# An Approach for Visual Attention Classification Based on Particle Swarm Optimization Using Eye Tracker Tests

Guillermo E. Medina Magdaleno, Marco A. Aceves Fernández and Dr. Jesús C. Pedraza Ortega

***Abstract*—Attention classification has been widely studied over the last decade, with methodologies and proposals for various purposes such as the early detection of autistic spectrum disorder, Attention Deficit Hyperactivity Disorder (ADHD) or just to have a reliable tool to determine whether a subject is being attentive or not. This work proposes a methodology for visual attention classification based on particle swarm optimization using eye tracker data. Firstly, the data was obtained through a series of visual tests applied to a certain number of adult subjects while an eye tracker acquires the eye coordinates. Then, the data was processed to extract the desired features to build the final dataset. To optimize the model, a particle swarm optimization with the K- Means algorithm was performed to generate the optimum groups for classification with KNN. Finally, the performance evaluation and comparison with other works from state of the art were carried out. The proposed methodology reached an accuracy of 97.78\% without using expensive or cumbersome equipment. Therefore, a reliable and comfortable tool for assessing visual attention was achieved using the proposed methodology.**

***Index Terms*— Classification, visual attention, particle swarm optimization, PSO.**

## I. Introduction

THE visual attention phenomenon has been studied widely for the last century. The first studies were technologically limited to simple observation or even introspection. Von Helmholtz observed the natural tendency of visual attention to deviate or stray away to new things, he was especially focused with eye movements to special locations or the "where" of visual attention [1]. In the ‘80’s, Posner stated that the visual attention was linked or could be compared with a "spotlight" or "lamp", referring to a limited space focused in the center. Other authors like Eriksen and Yeh make the same comparison but call it a "zoom" or other authors like Downing and Pinker to a Gaussian gradient [2].

This work was supported in part by the Mexican Government via the Scientific Secretariat (CONACYT).

G. E. Medina Magdaleno, Facultad de Ingeniería, Universidad Autónoma de Querétaro, 76010, Querétaro, México. (e-mail: guillermo.emedina1@gmail.com).

Dr. M. A. A. Fernández, Facultad de Ingeniería, Universidad Autónoma de Querétaro, 76010, Querétaro, México. (e-mail: marco.aceves@uaq.mx).

Dr. J. C. P. Ortega, Facultad de Ingeniería, Universidad Autónoma de Querétaro, 76010, Querétaro, México. (e-mail: caryoko@yahoo.com).

During the last decade, methodologies and proposals to classify attention for various goals have been developed, from the early detection of autistic spectrum disorder [3], attention deficit hyperactivity disorder [4], or simply to develop a reliable tool to determine whether a subject is being attentive or not [5].

However, some of these methods could be very invasive or challenging, like EEG application to determine attention levels [6] or ADHD classification in which noisy or high dimensional signals are present [7]. Using an eye tracker as a non-invasive method may be challenging, although it’s a viable option in terms of practicality and cost-effectiveness, taking advantage of the valuable data that can be acquired through gaze study and using that information adequately [8].

There are studies both recent and made in the last decade regarding eye- tracking and methodologies or algorithms based on artificial intelligence to classify attention or ASD using machine learning techniques such as random trees, random forests or genetic algorithms with support vector machines [4-5-25].

Furthermore, studies from the last decade, have demonstrated that swarm intelligence-based methodologies used for optimization, feature extraction or just to boost performance in general [9-13] yields satisfactory results, even when compared to standard machine learning- based methodologies [14], which suggests that using a swarm intelligence- based methodology could bring good results to the problem of classification with eye tracking data.

This work proposes a swarm intelligence-based methodology to classify attention levels via visual experiments using an eye tracker. The paper structure is as follows, in section II will be presented the theoretical background involving visual attention, eye tracking, and swarm intelligence; section III will include materials and methods; section IV will present the results and finally, section V Finish with some remarks as a conclusion for this work.

## II. Background

### A. Visual Attention

In the context of visual attention, there are two main functions of the human eye: tracking and fixation. fixation occurs when the eye is fixed into a particular visual objective. This allows the eyes to maximize the focus on the object.

