant's hands and fingers in space through two monochromatic stereo cameras and three infrared LEDs [27], such as it is shown in Fig. 2. In this way, the user can play the videos, change the scenario and move through the different modules.



Fig. 2. Use of the leap motion device. Source: Leap motion [28].

The training scenario presents a blank board, indicating the vowel to be written, where the child should use their hands to write on the board, through the Leap Motion device (Fig. 3); this allows the practice of the writing of the vowels and at the same time the fine-motor skills are worked by involving the movement of the hands and the hand-eye coordination. From this platform the images of the vowels are obtained by taking a screenshot of the training board.

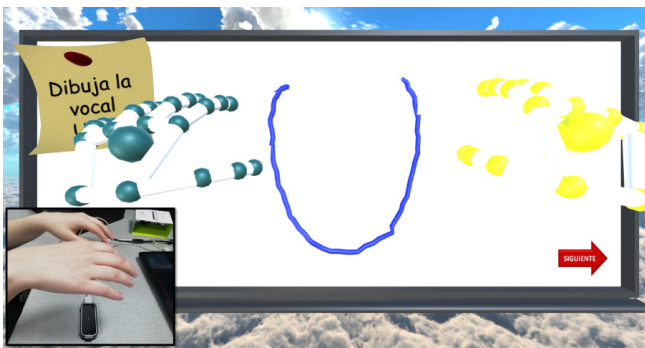


Fig. 3. Vowel writing in the immersive environment. Source: Authors.

B. Activities Feedback

Fig. 4 presents the approach to carry out the development of the module that corresponds to Activities Feedback. This component has as its starting point the training of machine learning models such as Support Vector Machine (SVM), K-Nearest Neighbor (K-NN) and Decision Tree (DT), using the Grid Search method, in order to perform hyperparameter optimization and obtain the best model of each technique according to the vowel training dataset in accordance with the nature and operation of the immersive environment presented previously. Additionally, an already defined model of convolutional neural network (CNN) will be trained, which has been taken from studies such as [15], exposing as a result a base model of a CNN, with the objective of recognizing images acquired from an immersive environment called AR-Sandbox, presenting an accuracy in prediction phase of 0.87. Once the models have been trained and selected, the evaluation and comparison of the models is carried out, using metrics such as confusion matrices, a classification report made up of precision score, f1 score and recall score and, finally, performing cross-validation.

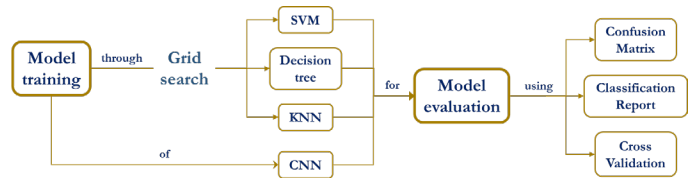


Fig. 4. General approach for Activities Feedback component. Source: Authors.

In order to perform the training of the SVM, K-NN, DT, and CNN models, a training dataset made up of 5 classes corresponding to the vowels (A, E, I, O, U) was used. The dataset is determined by a total of 1600 images, where, each class has a total of 320 images, taken from the internet or made in some basic drawing tool. Fig. 5 presents the training set used.

