Mathematical Biosciences and Engineering

http://www.aimspress.com/journal/MBE

MBE, 20(2): 3216–3236.
DOI: 10.3934/mbe.2023151
Received: 13 August 2022
Revised: 18 November 2022
Accepted: 21 November 2022
Published: 02 December 2022

Research article

Applying machine learning techniques to detect the deployment of spatial working memory from the spiking activity of MT neurons

Gayathri Vivekanandhan[1], Mahtab Mehrabbeik[2], Karthikeyan Rajagopal[3,4], Sajad Jafari[2,5], Stephen G. Lomber[6] and Yaser Merrikhi[6,*]

[1] Centre for Artificial Intelligence, Chennai Institute of Technology, India
[2] Department of Biomedical Engineering, Amirkabir University of Technology (Tehran Polytechnic), Iran
[3] Centre for Nonlinear Systems, Chennai Institute of Technology, India
[4] Department of Electronics and Communications Engineering and University Centre of Research & Development, Chandigarh University, Mohali 140413, Punjab
[5] Health Technology Research Institute, Amirkabir University of Technology (Tehran Polytechnic), Iran
[6] Department of Physiology, Faculty of Medicine, McGill University, Montreal, H3G 1Y6, Canada

* **Correspondence:** Email: yaser.merrikhiahangarkolaee@mcgill.ca.

Abstract: Neural signatures of working memory have been frequently identified in the spiking activity of different brain areas. However, some studies reported no memory-related change in the spiking activity of the middle temporal (MT) area in the visual cortex. However, recently it was shown that the content of working memory is reflected as an increase in the dimensionality of the average spiking activity of the MT neurons. This study aimed to find the features that can reveal memory-related changes with the help of machine-learning algorithms. In this regard, different linear and nonlinear features were obtained from the neuronal spiking activity during the presence and absence of working memory. To select the optimum features, the Genetic algorithm, Particle Swarm Optimization, and Ant Colony Optimization methods were employed. The classification was performed using the Support Vector Machine (SVM) and the K-Nearest Neighbor (KNN) classifiers. Our results suggest that the deployment of spatial working memory can be perfectly detected from spiking patterns of MT neurons with an accuracy of 99.65 ± 0.12 using the KNN and 99.50 ± 0.26 using the SVM classifiers.

Keywords: working memory; middle temporal area; firing rate signal; machine-learning algorithm;

classification

1. Introduction

Machine-learning algorithms are well-known approaches with a wide range of applications in biomedical signal processing, such as classification [1], regression [2], and optimization [3]. In medical applications, diagnosis, recognition, or prediction are the most important tasks that can be handled using classification techniques [4]. In literature, a wide range of applications can be found based on the classification techniques such as emotion recognition [5], text classification [6], activity recognition [7], and epileptic seizure classification [8]. In most classification techniques, the classification performance strongly depends on the input features of the classification algorithms. Hence, feature extraction, the process of summarizing the data into some indexes, and feature selection, the process of finding the optimum combination of the extracted features, play important roles in classification approaches [9].

Among the diverse features that can be extracted from any data, fractal-based features are known as powerful nonlinear tools for measuring data complexity [10]. For example, it was shown that an individual's eye movements are closely related to fractal patterns [11]. Also, it was revealed that the fractality of the electroencephalography (EEG) signals is significantly reduced in schizophrenia [12]. Furthermore, it found that the memory content could increase the fractality of the EEG signals [13]. Another class of powerful features is the features derived from transformed data. For instance, the features based on the time-frequency transforms strongly help to predict atrial fibrillation [14], diagnosis of heart failure [15], and detection of mitral valve prolapse [16]. Such transforms are sometimes not directly used for feature extraction; however, they can be beneficial in the data preparation processes prior to feature extraction [17].

The study of the brain's cognitive functions, such as working memory and over and covert visual attention, are interesting areas of research. Thus, many studies have been conducted to investigate how these top-down functions influence neural spiking activity in different areas of the brain [18–20]. Nonetheless, different studies claimed that no memory-related modulation could be found in the neural spiking activity of the middle temporal (MT) cortex [21–23]. On the other hand, a recent study showed that working memory increased the fractionality of the firing rate signals [24]. Many studies have shown the neural correlate of working memory as an increase in the firing rate of neurons in the prefrontal [19,25,26], parietal [27], and visual [25] cortices; however, no such memory-related change in the average spiking activity of neurons has been reported in extrastriate cortex including V4 [19] and the middle temporal (MT) cortex [19,25] even after applying machine learning techniques [25]. In this paper, we will focus on the maintenance of spatial information, which has been shown to be sent directly from frontal eye field to the extrastriate cortex through feedback connections in form of persistent spiking activity [19]. Although spatial working memory does not increase the average firing rate of extrastriate neurons, it significantly enhances the sensitivity of individual neurons to incoming visual stimuli [26], alters the correlated activity of the population of neurons [18], increases the power of the local field potential (LFP) in the frequency band of alpha-beta, enhances the spike-phase coherency of MT responses in the same frequency range [28] and increases the fractionality of MT spiking activity [24].

As mentioned above, a recent study revealed that when the spiking activity of neurons is mapped to the fractal dimension feature space, the content of working memory can be captured [24]. In this study, we examined whether a set of linear and/or nonlinear features could reveal the deployment of spatial working memory from the spiking activity of neurons in the area MT. In this regard, we used two different learners, three feature selection algorithms, and two cross-validation methods to show the robustness of our results. The remaining parts of the paper are arranged as follows: Section 2 describes the studied data, extracted features, selection methods, classification algorithms, and classification assessment criteria. Section 3 presents the results, and Section 4 discusses the results and concludes the paper.

2. Materials and methods

2.1. Data description

All experimental procedures were performed under the National Institutes of Health Guide for the Care and Use of Laboratory Animals, the Society for Neuroscience Guidelines and Policies. The protocols for all experimental, surgical, and behavioral procedures were approved by the Montana State University Institutional Animal Care and Use Committee.

In this study, the spiking activity of 131 neurons (stored at 32 kHz), recorded in 11 sessions using electrode arrays, was used. These signals were recorded from the area MT (Figure 1a) of the two male macaque monkeys' brains (five and seven years old). The monkeys, already acquainted with carrying out the memory-guided saccade (MGS) task, were positioned on a customized chair in front of a monitor (24 inches with 144 Hz refresh rate) at a distance of 28.5 cm from their eyes. Initially, during a surgery in which the monkeys were anesthetized, the recording chambers were mounted on the monkeys' skulls in the MT area. The recordings of single electrodes were used to confirm that the chambers are mounted in a desired (MT) area of the monkeys' brains, During the task, the monkeys' heads were restrained, and they received juice as a reward through a syringe pump. This reward delivery, as well as the visual stimulus presentation procedures, were controlled using the MonkeyLogic toolbox in MATLAB software. Moreover, a photodiode was used to record the actual time of the visual stimulus incidence. Then the recorded data were digitalized with a sampling frequency of 32 kHz and stored.

The MGS task commenced with the appearance of a fixation point (FP) in the center of the monitor. In this period, called the fixation period, the monkeys are required to fixate at the FP for 1000 ms. After that, while the monkeys were fixating on the FP, a visual stimulus appeared in one of the positions, whether IN (same visual hemifield as the neuron's receptive field; displayed with red dots) or OUT (opposite hemifield relative to the RF of the recorded neuron; displayed with a green dot) conditions shown in Figure 1b. The stimulus remained for 1000 ms and then disappeared (visual period). As soon as the visual cue disappeared, the memory period started and lasted for 1000 ms. During the memory period, the monkeys had to keep gazing at the FP and memorize the location of the disappeared cue. Finally, the monkeys were obliged to make a saccade to the remembered location after the disappearance of the FP to receive a reward. The MGS phases are simply portrayed in Figure 1c, along with the recorded spiking activity for a sample MT neuron during the MGS task (more details can be found in see [18,19]).