Tracking is the ability to stay focused on an object even when it is moving. This is important since most objects in the real world are moving. Without the ability to track, it would be very difficult to perceive anything. The visual attention mechanism must have at least one of these basic components [15]:

- The selection of a region of interest in the visual field.
- The selection of the characteristic dimensions and values of interest.
- The control of information flow through the neuron network that constitutes the visual system.
- The ability to jump from one selected region to another in time

### *B. Eye Tracking*

The device used to measure eye movement is called eye-tracker. Generally, there are two types of eye-tracking techniques, one that measures the eye position relative to the head, and another that measures eye orientation in a space or point of consideration [16]. Nowadays, eye- tracking technology applications range from video games and virtual reality to web publicity and research [17].

### *C. Swarm Intelligence*

Swarm intelligence algorithms are meta-heuristic algorithms that imitate the social behavior of insect colonies, where each agent presents a potential solution to a given problem. During each solution cycle, they change their positions and move within a domain with the objective of finding a better solution [18].

### *D. Particle Swarm Optimization*

PSO is an optimization algorithm that simulates the behavior of bird flocks. This works by randomly initializing a set of particles (in this context, a bird is a "particle", which is also a potential solution to a problem) over a given search space. The swarm moves towards a global best position at a certain velocity, with each iteration, the velocity for each particle changes based on the particle's momentum, the best position reached by the particle, and the best position of all the particles at the current stage, then, based on the velocity, the position for each particle changes. The position and velocity of the i th particle in the i th iteration is displayed in the vector $X = \left(x_{i1}^t, x_{i2}^t, x_{in}^t\right)$ and $V = \left(v_{i1}^t, v_{i2}^t, v_{in}^t\right)$ the personal best position of a particle is in the vector $P = \left(p_{i1}^t, p_{i2}^t, p_{in}^t\right)$ and finally, the global best with the vector $G = \left(g_{i1}^t, g_{i2}^t, gx_{in}^t\right)$. Finally, with every iteration, the velocity and position of a particle is updated using equations 1 and 2 as follows [17]:

$$v_i^{(t)} = \omega * V_i^{(t-1)} + c_1 * r_1\left(P_i - X_i^{(t-1)}\right) + c_2 * r_2\left(G_i - X_I^{(t-1)}\right) \quad (1)$$

$$X_i^{(t)} = X_i^{(t-1)} + V_i^{(t)} \quad (2)$$

Where $\omega$ is known as inertia weight and controls the impact of previous velocities of the particle on the current one, $r_1$ and $r_2$ are two variables raging from [0, 1], $c_1$ and $c_2$ are called acceleration constants, and are positive values that control the step size between iterations.

### *E. PSO-K Means Algorithm*

The algorithm used in this work is a hybrid between PSO and K-means for data clustering. This algorithm consists in seeding the initial swarm with K-means first, where K-means is terminated when certain criteria are met, the number of iterations is finished, or when the average change in centroid vectors is less than a user- set parameter. The result is then used as one of the particles, while the rest are generated randomly [19].

### *F. Clustering and K-Means*

Clustering techniques refer to organizing a certain group of objects that share similar characteristics, it's classified under the unsupervised machine learning methods, meaning there's no previous training from which to learn. The main idea in the K-Means algorithm is to define groups or clusters of data. Firsthand, a selected number of clusters is required, each of the clusters has a centroid, a point where the distance of the objects is to be calculated. The clusters are defined by an iterative process in the distances of objects closer to the centroid. To know which centroid is assigned to each object, the algorithm uses a measure of Euclidean distance. The sum of the squares is calculated with the square root of the Euclidean distances for each centroid of each cluster, the one with the smaller value is the cluster to which a certain object is assigned [20]. The K-Means algorithm can be summarized as follows [17]:

- Initialize randomly the k cluster centroid where $Z = \left(z_{1,} z_2, z_k\right)$
- Until a termination criteria is met, assign a cluster centroid to each data point that is closer to it, the distance of certain data point $y_p$ to the centroid in d-dimensional space is given as:
- $$D\left(y_p, z_j\right) = \sqrt{\sum_{i=1}^{d}\left(y_{pi} - z_{ji}\right)^2} \quad (3)$$
- Recalculate the centroids, the centroid for cluster j is determined by:
- $$z_j = \frac{1}{n_j}\sum_{\forall y_p \in c_j} y_p \quad (4)$$

  Where $c_j$ is the subset of data points belonging to the j-th cluster, and $n_j$ is the number of data points in the cluster.
- The algorithm then terminates when the number of iterations reaches certain parameter, when there if no change in cluster membership or when the cluster membership change is negligible.