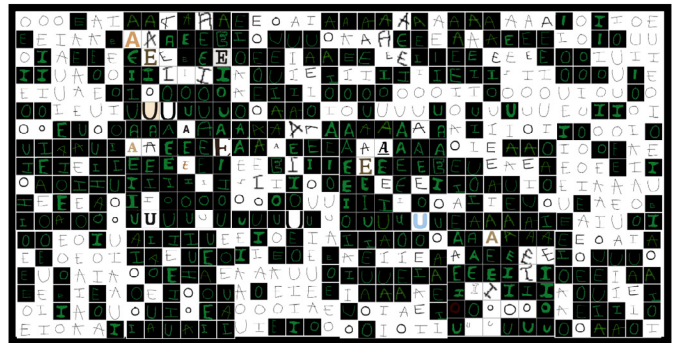


Fig. 5. Training dataset. Source: Authors

1. CNN Model

As mentioned above, the selection of the convolutional neural network model was based on a study carried out by [15], where the definition of a base model of a CNN was carried out, performing hyperparameter optimization by Random Search, defining a hyperparameters dictionary. These parameters are taken randomly in the base model, with the purpose of training a total of 50 models, to make the selection of a final model, having as decision criteria the accuracy, the average square error and the loss function, metrics presented in this study. The selected model had the purpose of recognizing and classifying geometric figures acquired directly from an immersive environment such as an AR-Sandbox. The model is presented in Fig. 6.

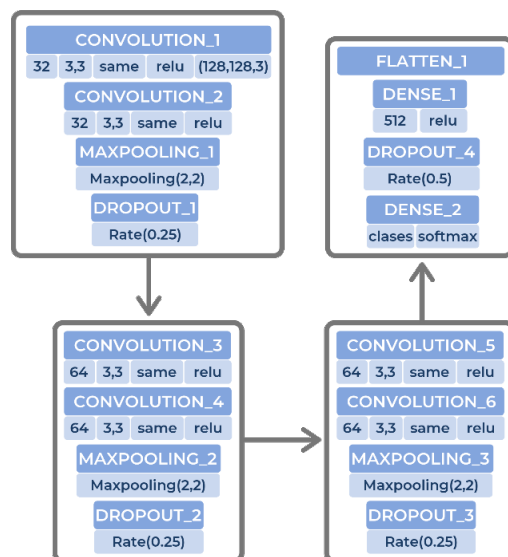


Fig. 6. CNN model. Source: Restrepo, Casas et al., 2019 [15]

When the base model is defined, we proceed to train it with the training dataset described above, obtaining an average of accuracy of 0.9758 and a loss function of 0.0874. Table I, presents the defined compilation and training hyperparameters.

TABLE I. COMPILATION AND TRAINING HYPERPARAMETERS

Hyperparameter	Value
Optimizer	RMSProp
Loss Function	Cross Entropy
Metrics	Accuracy
Batch_size	32
Epochs	10

Source: Authors

2. SVM Model

In order to obtain the SVM model by optimizing hyperparameters using Grid Search, a hyperparameters dictionary is defined, declaring variable parameters such as, the kernel that will use the algorithm and the penalty value for the error. Table II presents the defined hyperparameters dictionary.

TABLE II. SVM HYPERPARAMETER DICTIONARY

Hyperparameter	Possible Values
Kernel	'linear', 'poly', 'rbf', 'sigmoid'
Penalty parameter	1, 10, 100, 100, 1000

Source: Authors

As a result, a total of 20 models were trained, taking the accuracy value as the decision criterion, the model with the best performance had as hyperparameters a 'poly' type kernel and a *penalty value for the error* of 1.0, obtaining an average accuracy of 0.595. Fig. 7, presents the model scheme with its hyperparameters.

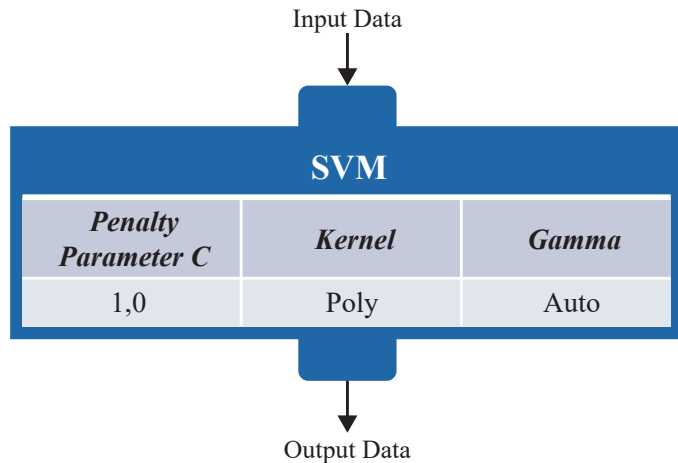


Fig. 7. SVM scheme. Source: Author

3. K-NN Model

On the other hand, in order to select the K-NN model using Grid Search, Table III presents the hyperparameters dictionary defined, declaring variable the number of neighbors and the weight function used in the prediction phase.