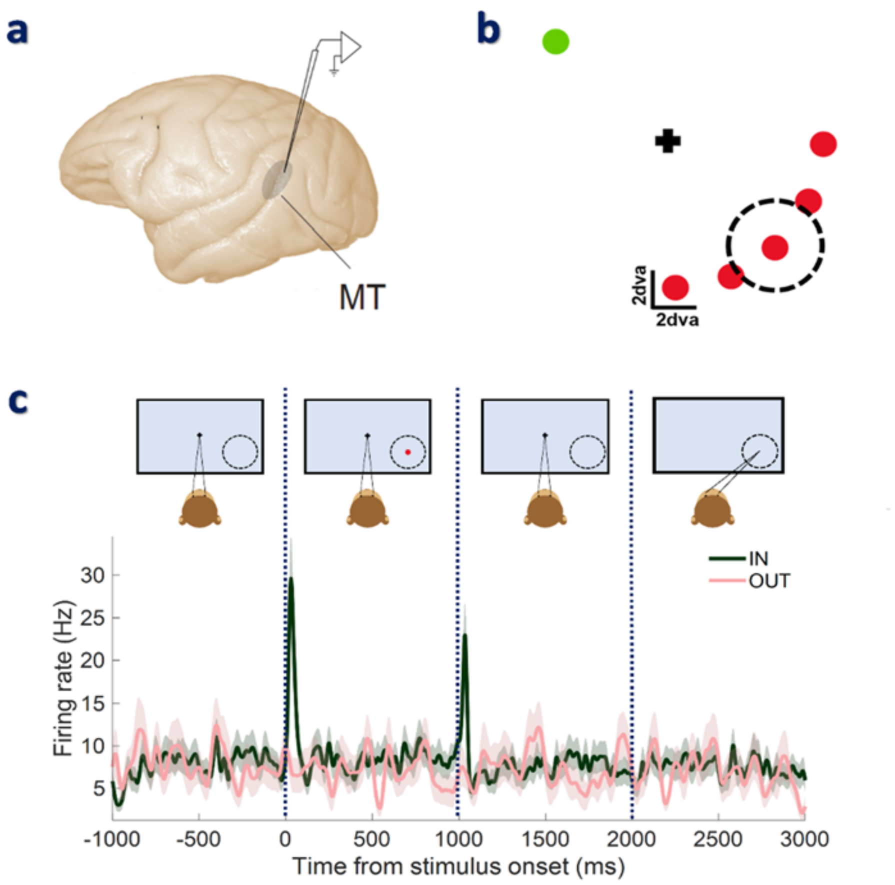

Figure 1. a) The area MT of the brain wherein the firing rate signals were recorded. **b)** Potential positions for the visual stimulus during the visual period of the MGS task. The red dots indicate the locations in the same hemifield (IN conditions), and the green dot indicates the location in the opposite hemifield (OUT condition) as the receptive field of the recorded neuron. The dashed line shows an imaginary receptive field of a sample neuron. **c)** The schematic representation of the MGS task, including the fixation (appearance of the FP), visual (appearance of visual stimulus), memory (disappearance of the visual stimulus), and saccade (after the disappearance of the FP) periods, as well as the corresponding response of a sample neuron. The average neural response during IN and OUT conditions are shown in green and pink, respectively.

2.2. Data preparation

Data preparation plays a vital role in obtaining the best results for machine-learning approaches. Here, in the first step, the average spiking activity of individual neurons across trials was obtained. In the next step, the average firing rate signals of each neuron in IN conditions and OUT were considered for extracting the features. In order to eliminate the spiking activity related to the disappearance of the visual stimulus in the memory period, the first 400 ms of this period were trimmed for the feature extraction step. In the final step, the smoothed signals were considered for the following processes.

2.3. Feature extraction

Extracting distinguishable features is an essential step in detecting the presence of the working memory. This section introduces the six most used fractal features and some statistical measures. Also, four frequently used transforms in signal processing are described.

2.3.1. Features based on fractal dimension

Fractal dimension (FD) is an index of complexity that refers to an object's non-integer dimension geometrically [29]. FD can be obtained using different algorithms; however, in general, it can be obtained based on the number of blocks forming a pattern or covering the data graph. Here, six famous algorithms for calculating the FD are described.

Higuchi fractal dimension (HFD): HFD is an accurate box-counting method that is the main algorithm for calculating the FD of a graph [30]. Any time series $x_t: x(1), x(2), \dots, x(N)$ with a finite number of samples (N) can be expressed as the k sets of x_m^k where

$$x_m(k) = \{x(m), x(m+k), x(m+2k), \dots, x(m+kR)\}. \tag{1}$$

Here, $m = 1,2, \dots, k$ is the initial time, $k = 1,2, \dots, k_m$ is the delay, and $R = \left[\frac{N-m}{k}\right]$. Note that $[\dots]$ denotes the integer part of the internal number. Accordingly, the curve length of each subset can be defined as

$$L_m^{\mathrm{k}} = \frac{N-1}{Rk^2} \sum_{i=1}^{R} \left| x_{m+ik} - x_{m+(i-1)k} \right|, \tag{2}$$

Letting $\langle L^k \rangle_m$ defines the average value of m curve lengths, it can be written that

$$\langle L^k \rangle_m \propto k^{-D}, \tag{3}$$

where D is FD obtained using the Higuchi algorithm. In this paper, $k_m = 30$ was determined based on trial and error.

Katz fractal dimension (KFD): KFD is a distance-based technique for obtaining the FD value that uses the averaged Euclidean distance of two consecutive points in a time series [31]. If L defines the sum of Euclidean distance between every two successive samples ($L = \sum_{i=1}^{N-1} dist_{Euclidean}(n_i, n_{i+1})$) and L_m denotes the maximum value of the Euclidean distance between the first and jth sample ($L_m = max\left(dist_{Euclidean}(n_1, n_j)\right)\ for\ j = 2, \dots, N$), the FD can be obtained as

$$D = \frac{\log(N)}{\log\left(\frac{Nd}{L}\right)}. \tag{4}$$

In the above equation, N is the number of samples in a given time series, and D is the FD value with the Katz technique.

Generalized Hurst exponent (GHE): GHE is known as a fractal index indicating the long-rage dependence of time series that uses the qth-order moment of the distribution [32]. Assuming $x(t)$ as the studied time series with N samples, the the qth-order moment of the distribution can be defined as

$$K_q^\tau = \frac{\langle |x(t) - x(t-\tau)|^q \rangle_t}{\langle |x(t)|^q \rangle}, \tag{5}$$

where τ is the delay and $\langle \ldots \rangle_t$ indicates the average value of the internal value over the total duration. Then, the fractal dimension of the time series can be obtained through

$$K_q^\tau \propto \left(\frac{\tau}{T_s}\right)^{qH}, \tag{6}$$

where T_s is the sampling time, and H refers to the GHE of the time series. In this study, the first-order moment was considered.