### *G. KNN Classifier*

The k-nearest neighbor classifier or KNN is a classification algorithm to classify unlabeled data based on the label of the nearest observations to a given point. There are two important parameters to take into account when using KNN, one is the calculated distance between an observation and its nearest neighbors, generally being a Euclidean distance as follows:

$$D(p,q)=\sqrt{(p_1-q_1)^2+(p_2-q_2)^2(p_n-q_n)^2} \quad (5)$$

where $p$ and $q$ are observations to be compared with n characteristics. And the other parameter is the "K" parameter, which determines how many neighbors with a given label are needed to classify the objective observation [21].

## III. Materials and Methods

### A. Data Acquisition and Construction

For the proposed model, the data was gathered through a series of visual tests that were applied to 41 male and female adult subjects, which registered the gaze coordinates using an eye tracker (Eye Tribe ET1000) with a sampling frequency of 60hz. There was a total of 11 visual tests that were applied to each subject. The experimental set-up consists of 6 D-48 domino tests and 6 Kohs cube tests, shown in figures 1 and 2, respectively. In these tests, the subject must observe the patterns presented and choose the correct answer.

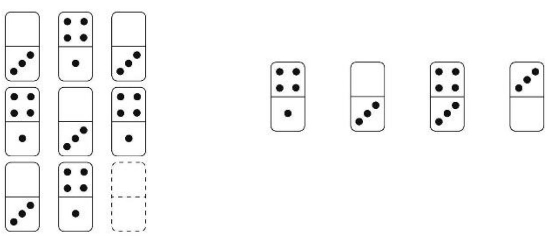

Fig. 1. A D-48 test.

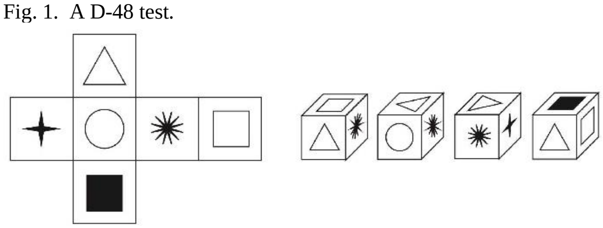

Fig. 2. A Kohs block test

Once the subjects carried out the visual tests, the coordinates of the eye tracker measurement were obtained as shown in the figure 3, from which the desired features ought to be extracted using the images yielded by the eye tracker containing the gaze points (Figure 4). One of the features was the percentage of gaze points within an area of interest (AOI) for each visual test. The other feature was the percentage of fixations for each of the visual tests, fixations being an eye movement where the gaze is fixed on a point [22]. And finally, the correct answers for each visual test. This is shown in figure 5.

| | timestamp | fix | avgx | avgy |
|---|---|---|---|---|
| 108 | 42:39.4 | False | 991.9171 | 713.4952 |
| 109 | 42:39.5 | False | 0 | 0 |
| 110 | 42:40.0 | False | 493.5139 | 525.2659 |
| 111 | 42:40.2 | False | 303.2414 | 411.4729 |
| 112 | 42:40.2 | False | 300.7155 | 411.318 |
| 113 | 42:40.3 | False | 245.6798 | 404.1431 |
| 114 | 42:40.5 | True | 263.472 | 437.4751 |
| 115 | 42:40.5 | True | 262.7594 | 450.1548 |
| 116 | 42:40.6 | True | 256.3541 | 453.4707 |
| 117 | 42:40.8 | False | 395.2192 | 581.2395 |
| 118 | 42:40.9 | False | 430.002 | 609.7447 |
| 119 | 42:41.0 | False | 426.0832 | 617.7189 |
| 120 | 42:41.1 | False | 388.7195 | 591.4795 |

Fig. 3. Raw data obtained from the eye tracker.