TABLE III. K-NN HYPERPARAMETER DICTIONARY

Hyperparameter	Possible Values
Number of neighbors	2, 3, 5, 15, 30, 60, 100
Weights function	'uniform', 'distance'

Source: Authors.

When defining the hyperparameters dictionary, a total of 14 models were trained, calculating their average accuracy score, considering it as a decision criterion. The model with the highest precision value was obtained by using **3 neighbors** and a *distance* type weight function, obtained 0.625 as the score value. Fig. 8, presents the model scheme with its hyperparameters.

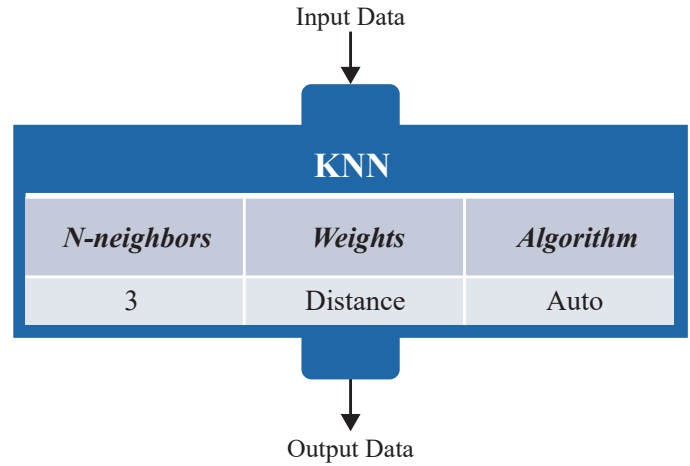


Fig. 8. K-NN scheme. Source: Authors.

4. DT model

Finishing the training of machine learning models, the decision tree model is selected using Grid Search, which is why a hyperparameters dictionary is defined, leaving the criteria to measure the quality of a division and the strategy to choose the division in each of the nodes. Table IV exposes the defined hyperparameters dictionary.

TABLE IV. DECISION TREE HYPERPARAMETER DICTIONARY

Hyperparameter	Possible Values
Criterion	'gini', 'entropy'
Splitter	'best', 'random'

Source: Authors.

According to the hyperparameters dictionary, only a total of 4 models were trained, where the best combination occurs when the criterion for measuring the quality of a division is made using the *Gini* function and the *best* function is used as a division strategy in each node. This model had an average accuracy value of 0.61. Fig. 9, shows the model scheme with its hyperparameters.

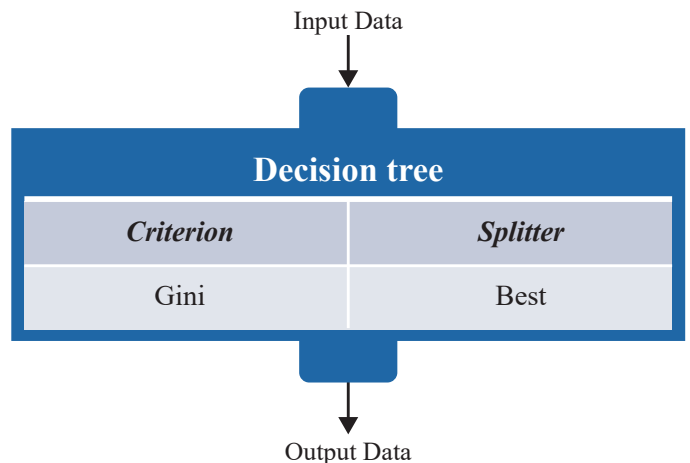


Fig. 9. Decision Tree scheme. Source: Authors.