Margaos and Sun fractal dimension (MSFD): MSFD is a morphological method for calculating FD since it tries to cover the data graph employing the morphological operators, i.e., erosion and dilation [33]. The support-limited erosions and dilations with the support set s $(s = 1,2, \ldots, N)$ and using the structuring element of b can be formulated as

$$\begin{gathered} erosion\colon \begin{cases} x_n \oplus_s b_k = max\{x_{n-1}, x_n, x_{n+1}\}\ for\ k = 1 \\ x_n \oplus_s b_k = max\{x_{n-1} \oplus_s b_k, x_{n+1} \oplus_s b_k\}\ for\ k \geq 2' \end{cases} \\ dilation\colon \begin{cases} x_n \ominus_s b^k = min\{x_{n-1}, x_n, x_{n+1}\}\ for\ k = 1 \\ x_n \ominus_s b^k = min\{x_{n-1} \ominus_s b_k, x_{n+1} \ominus_s b_k\}\ for\ k \geq 2' \end{cases} \end{gathered} \tag{7}$$

where $k = 1,2, \ldots, k_m$. Consequently, the morphological cover that surrounds the data graph can be obtained as

$$C_k = \sum_{n=1}^{N} [(x_n \oplus_s b_k) - (x_n \ominus_s b_k)]. \tag{8}$$

Here k_m is set according to the rule mentioned in [34]. Finally, the morphological FD of the time series can be acquired as the angular coefficient of linear regression of $ln\left(\frac{C_k}{\left(\frac{2k}{N}\right)^2}\right)$ vs. $ln\left(\frac{1}{\frac{2k}{N}}\right)$.

Leibovich and Toth fractal dimension (LTD): LTD is the fast implementation of the box-counting algorithm [35]. The box-counting algorithm is based on the number of blocks to which the data graph can be split. So, the FD can be found as

$$D = \lim_{\epsilon \to 0} \frac{\log\left(n_{blocks}(d)\right)}{\log\left(\frac{1}{d}\right)}, \tag{9}$$

where d is the size of the blocks and D is the FD base on the original box-counting method. The LTD algorithm implements the box-counting method by withdrawing too small and too large blocks; therefore, it is faster than the original algorithm [36].

Fractal Volatility (FV): FV computes the FD of a time series using the box-counting method with the rand-walk process [37]. In other words, it splits the data into blocks of size d by performing the random-walk process.

2.3.2. Features based on transforms

Discrete Wavelet Transform (DWT): DWT describes any data by a set of weighted orthonormal wavelets called mother wavelets [38] according to the following relation

$$X_{m,n} = \sum_{n=1}^{N} x_n \frac{1}{\sqrt{2^m}} \psi(2^{-m}t - n), \tag{10}$$

where N is the total number of samples, $X_{m,n}$ is the transformed data, ψ is the orthonormal wavelet. Moreover, m and n are two control parameters regarding the translation and dilation operation of the discrete wavelet and the data. DWT includes temporal information as well as frequency information in different scales. In this paper, the db4 was selected as the orthonormal wavelet for further analysis.

Discrete Fourier Transform (DFT): DFT is the most fundamental transformation that describes the data based on some weighted sinusoidal functions. DFT can be defined

$$X_k = \sum_{n=1}^{N} x_n \mathrm{e}^{-j\frac{2\pi}{N}kn}, \tag{11}$$

where $k = 1,2,\ldots,N$ and X_k is the transformes data for a specific frequency.

Discrete Short-Time Fourier Transform (DSTFT): DSTFT is the time-frequency version of the DFT transform that contains only the frequency information. Thus, DSTFT can be helpful when DFT becomes insufficient for analysis, such as nonstationary signals [39]. DSTFT is defined as

$$X_k = \sum_{n=1}^{N} x_n \omega(n - m) \mathrm{e}^{-j\frac{2\pi}{N}kn}, \tag{12}$$

where $\omega(L)$ is the temporal window of size L.

Discrete Stockwell Transform (DST): DST is the extension of the DWT while having a close relationship with DSTFT. Also, it provides a frequency-dependent resolution since the sinusoidal functions are fixed in time, and a scalable Gaussian window operates the dilation and translation [40]. DST can be described as

$$S\left(iT_s, \frac{n}{NT_s}\right) = \sum_{m=0}^{N-1} H\left(\frac{m+n}{NT_s}\right) e^{-\frac{2\pi^2 m^2}{n^2}} e^{\frac{j2\pi mi}{N}} \tag{13}$$

where T_s is the inverse sampling frequency, iT_s defines the window τ, $\frac{n}{NT_s}$ is the frequency domain, and $H(n)$ is the DFT of the input data.

2.3.3. Moments of distribution

Moments of distribution are basically statistical measures, the primary features that can be simply obtained from any data. The first-, second-, third-, and forth-order moments are called mean, variance, skewness, and kurtosis, which are described as follows

$$M_1 = \sum_{n=1}^{N} \frac{x_n}{N}, \tag{14}$$

$$M_2 = \sum_{n=1}^{N} \frac{(x_n - M_1)^2}{N}, \tag{15}$$

$$M_3 = \sum_{n=1}^{N} \frac{(x_n - \mu_1)^3}{N.M_2^{\frac{3}{2}}}, \tag{16}$$

$$M_4 = \sum_{n=1}^{N} \frac{(x_n - \mu_1)^4}{N.M_2^2}. \tag{17}$$

It should be noted that N is the number of data samples.

Median is another statistical measure that can be helpful whenever the mean of that data is not a good measure of the distribution. The median can be defined as follows

$$med = \begin{cases} x\left(\frac{n}{2}\right) & n \in 2k \\ \frac{1}{2}\left(x\left(\frac{n-1}{2}\right) + x\left(\frac{n+1}{2}\right)\right) & n \in 2k-1 \end{cases}, \tag{18}$$

where $k = 1,2,\dots,N$. Maximum and minimum values of the samples within a specific period are the other two famous statistical measures used in this paper.

2.4. Feature selection

In signal processing, feature selection is an optimization method leading to the optimum features for classification. Therefore, feature selection was used as a preprocessing step for machine-learning problems and is of particular importance when the data or extracted features are of high dimension [41]. In this subsection, three popular algorithms, namely, Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and Ant Colony Optimization (ACO) algorithms, are briefly described.

2.4.1. Genetic Optimization

GA is an evolutionary algorithm that can be applied for search and optimization problems inspired by species' natural selection and evolution. In this algorithm, a population of chromosomes, which are, in fact, the solution to the problem, are directly selected for the next generation. Based on the selected species, the new generation, including a few numbers of new chromosomes, is also created using crossover and mutation techniques. This process is heuristically repeated until the solution converges to the optimal value [42]. In this article, five primary chromosomes and 100 repetitions with a crossover rate of 0.6 and a mutation rate of 0.001 were selected as the initial parameters of the GA algorithm for selecting the optimum features.

2.4.2. Particle Swarm Optimization

PSO is an evolutionary computational algorithm that tries to find an optimized solution inspired by natural social behaviors, such as the group behavior observed in schools of fish and flocks of birds. According to this idea, some initial particles are localized in the search space, each having its own position and velocity. These particles move within the search space to find the best solution. In general, particles make their next moves based on their best-experienced position as well as the best-experienced position of the whole population of particles called a swarm. So, the movement of the particles in the search space is not dependent on the gradient, and therefore, it can be applied for differentiable and non-differentiable optimization problems [43]. In this paper, we set cognitive and social factors $C_1 = C_2 = 2$ and inertia weight $w = 1$ with five initial particles and 100 maximum iterations as the assumed parameters for the PSO algorithm used for obtaining the optimum features.