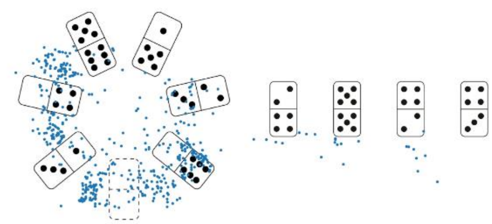

Fig. 4. Raw data image obtained from the Eye Tracker.

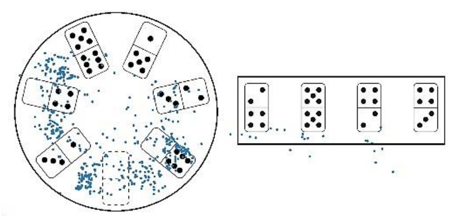

Fig. 5. Example of a region of interest for the domino test.

### B. Model Implementation

The stages of the model implementation are as follows (figure 6): Pre-processing, where the invalid columns are neglected, data normalization is performed, and finally dimensionality reduction using Principal Component Analysis (PCA). The following stage is, the label generation and clustering, where the elbow method is applied to obtain the optimum number of clusters to be generated as shown in equation 6.

$$J=\sum_{i=1}^{k}\sum_{x\in c_1}\left|x-c_1\right|^2 \quad (6)$$

Where $J$ is the cost function, $x$ is the element of the cluster $c_i$, and $k$ is the number of clusters. In this case, it was found that the optimum number of clusters was 3 as shown in figure 7.

Furthermore, the PSO-K Means algorithm is calculated to generate the clusters so the classifier algorithm can be performed using KNN. Finally, the evaluation of the model takes place using K- fold cross-validation. Comparisons with previous similar works from the state of the art take place to evaluate the performance of this algorithm.

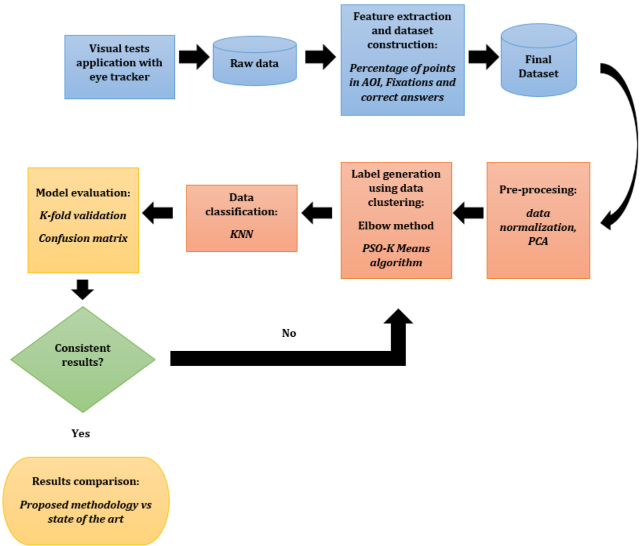


Fig. 6. Flow chart of the proposed methodology.

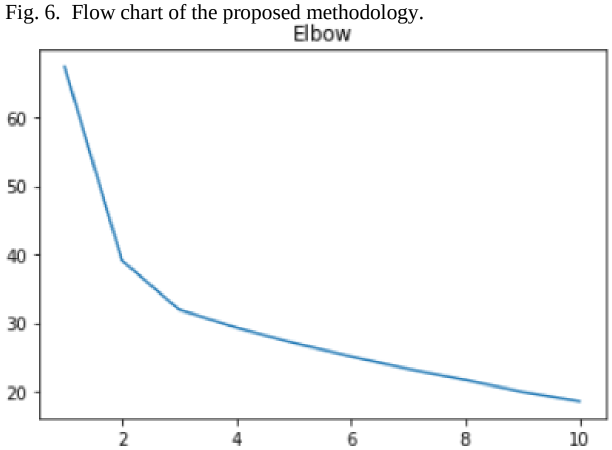


Fig. 7. Definition of the number of clusters using the elbow method.