Fig. 10. Test dataset. Source: Authors.

III. RESULTS AND DISCUSSIONS

In order to perform the evaluation and comparison of the models defined above, a total of 14 children were cited, who interacted with the immersive environment and made 40 images for each class, that is, 40 images for each of the vowels, forming a set of test data of a total of 200 images. Fig. 10 exposes the test dataset.

On the other hand, the models were evaluated by means of confusion matrices, a classification report, which is composed of metrics such as f1 score, recall score and precision score and, finally, by cross-validation, the above with the purpose of obtaining approximately the accuracy value both specifically by vowel and generally from each of the models, and after that, making a comparison between each of the techniques implemented.

A. Confusion Matrix

In order to obtain the confusion matrices of each model with the test dataset, the *sk-learn* library was used, which was implemented in Python.

1. CNN Model

Table V, presents the number of successes in the prediction phase with the CNN model with a total of 40 samples per class, where a success rate of 82.5% was obtained.

TABLE V. NUMBER OF HITS IN CNN

	A	E	I	O	U
A	36	2	0	1	1
E	3	35	1	1	0
I	2	6	30	0	2
O	4	3	0	31	2
U	0	0	3	4	33

Source: Authors.

2. SVM Model

Table VI exposes the number of hits in the prediction phase with the SVM model with a total of 40 samples per class, where a 59% hit rate was obtained.

TABLE VI. NUMBER OF HITS IN SVM

	A	E	I	O	U
A	20	4	5	9	2
E	1	24	6	5	4
I	5	2	19	9	5
O	3	3	0	34	0
U	3	5	1	10	21

Source: Authors.

3. K-NN Model

Table VII shows the number of hits in the prediction phase with the K-NN model with a total of 40 samples per class, where a 52.5% hit rate was obtained

TABLE VII. NUMBER OF HITS IN K-NN

	A	E	I	O	U
A	7	1	4	28	0
E	2	18	6	12	2
I	0	0	24	14	2
O	0	0	1	37	2
U	0	0	5	16	19

Source: Authors.

4. DT Model

Table VIII, presents the number of hits in the prediction phase with the DT model with a total of 40 samples for each class, where a 57.5% hit rate was obtained.

TABLE VIII. NUMBER OF HITS IN DT

	A	E	I	O	U
A	21	7	7	4	1
E	8	18	4	6	4
I	6	4	21	5	4
O	3	7	0	28	2
U	1	2	6	4	27

Source: Authors.

Based on the confusion matrices presented above, it can be seen that the model that shows the best performance is CNN, since it has a total of 82.5% success rate in its predictions, exposing an increase of 23.5% compared to SVM, 30% against K-NN and 25% against DT. A specific analysis of the confusion matrices is presented in Table IX, where H is the percentage of correct answers and F the percentage of errors.

As shown in Table IX, it can be evidenced that when performing a specific analysis of each of the vowels, the CNN model significantly overcomes the rest of the models, except in the O vowel, given that the SVM gets over it by a 0.75% and K-NN by 1.5%.

B. Classification Report

With the purpose of obtaining metrics such as: precision score, recall score and f1 score, the *sk-learn* library was used, through its classification report module, which includes the metrics mentioned above. Table X presents the average of these metrics for each of the trained models.

TABLE IX. CONFUSION MATRIX ANALYSIS

Method	Vowel									
	A		E		I		O		U	
	H	F	H	F	H	F	H	F	H	F
CNN	0.9	0.1	0.875	0.125	0.75	0.25	0.775	0.225	0.825	0.175
SVM	0.5	0.5	0.6	0.4	0.475	0.525	0.85	0.15	0.525	0.475
K-NN	0.175	0.825	0.45	0.55	0.6	0.4	0.925	0.075	0.4	0.6
DT	0.525	0.475	0.45	0.55	0.525	0.475	0.7	0.3	0.675	0.325

Source: Authors.