2.4.3. Ant Colony Optimization

ACO is a probabilistic and graph-based evolutionary algorithm that was first proposed based on the natural behavior of ants in hunting their food. In real life, ants leave pheromones on the pass to guide others to resources as they explore their surroundings, and thus, they can find the shortest pass to the food. Inspired by this cooperative-based technique, optimization problems can be solved and handled. First, some initial artificial ants are positioned in the parameter space that move to a solution stochastically. The pheromone trails, which specify the edges of the graph in the ACO algorithm, are obtained for each ant, and the best solution is selected. For the following steps, the edges of the graph become updated and guide the artificial ants toward the solution. This process is repeated iteratively until the solution converges to an optimal value [44]. In this paper, five initial artificial ants, 100 allowed iterations, $\alpha = \tau = \eta = 1$, $\rho = 0.2$, and $\beta = 0.1$ are selected as the parameters needed for applying the ACO algorithm to find optimum features.

2.5. Feature classification

Feature classification is the process of assigning or predicting the label of new data based on the trained model or information gained from the observed data [45]. The Support Vector Machine (SVM) and the K-Nearest Neighbor (KNN) classifiers are two supervised and most-used machine-learning algorithms for classification, which are briefly described below.

2.5.1. Support Vector Machine classifier

The original SVM algorithm can be used for classifying two classes using a linear boundary; however, it has been extended for the classification of multi-class data. The SVM classifier builds an optimum hyperplane with the largest margin—with the maximum distance from the nearest data to the decision boundary—that can distinguish the data with the highest accuracy. If the data is not linearly distinguishable, using nonlinear kernels, the data will be implicitly mapped into the higher dimension wherein the data is distinguishable with a linear boundary [44]. This paper uses the SVM classifier with a three-order polynomial kernel function to classify the neuronal spiking activity in fixation and memory periods.

2.5.2. K-Nearest Neighbor classifier

The KNN algorithm classifies each new data based on voting between the class of k-closest observed data to the new sample. Therefore, to classify any new data, the KNN classifier needs to find the k-nearest training data by computing the distances of the new sample from all other samples in the parameter space. Therefore, although the KNN algorithm is simple to implement, it may be time-consuming due to its computational costs. This paper employs the KNN classifier with $k = 3$ and standardized Euclidean distance to detect the presence of working memory using the firing rate data.

2.5.3. Classification evaluation

To assess the classification performance, different criteria are introduced in the literature [16]. Accuracy, the most well-known evaluation criterion for classification, is defined based on the number of samples correctly labeled by the classifier versus the number of samples (see Eq (19)). Sensitivity and specificity are more specific assessment criteria since they respectively show the performance of the classifier in detecting and not detecting the target class. Therefore, sensitivity is defined as the number of samples correctly labeled as the target class by the classifiers versus the actual number of samples in the target class (see Eq (20)). On the other hand, specificity is defined as the number of samples correctly labeled as the non-target class by the classifiers versus the actual number of samples in the non-target class (see Eq (21)).

$$Accuracy = \frac{tp + tn}{tp + tn + fp + fn}, \tag{19}$$

$$Sensitivity = \frac{tp}{tp + fn}, \tag{20}$$

$$Specificity = \frac{tn}{tn + fp}. \tag{21}$$

Here, tp, tn, fp, and fn refer to the true positive, true negative, false positive, and false negative indexes.

3. Results

Figure 2 shows the average normalized response of 131 MT neurons during the MGS task (see Methods). According to this figure, on average, no change in spiking activity is observed when comparing the response of neurons before and after the visual stimulus (i.e., fixation and memory periods, respectively). The inset bar graph in Figure 2 reveals no significant difference between the average spiking activity of MT neurons in memory vs. fixation periods in both IN and OUT memory conditions ($p_{fixation\ IN\ vs,memory\ IN} = 0.385$ and $p_{fixation\ OUT\ vs,memory\ OUT} = 0.385$).

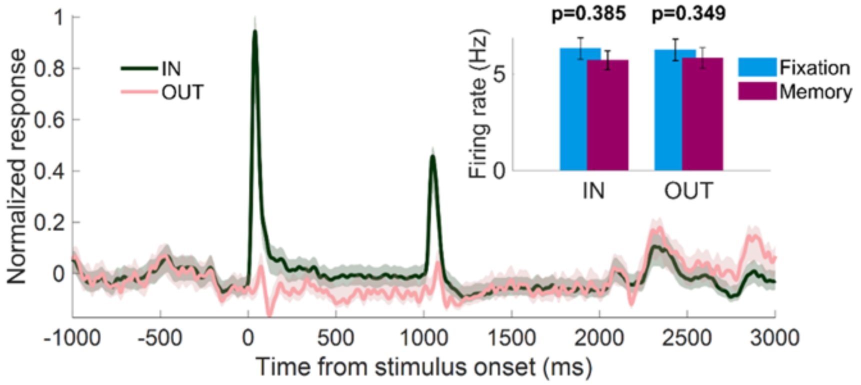

Figure 2. The normalized response of 131 MT neurons during the MGS task, including the fixation, visual, memory, and saccade periods. The average response in IN and OUT conditions are shown in green and pink, respectively. The inset bar graphs show the average firing rate of MT neurons in IN and OUT conditions during the fixation and memory periods.

To detect the presence of the memory, we used the neural spiking activity in IN conditions during the fixation (when no working memory is involved) and memory (when deployment of top-down working memory signals is present) periods for feature extraction, selection, and classification. In the feature extraction step, 41 features are extracted from the IN conditions in fixation and memory periods. Therefore, the feature vector can be described as:

- Index 1–6 are the fractal-based features, including HFD, KFD, GHE, MSFD, LTFD, and FV.
- Index 7–34 are the transform-based features, including mean (index 7–10), variance (index 11–14), kurtosis (index 15–18), skewness (index 19–22), median (index 23–26), maximum (index 27–30), and minimum (index 31–34) of the components of DWT, DFT, DSTFT, and DST, respectively.
- Index 35–41 are statistical features, including mean, variance, kurtosis, skewness, median, maximum, and minimum values of the firing rate signals within the fixation and memory periods.

Finally, the selected features from the fixation IN and memory IN were classified using the SVM and KNN classifiers. Based on the selection method, classification was performed in four cases. To estimate the performance of the classifiers on the data, the k-fold cross-validation method (with $k = 10$) and its balanced version called the A-test algorithm (with $k = 10$ and ten iterations) were employed as the procedure for performing the classification. Unlike the K-fold cross-validation method, the A-test algorithm ensures that the same number of samples from each class is included in all training and testing folds. Therefore, the A-test method might be more reliable, especially when classes contain unequal data samples.

3.1. All extracted features combined

In the first case, all extracted features, including 41 linear and nonlinear features, were used for the classification step. The results of the classification performance are shown in Figure 3. The details of the results can be found in Table 1.

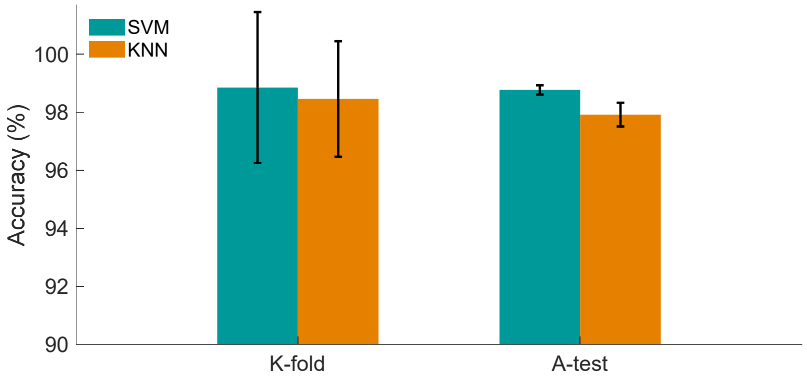

Figure 3. The average performance of the classifiers SVM and KNN using the K-fold (with $K = 10$) and A-test (with $K = 10$ and 10 iterations) cross-validation methods. In this case, the classification was performed using all extracted features.