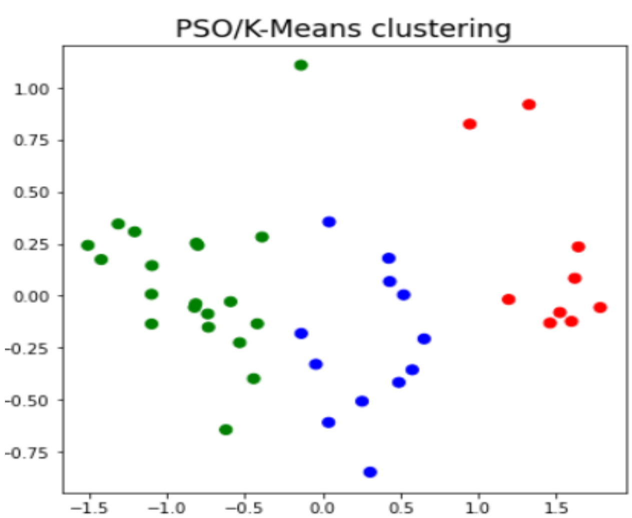


Fig. 8. PSO/K-Means generated clusters.

## IV. Results

In this section, the results of the proposed model are presented. It is important to mention that the tests were applied to 41 adult subjects of various ages and genders. The classification was done with a KNN classifier using a K of 7. Once the classification of the three generated labels was complete, the model was evaluated using a K fold cross validation with a K value of 5, ensuring that 20% of the data is used for testing and the remaining 80% for training in each fold. After the K fold implementation, the average accuracy of every fold was taken. Finally, the results are presented in comparison with similar works to see how well our methodology performs compared to other works.

In the following table, the results of the proposed model are presented. The following metrics were obtained from the cross-validation and the average of each metric's average is presented and the average accuracy obtained. The metrics used were the macro and weighted precision, recall, and F1 score as presented in Table 1.

TABLE I
Results Using the Proposed Methodology

| Value | Average value for K-fold with K=5 | | |
|---|---|---|---|
| | Precision | Recall | F1 Score |
| Macro Average | 98.67% | 97.78% | 97.92% |
| Variance | 0.07% | 0.19% | 0.17% |
| Weighted Average | 98.22% | 97.78% | 97.68% |
| Variance | 0.08% | 0.20% | 0.18% |
| | Accuracy = 97.78% | | |
| | Variance = 0.20% | | |

Table 1 shows the average results of the proposed methodology using K-foldcross-validation. In this order of ideas, ten trials were carried out afterwards, in which the data in the training and testing sets were set randomly after each trial to ensure that the model is not biased or overfit the results. Figure 9, figure 10 and Table 2 shows the results for these trials, which shows the metrics for accuracy, precision, recall, and F1-score.

TABLE II
Results Using the Proposed Methodology with 10 Trial Test

| Value | Average value per 10 trials | | |
|---|---|---|---|
| | Precision | Recall | F1 Score |
| Macro Average | 96.38% | 97.67% | 96.40% |
| Variance | 34.03% | 14.43% | 31.49% |
| Weighted Average | 95.79% | 94.44% | 94.25% |
| Variance | 48.55% | 80.24% | 85.80% |
| | Accuracy = 96.67% | | |
| | Variance = 25.92% | | |

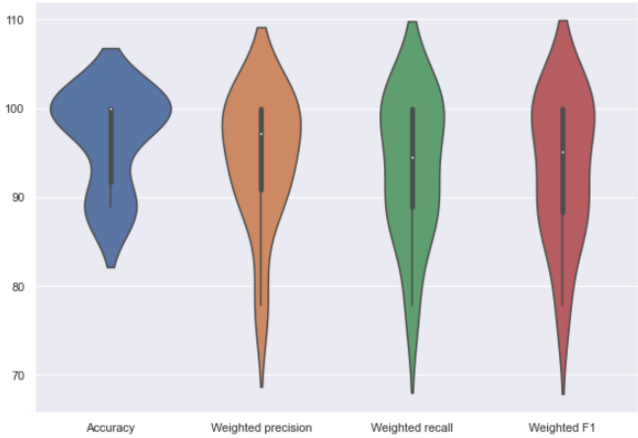


Fig. 9. Weighted values of the 10 trial test.