TABLE X. AVERAGE METRICS FOR CLASSIFICATION REPORT

Model	Precision_score	Recall_score	F1_score
CNN	0.830	0.826	0.824
SVM	0.610	0.586	0.586
K-NN	0.700	0.516	0.510
DT	0.584	0.594	0.582

Source: Authors.

Additionally, Fig. 11. Classification report values. Source: Authors.11 graphically represents the consolidated data in Table X. As of Fig. 11. Classification report values. Source: Authors.11, it can be seen that the CNN model presents a higher level of performance compared to the rest of the techniques, in addition, this means that for the SVM, K-NN and DT models there are a greater number of false positives, therefore, the values in these metrics tend to be lower 24.13% on average compared to the CNN model.

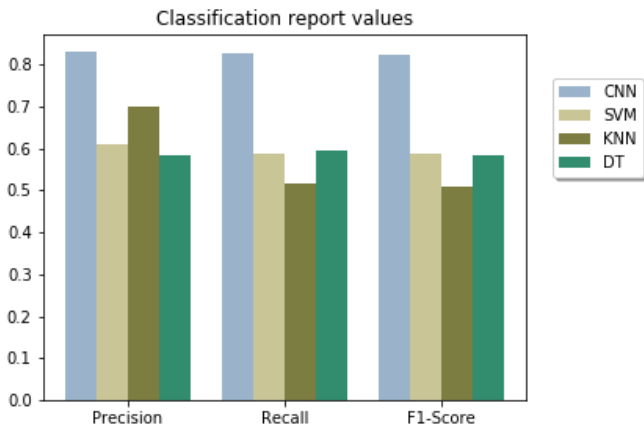


Fig. 11. Classification report values. Source: Authors.

C. Cross Validation

In order to carry out the cross-validation process, use was made of both the training dataset and the test dataset, that is 1800 images. Additionally, a total of 10 randomly divisions were established, using functions such as *KFold* and *cross_val_score* of the *model_selection* module of the *sk-learn* library. Table XI, exposes the average of the cross-validation score and its standard deviation for each of the models and for the number of divisions defined.

TABLE XI. SCORE AND STANDARD DEVIATION IN CROSS VALIDATION

Model	Average Cross Validation Score	Standard Deviation
CNN	0.8651	0.0451
SVM	0.5306	0.0485
K-NN	0.5137	0.0255
DT	0.5638	0.0457

Source: Authors

As of Fig. 12. Model comparison by cross validation. Source: Authors.12, it can be seen that when applying cross-validation and calculating its score, the CNN model shows an increase of 0.3345 against SVM, 0.3514 against K-NN and 0.3013 against DT, on average it increases by 32.9%, in addition, this model exposes a high score above 0.85. On the other hand, from Fig. 12. Model comparison by cross validation. Source: Authors.12, which presents the consolidated data in Table XI through a box diagram, it can be stated that, as a first instance, the median of the CNN model is located approximately 0.25 points above the median of the other models. In addition, it has a relatively close maximum value of 0.9 and a minimum value approximately close to 0.8, therefore its values do not vary significantly, that is, its precision value calculated when applying cross validation remains stable.

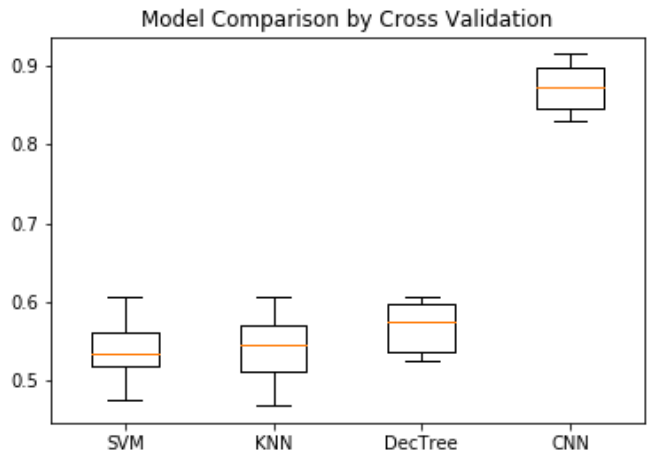


Fig. 12. Model comparison by cross validation. Source: Authors.