Table 1. Classification performance results $(mean\% \pm STD\%)$ of fixation IN versus memory IN using all extracted features.

Classifier	Assessment criterion	K-fold ($\boldsymbol{K = 10}$)	A-test ($\boldsymbol{K = 10, iter = 10}$)
SVM	Accuracy	$\mathbf{98.85 \pm 2.60}$	98.77 ± 0.16
	sensitivity	97.69 ± 5.19	97.54 ± 0.32
	specificity	100 ± 0	100 ± 0
KNN	Accuracy	$\mathbf{98.46 \pm 1.99}$	97.92 ± 0.41
	sensitivity	97.69 ± 3.72	97 ± 0.67
	specificity	99.23 ± 2.43	98.85 ± 0.54

Figure 3 shows that the SVM and KNN classifiers SVM and KNN classifiers performed closely ($Accuracy_{SVM} = 98.85 \pm 2.6$; $Accuracy_{KNN} = 98.46 \pm 1.99$) in distinguishing fixation IN and memory IN data. However, SVM performed slightly better than KNN ($Accuracy_{SVM} - Accuracy_{KNN} < 1\%$) in both cross-validation approaches. Table 1 reveals that the average accuracy is higher in the K-fold approach; however, the standard deviation mentioned in the A-test method is considerably lower. This shows that in different iterations, the average accuracy is not changed remarkably, and thus, the results are valid.

3.2. GA selecting method

Here, 20 features were selected using the GA feature selection method, including three fractal-based features (including the HFD, the KFD, and the FV), 14 transform-based features (including the mean of DFT, the mean of DST, the variance of DFT, the variance of DSTFT, the variance of DST, the kurtosis of DWT, the kurtosis of DSTFT, the skewness of DFT, the skewness of DSTFT, the median of DWT, the median of DFT, the median of DST, the maximum of DWT, and the minimum DWT), and three statistical features (including the skewness, the median, and the minimum values). Figure 4 and Table 2 demonstrate the classification results using the 20 GA-based selected features.

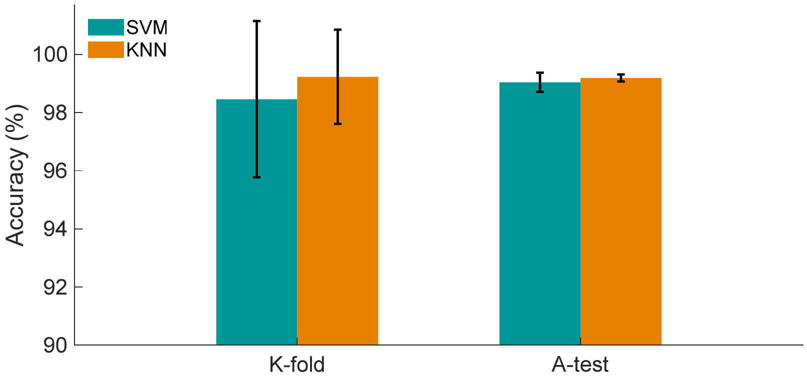

Figure 4. The average performance of the classifiers SVM and KNN using the K-fold (with $K = 10$) and A-test (with $K = 10$ and 10 iterations) cross-validation methods. In this case, the classification was performed using the feature selected by GA.

Table 2. Classification performance results $(mean\% \pm STD\%)$ of fixation IN versus memory IN using the GA as the feature selection method.

Classifier	Assessment criterion	K-fold ($\boldsymbol{K = 10}$)	A-test ($\boldsymbol{K = 10, iter = 10}$)
SVM	Accuracy	98.46 ± 2.69	$\mathbf{99.04 \pm 0.33}$
	sensitivity	96.92 ± 5.38	98.08 ± 0.65
	specificity	100 ± 0	100 ± 0
KNN	Accuracy	$\mathbf{99.23 \pm 1.62}$	99.19 ± 0.12
	sensitivity	98.46 ± 3.24	98.38 ± 0.24
	specificity	100 ± 0	100 ± 0

According to Figure 4, unlike Figure 3, the KNN classifier performed slightly more effectively than the SVM classifier ($Accuracy_{KNN} - Accuracy_{SVM} < 1\%$). Table 2 shows that the GA-selected feature improved the classification performance, especially for the KNN classifier ($Accuracy_{KNN} = 99.23 \pm 1.62$). In addition, the KNN classifier has more reliable results since it generally has a lower standard deviation.

3.3. PSO selecting method

Employing the PSO feature selection method, 17 features were selected, including three fractal-based features (including the HFD, the KFD, and the FV), 12 transform-based features (including the mean of DFT, the mean of DST, the variance of DFT, the variance of DSTFT, the kurtosis of DWT, the kurtosis of DSTFT, the skewness of DFT, the median of DWT, the median of DFT, the median of DST, the maximum of DWT, the minimum of DWT), and two statistical features (including the variance, and the maximum values). Similar to the previous subsections, Figure 5 and Table 3 illustrate the classification performance considering the PSO-based selected features.

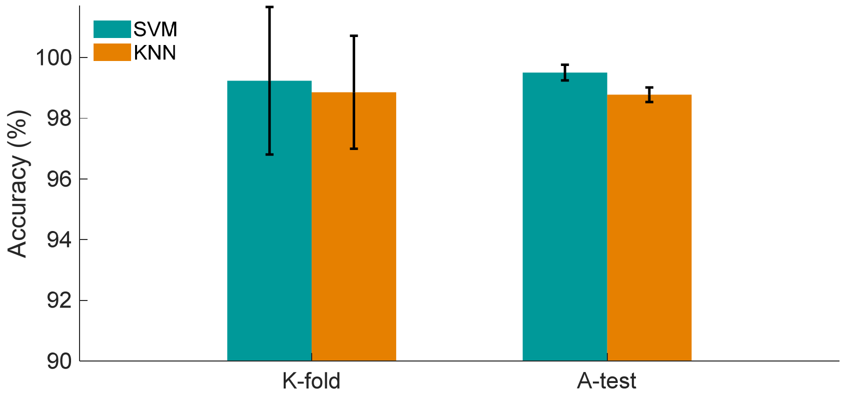

Figure 5. The average performance of the classifiers SVM and KNN using the K-fold (with $K = 10$) and A-test (with $K = 10$ and 10 iterations) cross-validation methods. In this case, the classification was performed using the feature selected by PSO.

Table 3. Classification performance results $(mean\% \pm STD\%)$ of fixation IN versus memory IN using the PSO as the feature selection method.

Classifier	Assessment criterion	K-fold ($\boldsymbol{K = 10}$)	A-test ($\boldsymbol{K = 10, iter = 10}$)
SVM	Accuracy	99.23 ± 2.43	$\mathbf{99.50 \pm 0.26}$
	sensitivity	98.46 ± 4.87	99.00 ± 0.52
	specificity	100 ± 0	100 ± 0
KNN	Accuracy	$\mathbf{98.85 \pm 1.86}$	98.77 ± 0.24
	sensitivity	97.69 ± 3.72	97.77 ± 0.24
	specificity	100 ± 0	99.77 ± 0.37

In contrast to Figure 3, in this case, Figure 5 shows that the PSO-selected features helped the SVM classifier improve its performance ($Accuracy_{SVM} = 99.50 \pm 0.26$) more considerably than the KNN classifier ($Accuracy = 98.85 \pm 1.86$). Besides, Table 3 reveals that the KNN classifier has more reliable results due to the lower standard deviation on average.