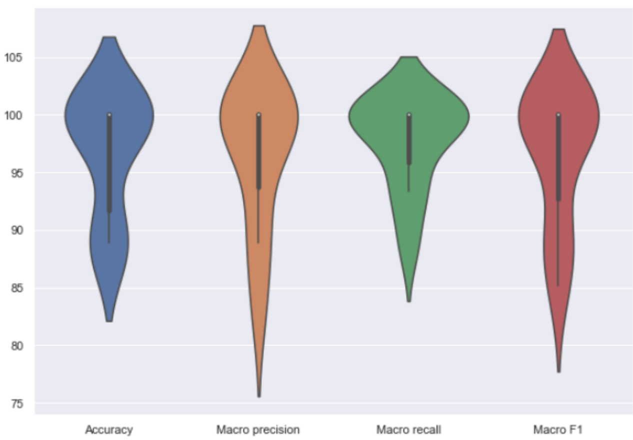


Fig. 10. Macro values of the 10 trial test.

The average accuracy was of 96.67% this time, slightly lower than in the previous k fold test but staying above 95%. The rest of the metrics were also lightly different from the previous test while at the same time little difference is observed between them.

In the previous works from the state of the art, the sensors used for data acquisition to explore attention statuses range from EEG to eye tracker data. The following table compares the results from the previous works to the proposed one. The proposed methodology outperforms the previous works with 97.78% as shown in Table 3.

TABLE III
RESULTS COMPARISON

| Author | *Accuracy* | Comparison *Data Acquisition* | *Method proposed* |
|---|---|---|---|
| De Silva *et. al*. (2019) | 85.31% | Eye Tracker | Random Forest |
| Alirezaei *et. al*. (2017) | 92.8% | EEG | C-SVM |
| Chen *et. al.* (2017) | 93.1% | Eye Tracker | GA-SVM |
| Proposed work | 97.78% | Eye Tracker | KMeans/PSO-KNN |

## V. CONCLUSION

In this work, a proposed methodology to classify attention levels via eye tracker obtained data from a relatively affordable eye tracker is developed as a means to classify attention without the need of cumbersome or expensive equipment required normally to achieve this task while keeping a consistent performance level of 97.78% in accuracy and a variance no greater than 0.20%. The results obtained and shown in this work have proven to be consistent through the various tests performed and even outperformed previous works from the state of the art. This was achieved through a methodology based on PSO to optimize the label generation and clustering stage crucial for the upcoming classification task. In the end, a consistent and accurate methodology to classify attention levels without the need for highly expensive or uncomfortable equipment that will not have the need for the subject or patient to be wearing any kind of accessory while taking the visual tests makes our system easy to use for both the operating personnel and the subjects taking the test. As future work, it may be important to determine the extent to which a cost-effective eye tracker may be used for all types of tests that are not necessarily attention dominant.

## REFERENCES


[1] A. T. Duchowski, Eye tracking methodology. Cham: *Springer International Publishing*, 2017.

[2] M. Carrasco, "Visual attention: the past 25 years," Vision Res., vol. 51, no. 13, pp. 1484–1525, 2011.

[3] Mestas Hernández, L., Gordillo León, F., Castillo Parra, G., & Escotto Córdova, E., "Diagnóstico y tratamiento de los trastornos del espectro autista: Eye-traking e integración sensorial," EduPsykhé: Revista de Psicología y Psicopedagogía, vol. 1, no. 15, pp. 93–110, 2016.

[4] S. De Silva et al., "A rule-based system for ADHD identification using eye movement data," *in 2019 Moratuwa Engineering Research Conference (MERCon)*, 2019.

[5] O. T. C. Chen, P. C. Chen, and Y. T. Tsai, "Attention estimation system via smart glasses," *in 2017 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, 2017.

[6] C.-M. Chen, J. Y. Wang, and C. M. Yu, "Assessing the attention levels of students by using a novel attention aware system based on brainwave signals: Novel attention aware system based on brainwave signals," Br. *J. Educ. Technol*., vol. 48, no. 2, pp. 348–369, 2017.

[7] M. A. Aceves Fernandez, "Methodology proposal of ADHD classification of children based on cross recurrence plots," Nonlinear Dyn., vol. 104, no. 2, pp. 1491–1505, 2021.

[8] R. J. Leigh and D. S. Zee, Neurology of eye movements. F. A. Davis Company, 1983.

