

## Article

# Energy-Efficient EEG-Based Scheme for Autism Spectrum Disorder Detection Using Wearable Sensors

Sarah Alhassan <sup>1,2,\*</sup> , Adel Soudani <sup>1,\*</sup>  and Manan Almusallam <sup>2</sup> <sup>1</sup> Department of Computer Science, College of Computer and Information Science, King Saud University, Riyadh 11362, Saudi Arabia<sup>2</sup> Department of Computer Science, College of Computer and Information Science, Imam Mohammad Ibn Saud Islamic University, Riyadh 11564, Saudi Arabia

\* Correspondence: salhassan.ccis@gmail.com (S.A.); asoudani@ksu.edu.sa (A.S.)

**Abstract:** The deployment of wearable wireless systems that collect physiological indicators to aid in diagnosing neurological disorders represents a potential solution for the new generation of e-health systems. Electroencephalography (EEG), a recording of the brain's electrical activity, is a promising physiological test for the diagnosis of autism spectrum disorders. It can identify the abnormalities of the neural system that are associated with autism spectrum disorders. However, streaming EEG samples remotely for classification can reduce the wireless sensor's lifespan and creates doubt regarding the application's feasibility. Therefore, decreasing data transmission may conserve sensor energy and extend the lifespan of wireless sensor networks. This paper suggests the development of a sensor-based scheme for early age autism detection. The proposed scheme implements an energy-efficient method for signal transformation allowing relevant feature extraction for accurate classification using machine learning algorithms. The experimental results indicate an accuracy of 96%, a sensitivity of 100%, and around 95% of F1 score for all used machine learning models. The results also show that our scheme energy consumption is 97% lower than streaming the raw EEG samples.



**Citation:** Alhassan, S.; Soudani, A.; Almusallam, M. Energy-Efficient EEG-Based Scheme for Autism Spectrum Disorder Detection Using Wearable Sensors. *Sensors* **2023**, *23*, 2228. <https://doi.org/10.3390/s23042228>

Academic Editors: Ki H. Chon and Antonis A. Armoundas

Received: 29 December 2022

Revised: 6 February 2023

Accepted: 15 February 2023

Published: 16 February 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** Autism Spectrum Disorder detection; wearable sensors; EEG signal; on-node feature extraction and classification; embedded machine learning

## 1. Introduction

Autism Spectrum Disorder (ASD) is among the most common childhood neurodevelopmental disorders (approximately 1 in 44 children) [1,2]. According to the American Psychological Association [3], ASD subjects often have restricted and repetitive activity patterns that characterize social, communication, and interaction difficulties. Early diagnosis of ASD can improve the quality of life for autistic children and their families [4] and significantly reduce the severity of later effects [5]. The situation is further complicated as the eligibility of many children for early intervention therapy lapses as they reach school age [6].

Current ASD diagnosis is based on subjective behavioural assessment derived from parent interviews and observations [7]. For an accurate diagnosis, an exhaustive analysis of the child's skills is necessary, which might take months or even years, delaying the starting of accommodation therapy [8]. Furthermore, behaviour-based diagnosis is challenged by the fact that ASD symptoms may overlap with other neurodevelopmental disorders, especially in mild ASD cases [9].

An open research target is the investigation of new autism biomarkers to be used as accurate diagnostic tools that provide early ASD detection and informed clinicians' decisions [10]. There is a need for a scalable ASD biomarker that can be used during standard check-ups [7,11]. This biomarker should be simple, affordable, behaviour-independent, and suitable for routine examinations [8].

Several neuroimaging and neurophysiological methods have been deployed to study the association between brain functionality and autism behaviour [12] to remove the bias of the current subjective behaviour-dependent diagnosis process. Among these methods, electroencephalography (EEG), a measurement of potentials through electrodes placed on the scalp that reflect the brain's electrical activity [13,14], can serve as an investigational tool for characterizing neurodevelopmental disorders [15] and, in particular, for ASD diagnosis [5,9,16]. In fact, EEG-based early brain development evaluation can indicate ASD even before the onset of behavioural symptoms [7]. Even in the absence of ASD-related behaviour, the EEG signal contains temporal, frequency, and phase information that can aid in recognizing abnormal neural activity [17]. The diagnostic potential of EEG signals should, therefore, be capitalized in the development of assistive automatic EEG-based schemes for ASD detection.