3.4. ACO selecting method

Using the ACO algorithm for selecting the optimum features, a total of 20 features were selected, including three fractal-based features (including the HFD, the GHE, and the FV), 14 transform-based features (including the skewness of DSFTF, the variance of DWT, the kurtosis of DSTFT, the skewness of DWT, the kurtosis of DFT, the mean of DST, the skewness of DFT, the min of DFT, the mean of DWT, the variance of DSTFT, the variance of DST, the maximum of DFT), and five statistical features (including the skewness, the maximum, the variance, the kurtosis, and the minimum values). Figure 6, as well as Table 4, contains the results of performing the classification using the ACO-based selected features.

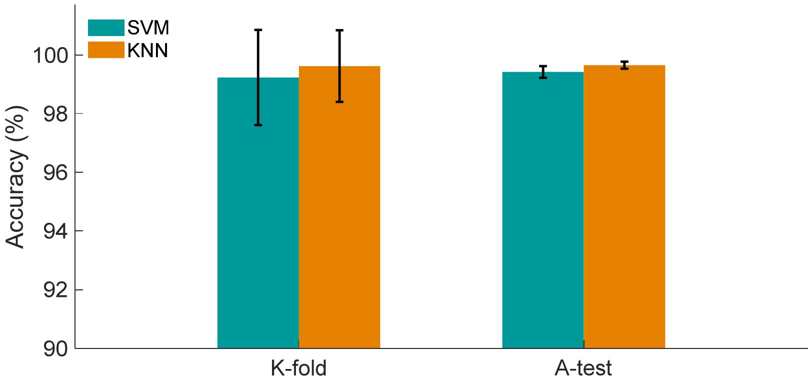

Figure 6. The average performance of the classifiers SVM and KNN using the K-fold (with $K = 10$) and A-test (with $K = 10$ and 10 iterations) cross-validation methods. In this case, the classification was performed using the feature selected by ACO.

Table 4. Classification performance results ($mean\% \pm STD\%$) of fixation IN versus memory IN using the ACO as the feature selection method.

Classifier	Assessment criterion	K-fold ($\boldsymbol{K = 10}$)	A-test ($\boldsymbol{K = 10, iter = 10}$)
SVM	Accuracy	99.23 ± 1.62	$\mathbf{99.42 \pm 0.20}$
	sensitivity	98.46 ± 3.24	98.85 ± 0.41
	specificity	100 ± 0	100 ± 0
KNN	Accuracy	99.62 ± 1.22	$\mathbf{99.65 \pm 0.12}$
	sensitivity	99.23 ± 2.43	99.31 ± 0.24
	specificity	100 ± 0	100 ± 0

Figure 6 illustrates that using the ACO-selected features, the average performance of both classifiers in detecting the presence of working memory is significantly enhanced compared to Figure 3, wherein all features were involved in the classification procedure ($Accuracy_{SVM} = 99.42 \pm 0.20$; $Accuracy_{KNN} = 99.65 \pm 0.12$). From Table 4, it can be seen that the KNN classifier not only has slightly better performance ($Accuracy_{KNN} - Accuracy_{SVM} < 0.5\%$) but also, due to the lower standard deviation, has a more reliable performance.

4. Discussion

The brain is the most complex system in the human body. This complexity is reflected in the signals recorded from the brain. Thus, brain-associated data such as EEG or the spiking activity of neurons predominantly have nonlinear properties. This nonlinearity can be captured by FD, which is an index of complexity. To obtain FD of a time series, different algorithms have been proposed, such as HFD [30], KFD [31], GHE [32], MSFD [33], LTD [35], and FV [37]. However, transform-based features and statistical indexes are popular in signal processing. The main objective of this paper was to examine the ability of machine-learning methods to detect the presence of working memory using various linear and nonlinear features. Therefore, we used several algorithms to obtain the FD value (including HFD, KFD, GHE, MSFD, LTFD, and FV) and different transforms to obtain the frequency and/or time-frequency components (including statistical measures of DWT, DFT, DSTFT, and DST) of the average spiking activity of MT neurons. Also, we included some of the important statistical indexes (including mean, variance, kurtosis, skewness, median, maximum, and minimum values) in the feature set.

Selecting the optimum features can be an optimization problem that, in most cases, results in improving the classifiers' performance. Such methods mainly focus on finding the best set of features that can lead to the best classification result. In this way, it can also help reduce the feature space's dimensionality and simplify the classification problem. Hence, after performing the classification with all extracted features, we examined whether the feature selection method could enhance the classification performance. Accordingly, we used GA, PSO, and ACO algorithms to select the optimum features for detecting the presence of working memory. In the classification step, the ability of two machine-learning algorithms, namely SVM and KNN classifiers, were employed to detect the presence of memory. KNN is typically considered a nonlinear classifier as it has a nonlinear decision boundary whilst SVM can be a nonlinear classifier if it uses a nonlinear kernel function. In general, the nonlinearity of the decision boundary enables a classifier to learn and distinguish the data classes more precisely and define the membership's probability to each data class for new data. For this reason, we used SVM with a three-order polynomial kernel function and KNN with three nearest neighbors. It should be noted that cross-validation methods can determine the validity of classification results, particularly when the number of samples is not too high. Therefore, K-fold (with 10 folds) and A-test (with 10 folds and 10 iterations) cross-validation methods were performed to show the results' robustness.

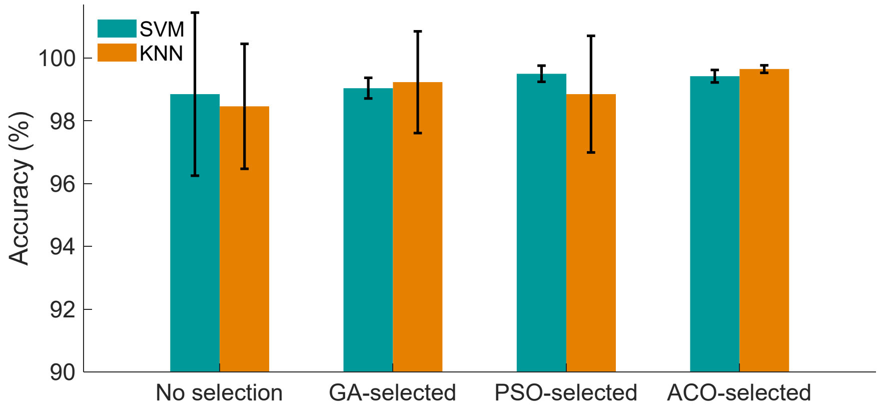

Figure 7. The best performance of KNN and SVM classifiers employing no selection algorithm as well as the GA, PSO, and ACO selection methods.

The best classification performance of the SVM and KNN classifiers is summarized in Figure 7. According to this figure, when no selection algorithms were employed, the SVM classifiers led to the best performance ($Accuracy_{SVM} = 98.85 \pm 2.60$). The same result can be seen in the case where the PSO algorithm was used to select the features ($Accuracy_{SVM} = 99.50 \pm 0.26$). In contrast, when the GA- and ACO-selected features were used as the input of the classification algorithms, the KNN classifier reached higher average accuracy ($Accuracy_{KNN} = 99.23 \pm 1.62$ and 99.65 ± 0.12, respectively). Moreover, Figure 7 shows that employing the selection method can improve the classification performance since, in all cases, the average accuracy was grown compared to the case where no selecting method was employed.