[9] V. Agrawal and S. Chandra, "Feature selection using Artificial Bee Colony algorithm for medical image classification," *in 2015 Eighth International Conference on Contemporary Computing (IC3)*, 2015.

[10] W. C. Kao, C.-Y. Lee, C. Y. Lin, T. Y. Su, B. Y. Ke, and C. Y. Liao, "Gaze tracking with particle swarm optimization," *in 2015 International Symposium on Consumer Electronics (ISCE)*, 2015.

[11] B. Ordóñez De León, M. A. Aceves Fernandez, S. M. Fernandez Fraga, J. M. Ramos Arreguín, and E. Gorrostieta Hurtado, "An improved particle swarm optimization (PSO): method to enhance modeling of airborne particulate matter (PM10)," Evol. syst., 2019.

[12] S. M. Fernandez Fraga, M. A. Aceves Fernandez, J. C. Pedraza Ortega, S. Tovar Arriaga, "Feature Extraction of EEG Signal upon BCI Systems Based on Steady-State Visual Evoked Potentials Using the Ant Colony Optimization Algorithm", Discrete Dynamics in Nature and Society, vol. 2018, Article ID 2143873, 19 pages, 2018. https://doi.org/10.1155/2018/2143873.

[13] Martinez Zeron, E., Aceves Fernandez, M. A., Gorrostieta Hurtado, E., Sotomayor Olmedo, A., & Ramos Arreguín, J. M. (2014). Method to

improve airborne pollution forecasting by using ant colony optimization and neuro-fuzzy algorithms. *International Journal of Intelligence Science, 4(04), 81.S. P. Bingulac*, "On the compatibility of adaptive controllers (Published Conference Proceedings style)," in Proc. 4th Annu. Allerton Conf. Circuits and Systems Theory, New York, 1994, pp. 8–16.

[14] M. Etehadi Abari, "A Novel Eye Gaze Estimation Method Using Ant Colony Optimizer," *Journal of Computer & Robotics*, vol. 1, no. 12, pp. 113–122, 2019.

[15] F. Francis and A. Suresh, "A cognitive model for analyzing visual attention using ocular movements," *in 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS)*, 2017.

[16] A. Kar and P. Corcoran, "A review and analysis of eye-gaze estimation systems, algorithms and performance evaluation methods in consumer platforms," *IEEE Access*, vol. 5, pp. 16495–16519, 2017.

[17] M. P. Saka, E. Doğan, and I. Aydogdu, "Analysis of swarm intelligence–based algorithms for constrained optimization," in Swarm Intelligence and Bio-Inspired Computation, Elsevier, 2013, pp. 25–48.

[18] A. Ahmadyfard and H. Modares, "Combining PSO and k-means to enhance data clustering," in 2008 *International Symposium on Telecommunications*, 2008.

[19] D. W. van der Merwe and A. P. Engelbrecht, "Data clustering using particle swarm optimization," in The 2003 Congress on Evolutionary Computation, 2003. CEC '03, 2004.

[20] J. Bell, Machine learning: Hands-on for developers and technical professionals, 2nd ed. Nashville, TN: John Wiley & Sons, 2020.

[21] Z. Zhang, "Introduction to machine learning: k-nearest neighbors," Ann. Transl. Med., vol. 4, no. 11, p. 218, 2016.

[22] L. Larsson, M. Nyström, R. Andersson, and M. Stridh, "Detection of fixations and smooth pursuit movements in high-speed eye-tracking data," Biomed. Signal Process. Control, vol. 18, pp. 145–152, 2015.

[23] Alirezaei, M., & Sardouie, S. H., "Detection of human attention using EEG signals," *in 24th National and 2nd International Iranian Conference on Biomedical Engineering (ICBME),* 2017, pp. 1–5.

[24] M. A. Aceves Fernandez, "Methodology proposal of ADHD classification of children based on cross recurrence plots," Nonlinear Dyn., vol. 104, no. 2, pp. 1491–1505, 2021.

[25] Z. Zhao, H. Tang, X. Zhang, X. Qu, X. Hu, and J. Lu, "Classification of children with autism and typical development using eye-tracking data from face-to-face conversations: Machine learning model development and performance evaluation," *J. Med. Internet Res*., vol. 23, no. 8, p. e29328, 2021.