Several studies discussed in [9] demonstrated evidence of EEG-based ASD detection. The extracted features from prior investigations are categorized into three primary groups [9]: information dynamics [7], functional connectivity [18–20], and spectral power [5,21,22].

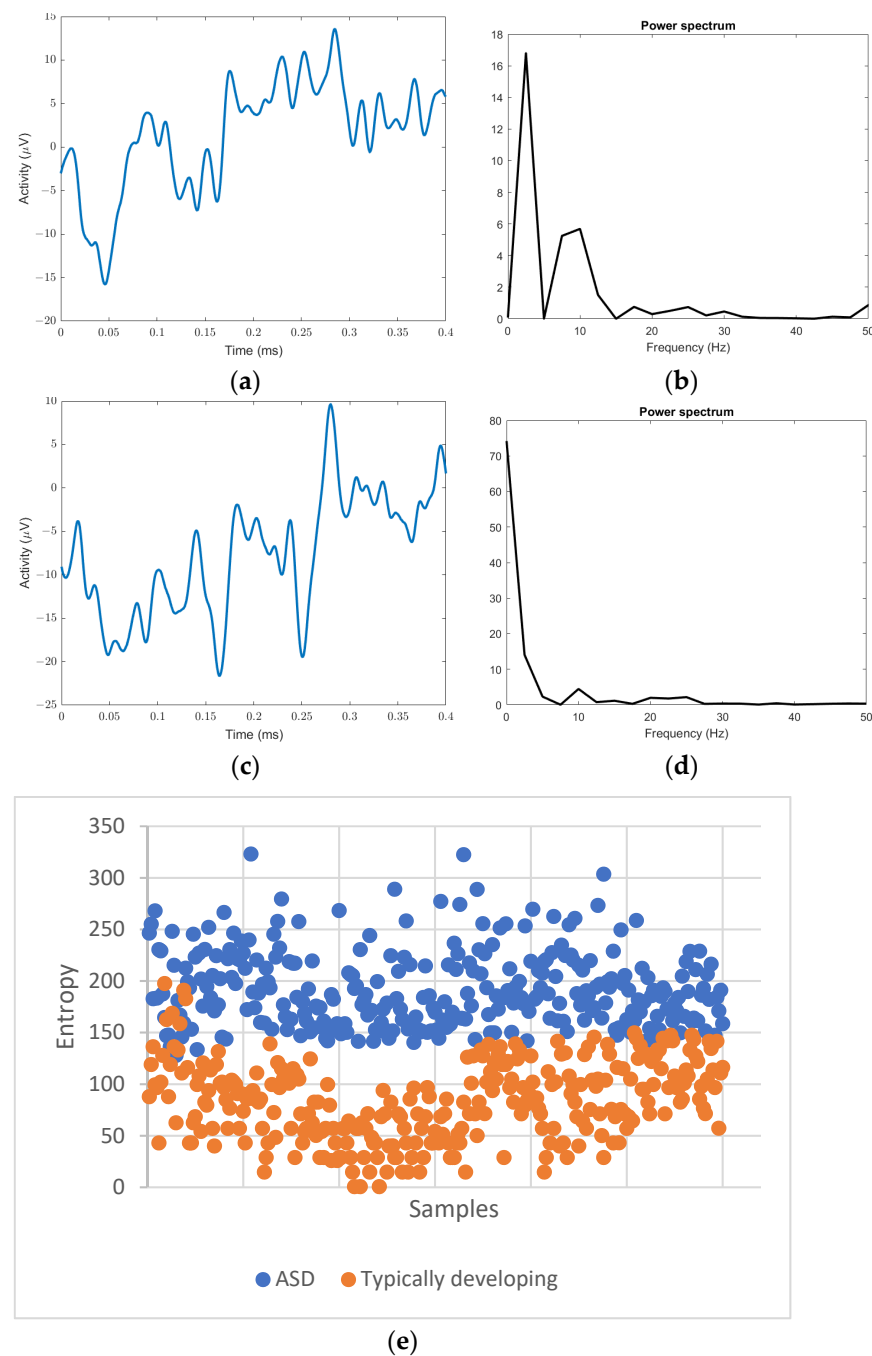
To confirm this statement, we illustrate in Figure 1 the distribution of power spectrum and entropy dynamics features extracted from EEG signals recorded for ASD cases and typically developing groups. Due to nonlinearity, dynamics, and complex characteristics, the EEG signals are difficult to observe and interpret visually [23]. However, the use of power spectrum and entropy dynamic features (Figure 1a–e) shows different distributions between ASD and typically developing subjects in brain function, which attests the feasibility of using EEG-based features for ASD detection.