In total, Figure 7 reveals that when the features were selected using the ACO algorithm, the KNN classifier could detect the presence of working memory with the accuracy of 99.65% and the standard deviation of 0.12 (using the A-test cross-validation method), which is the highest obtained average accuracy among the studied cases. In this case, three out of six FD-base features (including the HFD, the GHE, and the FV), 14 out of 28 transform-based features (including the skewness of DSFTF, the variance of DWT, the kurtosis of DSTFT, the skewness of DWT, the kurtosis of DFT, the mean of DST, the skewness of DFT, the min of DFT, the mean of DWT, the variance of DSTFT, the variance of DST, the maximum of DFT), and five out of seven statistical-based features (including the skewness, the maximum, the variance, the kurtosis, and the minimum values) were involved in the classification procedures.

To obtain the neural code for spatial working memory represented in the firing rate of visual neurons, we compared the neural responses of MT neurons during the memory period (where the monkey is actively memorizing a location) with the neural activity during the fixation period (when no working memory is present). This approach could be questioned as one can argue that any differences between the neural responses during memory and fixation periods could occur due to other cognitive signals such as arousal or expectation than spatial working memory. Here we review a series of neurophysiological evidence revealing the dependence of these response changes (i.e., the differences of neural responses between memory and fixation period) on the content of working memory:

- By measuring the neural responses of extrastriate neurons to visual probes presented during fixation and memory periods, a recent study showed a strong modulation of RF profile in V4 and MT neurons that were dependent on the content of working memory [19]. It was shown that during the maintenance of spatial information, only the neurons whose RFs during the fixation period were close to the remembered location, expanded and shifted their RFs towards that location during the memory period.
- It was also found that the encoding of the visual probe's location by the population of MT neurons was enhanced during the memory period compared to the fixation period. In detail, this was measured by the ability of MT individual neurons' firing activity to discriminate two different visual probes (two-point discriminability). In fact, this memory-related enhancement in two-point discriminability only occurred to those visual probes that were presented near the locus of working memory.
- At the level of LFP, it has been shown that the amount of information regarding the visual input conveyed by the alpha-beta phase of spike times increases during the memory period compared to the fixation period. This phenomenon only occurred to those visual probes which were presented near the remembered location.
- Furthermore, it was shown that the discrimination between visual probes, based on the phase of each spike in the alpha-beta frequency range, is enhanced during the memory period compared to fixation. This discrimination enhancement was observed for the visual probes presented near the locus of working memory.

As the differences between the neural responses during memory and fixation periods occur in a spatial-specific manner (i.e., near the locus of working memory), it would be very unlikely to relate these response changes to any other cognitive signals, such as arousal, than spatial working memory.

Acknowledgments

Yaser Merrikhi collected this dataset in the laboratory of Dr. Behrad Noudoost at Montana State University, Bozeman, MT, USA. The lab's experiments were supported by MSU start-up fund, Whitehall 2014-5-18, NIH R01EY026924, and NSF143221 and 1632738 grants to Dr. Noudoost. We would like to thank Dr. Noudoost for sharing this dataset. This work is funded by the Centre for Nonlinear Systems, Chennai Institute of Technology, India, vide funding number CIT/CNS/2022/RP-006.

Conflict of interest

The authors declare there is no conflict of interest.

References

1. A. E. Maxwell, T. A. Warner, F. Fang, Implementation of machine-learning classification in remote sensing: An applied review, *Int. J. Remote Sens.*, **39** (2018), 2784–2817. https://doi.org/10.1080/01431161.2018.1433343
2. D. Maulud, A. M. Abdulazeez, A review on linear regression comprehensive in machine learning, *J. Appl. Sci. Technol. Trends*, **1** (2020), 140–147. https://doi.org/10.38094/jastt1457
3. S. Sun, Z. Cao, H. Zhu, J. Zhao, A survey of optimization methods from a machine learning perspective, *IEEE Trans. Cybern.*, **50** (2019), 3668–3681. https://doi.org/10.1109/TCYB.2019.2950779
4. I. M. Ibrahim, A. M. Abdulazeez, The role of machine learning algorithms for diagnosing diseases, *J. Appl. Sci. Technol. Trends*, **2** (2021), 10–19. https://doi.org/10.38094/jastt20179
5. J. Wang, M. Wang, Review of the emotional feature extraction and classification using EEG signals, *Cognit. Rob.*, **1** (2021), 29–40. https://doi.org/10.1016/j.cogr.2021.04.001
6. A. I. Kadhim, Survey on supervised machine learning techniques for automatic text classification, *Artif. Intell. Rev.*, **52** (2019), 273–292. https://doi.org/10.1007/s10462-018-09677-1
7. C. Jobanputra, J. Bavishi, N. Doshi, Human activity recognition: A survey, *Procedia Comput. Sci.*, **155** (2019), 698–703. https://doi.org/10.1016/j.procs.2019.08.100
8. O. Karabiber Cura, S. Kocaaslan Atli, H. S. Türe, A. Akan, Epileptic seizure classifications using empirical mode decomposition and its derivative, *Biomed. Eng. Online*, **19** (2020), 1–22. https://doi.org/10.1186/s12938-020-0754-y
9. R. Zebari, A. Abdulazeez, D. Zeebaree, D. Zebari, J. Saeed, A comprehensive review of dimensionality reduction techniques for feature selection and feature extraction, *J. Appl. Sci. Technol. Trends*, **1** (2020), 56–70. https://doi.org/10.38094/jastt1224
10. H. Namazi, A. Daneshi, H. Azarnoush, S. Jafari, F. Towhidkhah, Fractal-based analysis of the influence of auditory stimuli on eye movements, *Fractals*, **26** (2018), 1850040. https://doi.org/10.1142/S0218348X18500408
11. H. Alipour, F. Towhidkhah, S. Jafari, A. Menon, H. Namazi, Complexity-based analysis of the relation between fractal visual stimuli and fractal eye movements, *Fluctuation Noise Lett.*, **18** (2019), 1950012. https://doi.org/10.1142/S0219477519500123