IV. CONCLUSIONS

Based on the results obtained in the evaluation of the confusion matrix metrics, precision score, recall score, F1 score and cross validation, it is possible to conclude that the CNN model implemented has a better performance compared to traditional methods such as K-NN, SVM and DT, even obtaining the best models of these techniques through hyperparameter optimization. Methods such as K-NN and SVM presented a greater number of successes in the prediction of the O vowel, compared to the CNN method, when analyzing the confusion matrix; nevertheless, this may imply that these methods created a predisposition to the vowel in its phase of extracting features and training.

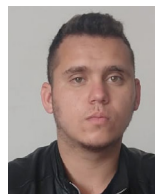
Given the importance of the educational and motor training of children in the first school years, it is necessary to have a high level of percentage of successes in the model implemented to carry out the evaluation and feedback of the exercises performed in an immersive environment that supports the development of these skills.

Although the recognition of images associated with vowels, geometric figures, among others, has been presented in a large number of previous studies, it is important to consider that the immersive environment used involves a variety of multimedia elements, which can generate noise. Despite the above, the CNN model presented a high level of precision in its prediction stage; however, this value must be refined and increased, perhaps as the authors of [15] do, using an image cleaning technique.

The immersive environment presented in this study was taken from the work carried out by [24], where the authors have as their main objective the reinforcement of fine motor skills through the writing of vowels; however, since there is no way to evaluate the exercise, it is not guaranteed that the children have fulfilled the objective. That is why when obtaining a method with a high level of success to carry out feedback on the exercises performed in immersive environments such as the one mentioned above, the design and implementation of a monitoring and recommendation system for students is considered as future work, being applied from its early stages of education which can be decisive for its formation. Additionally, it is possible to include new scenarios by implementing different exercises by broadening the model recognition spectrum.