12. H. Namazi, E. Aghasian, T. S. Ala, Fractal-based classification of electroencephalography (EEG) signals in healthy adolescents and adolescents with symptoms of schizophrenia, *Technol. Health Care*, **27** (2019), 233–241. https://doi.org/10.3233/THC-181497
13. H. Namazi, R. Khosrowabadi, J. Hussaini, S. Habibi, A. Farid, V. V. Kulish, Analysis of the influence of memory content of auditory stimuli on the memory content of EEG signal, *Oncotarget*, **7** (2016), 56120–56128. https://doi.org/10.18632/oncotarget.11234
14. A. Narin, Y. Isler, M. Ozer, M. Perc, Early prediction of paroxysmal atrial fibrillation based on short-term heart rate variability, *Physica A*, **509** (2018), 56–65. https://doi.org/10.1016/j.physa.2018.06.022
15. Y. Isler, A. Narin, M. Ozer, M. Perc,, Multi-stage classification of congestive heart failure based on short-term heart rate variability, *Chaos, Solitons Fractals*, **118** (2019), 145–151. https://doi.org/10.1016/j.chaos.2018.11.020
16. M. Mehrabbeik, S. Rashidi, A. Fallah, E. R. Khoshnood, Phonocardiography-based mitral valve prolapse detection with using fractional fourier transform, *Biomed. Phys. Eng. Express*, **7** (2020), 015003. https://doi.org/10.1088/2057-1976/abcaab
17. S. Ansari Nasab, S. Panahi, F. Ghassemi, S. Jafari, K. Rajagopal, D. Ghosh, et al., Functional neuronal networks reveal emotional processing differences in children with ADHD, *Cogn. Neurodyn.*, **16** (2022), 91–100. https://doi.org/10.1007/s11571-021-09699-6
18. Y. Merrikhi, K. Clark, B. Noudoost, Concurrent influence of top-down and bottom-up inputs on correlated activity of Macaque extrastriate neurons, *Nat. Commun.*, **9** (2018), 5393. https://doi.org/10.1038/s41467-018-07816-4
19. Y. Merrikhi, K. Clark, E. Albarran, M. Parsa, M. Zirnsak, T. Moore, et al., Spatial working memory alters the efficacy of input to visual cortex, *Nat. Commun.*, **8** (2017), 15041. https://doi.org/10.1038/ncomms15041
20. Y. Merrikhi, M. Shams-Ahmar, H. Karimi-Rouzbahani, K. Clark, R. Ebrahimpour, B. Noudoost, Dissociable contribution of extrastriate responses to representational enhancement of gaze targets, *J. Cognit. Neurosci.*, **33** (2021), 2167–2180. https://doi.org/10.1162/jocn_a_01750
21. D. Zaksas, T. Pasternak, Directional signals in the prefrontal cortex and in area MT during a working memory for visual motion task, *J. Neurosci.*, **26** (2006), 11726–11742. https://doi.org/10.1523/JNEUROSCI.3420-06.2006
22. J. W. Bisley, D. Zaksas, J. A. Droll, T. Pasternak, Activity of neurons in cortical area MT during a memory for motion task, *J. Neurophysiol.*, **91** (2004), 286–300. https://doi.org/10.1152/jn.00870.2003
23. L. Chelazzi, E. K. Miller, J. Duncan, R. Desimone, Responses of neurons in macaque area V4 during memory-guided visual search, *Cereb. Cortex*, **11** (2001), 761–772. https://doi.org/10.1093/cercor/11.8.761
24. M. Mehrabbeik, M. Shams-Ahmar, A. T. Levine, S. Jafari, Y. Merrikhi, Distinctive nonlinear dimensionality of neural spiking activity in extrastriate cortex during spatial working memory; a Higuchi fractal analysis, *Chaos, Solitons Fractals*, **158** (2022), 112051. https://doi.org/10.1016/j.chaos.2022.112051
25. D. Mendoza-Halliday, S. Torres, J. C. Martinez-Trujillo, Sharp emergence of feature-selective sustained activity along the dorsal visual pathway, *Nat. Neurosci.*, **17** (2014), 1255–1262. https://doi.org/10.1038/nn.3785

26. K. M. Armstrong, M. H. Chang, T. Moore, Selection and maintenance of spatial information by frontal eye field neurons, *J. Neurosci.*, **29** (2009), 15621–15629. https://doi.org/10.1523/JNEUROSCI.4465-09.2009
27. X. Zhou, F. Katsuki, X. L. Qi, C. Constantinidis, Neurons with inverted tuning during the delay periods of working memory tasks in the dorsal prefrontal and posterior parietal cortex, *J. Neurophysiol.*, **108** (2012), 31–38. https://doi.org/10.1152/jn.01151.2011
28. Z. Bahmani, M. R. Daliri, Y. Merrikhi, K. Clark, B. Noudoost, Working memory enhances cortical representations via spatially specific coordination of spike times, *Neuron*, **97** (2018), 967–979. https://doi.org/10.1016/j.neuron.2018.01.012
29. R. Esteller, G. Vachtsevanos, J. Echauz, B. Litt, A comparison of waveform fractal dimension algorithms, *IEEE Trans. Circuits. Syst. I Fundam. Theor. Appl.*, **48** (2001), 177–183. https://doi.org/10.1109/81.904882
30. T. Higuchi, Approach to an irregular time series on the basis of the fractal theory, *Physica D*, **31** (1988), 277–283. https://doi.org/10.1016/0167-2789(88)90081-4
31. M. J. Katz, Fractals and the analysis of waveforms, *Comput. Biol. Med.*, **18** (1988), 145–156. https://doi.org/10.1016/0010-4825(88)90041-8
32. R. Morales, T. Di Matteo, R. Gramatica, T. Aste, Dynamical generalized Hurst exponent as a tool to monitor unstable periods in financial time series, *Physica A*, **391** (2012), 3180–3189. https://doi.org/10.1016/j.physa.2012.01.004
33. C. Gómez, R. Hornero, Entropy and complexity analyses in Alzheimer's disease: An MEG study, *Open Biomed. Eng. J.*, **4** (2010), 223–235. https://doi.org/10.2174/1874120701004010223
34. P. Maragos, F. K. Sun, Measuring the fractal dimension of signals: Morphological covers and iterative optimization, *IEEE Trans. Signal Process.*, **41** (1993), 108. https://doi.org/10.1109/TSP.1993.193131
35. L. S. Liebovitch, T. Toth, A fast algorithm to determine fractal dimensions by box counting, *Phys. Lett. A*, **141** (1989), 386–390. https://doi.org/10.1016/0375-9601(89)90854-2
36. B. Klinkenberg, A review of methods used to determine the fractal dimension of linear features, *Math. Geol.*, **26** (1994), 23–46. https://doi.org/10.1007/BF02065874
37. K. Suganthi, G. Jayalalitha, Geometric brownian motion in stock prices, *J. Phys. Conf. Ser.*, **1377** (2019), 012016. https://doi.org/10.1088/1742-6596/1377/1/012016
38. R. M. Rangayyan, *Biomedical Signal Analysis*, John Wiley & Sons, 2015. https://doi.org/10.1002/9781119068129
39. H. H. Giv, Directional short-time Fourier transform, *J. Math. Anal. Appl.*, **399** (2013), 100–107. https://doi.org/10.1016/j.jmaa.2012.09.053
40. R. G. Stockwell, L. Mansinha, R. P. Lowe, Localization of the complex spectrum: the S transform, *IEEE Trans. Signal Process.*, **44** (1996), 998–1001. https://doi.org/10.1109/78.492555
41. V. Kumar, S. Minz, Feature selection: a literature review, *Smart Comput. Rev.*, **4** (2014), 211–229. https://doi.org/10.6029/smartcr.2014.03.007
42. L. Haldurai, T. Madhubala, R. Rajalakshmi, A study on genetic algorithm and its applications, *Int. J. Comput. Sci. Eng.*, **4** (2016), 139.
43. B. Xue, M. Zhang, W. N. Browne, Particle Swarm Optimization for feature selection in classification: A multi-objective approach, *IEEE Trans. Cybern.*, **43** (2013), 1656–1671. https://doi.org/10.1109/TSMCB.2012.2227469

44. B. Chen, L. Chen, Y. Chen, Efficient ant colony optimization for image feature selection, *Signal Process.*, **93** (2013), 1566–1576. https://doi.org/10.1016/j.sigpro.2012.10.022
45. A. A. B. Pessa, R. S. Zola, M. Perc, H. V. Ribeiro, Determining liquid crystal properties with ordinal networks and machine learning, *Chaos, Solitons Fractals*, **154** (2022), 111607. https://doi.org/10.1016/j.chaos.2021.111607