REFERENCES

- [1] Grissmer, D., Grimm, K. J., Steele, J. S., Aiyer, S. M., and Murrah, W. M. (2010). Fine Motor Skills and Early Comprehension of the World: Two New School Readiness Indicators. *Developmental Psychology*, 46(5), pp. 1008–1017. doi: 10.1037/a0020104.
- [2] Case-Smith, J. (2013). *Encyclopedia of Autism Spectrum Disorders*. New York, United States: Volkmar, Fred R.
- [3] Perera H, Shiratuddin MF, Wong KW and Fullarton K. (2018) EEG Signal Analysis of Writing and Typing between Adults with Dyslexia and Normal Controls. *International Journal of Interactive Multimedia and Artificial Intelligence*, 5(1), pp. 62-67. doi: 10.9781/ijimai.2018.04.005.
- [4] Che Hashim, N., Abd Majid, N. A., Arshad, H. and Khalid Obeidy, W. (2018). User Satisfaction for an Augmented Reality Application to Support Productive Vocabulary Using Speech Recognition. *Advances in Multimedia*, 2018, pp. 1-10. doi: 10.1155/2018/9753979.
- [5] Sveistrup, H. (2004). Motor rehabilitation using virtual reality. *Journal of NeuroEngineering and Rehabilitation*, 1(10), pp. 1-8. doi: 10.1186/1743-0003-1-10.
- [6] Garcia Gaona, P. A., Martin Moncunill, D., Gordillo, K. and González Crespo, R. (2016). Navigation and Visualization of Knowledge Organization Systems using Virtual Reality Glasses. *IEEE Latin America Transactions*, 14(6), pp. 2915-2920. doi: 10.1109/TLA.2016.7555275.
- [7] Garcia Gaona, P. A., Martin Moncunill, D., Gordillo, K. and Montenegro, C. (2016, June). "Navigation and Visualization of Knowledge Organization Systems using Virtual Reality Glasses: first insights". Presented in *UNESCO-UNIR ICT & Education Latam 2016*, Bogotá, Colombia.
- [8] Rehman, I., Ullah, S. and Raees, M. (2019). Two Hand Gesture Based 3D Navigation in Virtual Environments. *International Journal of Interactive Multimedia and Artificial Intelligence*, 5(4), pp. 128-140. doi: 10.9781/ijimai.2018.07.001.
- [9] Medina, M. A., García, C. F. J. and Olguín, M. J. A. (2018). Planning and Allocation of Digital Learning Objects with Augmented Reality to Higher Education Students According to the VARK Model. *International Journal of Interactive Multimedia and Artificial Intelligence*, 5(2), pp. 53-57. doi: 10.9781/ijimai.2018.02.005
- [10] Pasaréti, O, Hajdú, H., Matuszka, T., Jámbori, A., Molnár, I. and Turcsányi-Szabó, M. (2012). Augmented Reality in education. pp. 1-10.
- [11] Chen, C.-H., Huang, C.-Y. and Chou, Y.-Y. (2019). Effects of augmented reality-based multidimensional concept maps on students' learning achievement, motivation and acceptance. *Universal Access in the Information Society*, 18(2), pp. 257–268. doi: 10.1007/s10209-017-0595-z.
- [12] Gutiérrez de Ravé, E., Jiménez-Hornero, F. J., Ariza-Villaverde, A. B. and Taguas-Ruiz, J. (2016). DiedricAR: a mobile augmented reality system designed for the ubiquitous descriptive geometry learning. *Multimedia Tools and Applications*, 75(16), pp. 9641–9663. doi: 10.1007/s11042-016-3384-4.
- [13] Zhu, Y., Yang, X. and Wang, S. J. (2017). Augmented Reality Meets Tangibility: A New Approach for Early Childhood Education. *EAI Endorsed Transactions on Creative Technologies*, 4(11), pp. 1-8. doi: 10.4108/eai.5-9-2017.153059.
- [14] Cascales, A., Laguna, I., Pérez-López, D., Perona, P. and Contero, M. (2013, July). "An Experience on Natural Sciences Augmented Reality Contents for Preschoolers". Presented in *International Conference on Virtual, Augmented and Mixed Reality VAMR 2013: Virtual, Augmented and Mixed Reality. Systems and Applications*, Las Vegas, NV, USA.
- [15] Restrepo Rodríguez, A. O., Casas Mateus, D. E., Gaona García, P. A., Montenegro Marín, C. E. and González Crespo, R. (2018). 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(743), pp. 1-16. doi: 10.3390/sym10120743.
- [16] Adankon, M. and Cheriet, M. (2009). *Encyclopedia of Biometrics*. Boston, MA: Li S.Z., Jain A.
- [17] Ault, Aaron & Zhong, Xuan & Coyle, Edward. (2005). K-nearest-neighbor analysis of received signal strength distance estimation across environments.
- [18] Ketkar, N. (2017). *Deep Learning with Python*. Apress, Berkeley, CA: Springer, pp. 63-78.
- [19] Ganten, D. (2006). *Encyclopedic Reference of Genomics and Proteomics in Molecular Medicine*. Berlin, Heidelberg: Springer.
- [20] Lin, P., Li, X. L., Chen, Y. M. and He, Y. (2018). A Deep Convolutional Neural Network Architecture for Boosting Image Discrimination Accuracy of Rice Species. *Food and Bioprocess Technology*, 11(4), pp. 765–773. doi: 10.1007/s11947-017-2050-9.
- [21] Wang, Y. and Xin, M. (2019). Research on image classification model based on deep convolution neural network. *EURASIP Journal on Image and Video Processing*, 2019(40). doi: 10.1186/s13640-019-0417-8.
- [22] Dyrmann, M., Karstoft, H. and SkovMidtby, H. Plant species classification using deep convolutional neural network. *Biosystems Engineering*, (151), pp. 72-80. doi: 10.1016/j.biosystemseng.2016.08.024.
- [23] Ariza Riaño, M., Restrepo Rodríguez, A. O., Gaona García, P. and Montenegro Marín, C. (2018, november). "3D Immersive Environments to Support Training Activities Through Fine Motor Therapy". Presented in *11th annual International Conference of Education, Research and Innovation*, Seville, Spain. doi: 10.21125/iceri.2018.1675.
- [24] Allix, K., Bissyandé, T. F., Jérôme, Q., Klein, J., State, R. and Le Traon, Y. (2016). Empirical assessment of machine learning-based malware detectors for Android. *Empirical Software Engineering*, 21(1), pp. 183–211. doi: 10.1007/s10664-014-9352-6.
- [25] Bisoi, R., Dash, P. K. and Das, P. P. (2018). Short-term electricity price forecasting and classification in smart grids using optimized multikernel extreme learning machine. *Neural Computing and Applications*, pp. 1-24. doi: 10.1007/s00521-018-3652-5.
- [26] Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D. and Stefanovic, D. (2016). Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification. *Computational Intelligence and Neuroscience*, 2016, pp. 1-11. doi: 10.1155/2016/3289801.
- [27] Tung, J. Y., Lulic, T., Gonzalez, D. A., Tran, J., Dickerson, C. R. and Roy, E. A. (2015). Evaluation of a portable markerless finger position capture device: accuracy of the Leap Motion controller in healthy adults. *Physiological measurement*, 36(5), pp. 1025-1035. doi: 10.1088/0967-3334/36/5/1025.
- [28] Ultrahaptics Ltd. (2019). *Leap Motion*. In: <https://www.leapmotion.com/>.



Andrés Ovidio Restrepo-Rodríguez

Is an active member of Interactive Multimedia and Digital Animation research group at Engineering Faculty of Universidad Distrital Francisco José de Caldas, Bogotá, Colombia. He finished a degree on System Engineer on 2019 and currently he is doing an MSc in information science and Communication at Universidad Distrital Francisco José de Caldas. His research interests include Artificial Intelligence, Data Mining, Heterogenous Computing, Virtual Environments, Learning Analytics.



Maddyzeth Ariza-Riaño

She is a Systems Engineer and a master's student on Information Science and Communication at Engineering Faculty of Universidad Distrital Francisco José de Caldas, Bogotá, Colombia. Her research interests include Data Mining, Software Architecture and Semantic Web.



Paulo Alonso Gaona-García

Is Associate Professor and active member of GIIRA research group at Engineering Faculty of Universidad Distrital Francisco José de Caldas, Bogotá – Colombia. He obtained his Ph.D. in Information of Engineering and Knowledge at University of Alcalá in 2014. He finished a degree on System Engineer on 2003 and obtained an MSc in Information Science and Communication in 2007 at Universidad Distrital Francisco José de Caldas. His research interest includes Web science, network and communications, information security, e-learning, information visualisation and semantic Web.



Carlos Enrique Montenegro-Marin

Is a Ph.D. in systems and computer services for internet from University of Oviedo, Asturias, Spain (2012). He has a Diploma of advanced studies 2008 of the Pontifical University of Salamanca. He is MSc. Science in Information and Communication Systems from the Universidad Distrital Francisco José de Caldas. He is a System engineer. His research interests include Object-Oriented technology, Language Processors, Modeling Software with, DSL and MDA.



Íñigo Sarría

Íñigo Sarría holds a Doctorate (Cum Laude and special prize) from the International University of La Rioja, a Licentiate degree in Mathematics from the University of the Basque Country, a University Expert qualification in Analysis of the Knowledge Society from The International University of La Rioja, and a Certificate in Teaching from the Complutense University of Madrid. He is Director of the Computer Science and Technology Area of the Higher School of Engineering and Technology at the International University of La Rioja and is the Coordinator of the Master's in Mathematical Engineering and Computing. ORCID code: <https://orcid.org/0000-0002-2584-9671>