Previous research contributions studied the analysis of EEG signal for ASD detection with different approaches and based on a variable set of features. Information dynamics features entail nonlinear approaches to find group differences, such as entropy features [9]. Bosl et al. [7] extracted nonlinear measures and claimed it was the first study to apply these measures to developmental neurobiology.

Spectral analysis is considered the most prevalent EEG approach [9]. It was used by Gabard-Durnam et al. [5]. In their approach, the EEG signal was decomposed using FFT, and then the summed power was calculated across all frequencies. Individual alpha peak frequency (iAPF) and individual alpha absolute power (iABP) features were extracted by Zhao et al. [21]. Few studies utilized machine learning in the classification. Bosl et al. [7] utilized the support vector machine (SVM) to produce classification results exceeding 95% in terms of specificity, sensitivity, and positive predictive value. In addition, Zhao et al. [21] employed the support vector machine to achieve 92% accuracy. Gabard-Durnam et al. [5] used logistic regression and reached 90% sensitivity.

Nevertheless, prior studies on EEG-based autism detection relied on wired sensors and did not address the challenges of EEG recording for autistic children. Extensive wiring and lengthy recording sessions often restrict the use of wired EEG sensors to laboratory settings [24,25]. In addition, a lengthy EEG preparation procedure is difficult and unpleasant for young children, which reduces the time available for practical testing [26].

On the other hand, wearable EEG sensors [24] offer inexpensive, non-invasive, and portable real-time monitoring of human brain activity. These sensors are key components in Internet of Medical Things (IoMT) [27] that emerged as the next-generation bioanalytical tools designed to improve the functionality and decision-making ability of healthcare applications. In depth, using wearable EEG sensors reduces the setup burden for EEG specialists and facilitates testing in children's convenient environments, thus eliminating the intimidation effects related to lab settings. Furthermore, it allows multiple analyses at multiple times/days and multiple experimental conditions that more closely simulate everyday experiences [28].

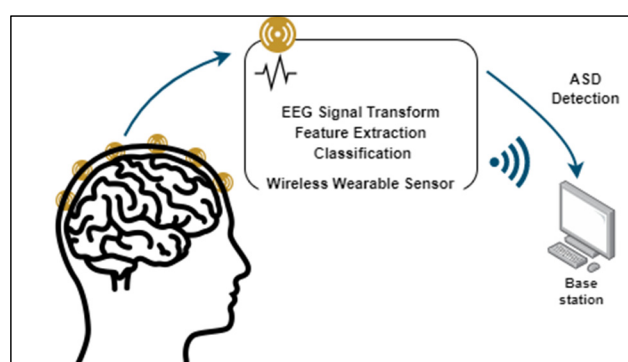


**Figure 1.** Sample EEG signals, their power spectrum, and the distribution of entropy dynamics between ASD and typically developing subjects. (a) ASD EEG sample, (b) power spectrum of ASD EEG sample, (c) typically developing EEG sample, (d) power spectrum of typically developing EEG sample, (e) distribution of entropy dynamics.

In a typical architecture of a wearable health monitoring system [29], wireless EEG sensors are programmed to continuously transmit EEG signal samples to a nearby edge node or a remote cloud server. This approach is challenged by the trade-off between the energy consumption of the wireless sensor and the size of transmitted data [30,31]. Therefore, it is necessary to reduce the amount of data transmitted by the sensor to increase the sensor's lifetime. Alvarez et al. [32] have experimentally validated the energy saving from low-complexity on-sensor EEG compression that reduces the amount of data transmitted. Yet, the inevitable distortion caused by the compression algorithm questions the clinical utility of a reconstructed signal and the preservation of relevant EEG patterns.

On the other hand, embedded machine learning (EML) [33] is a promising solution where classical and deep learning models are designed to be executed within resource-constrained wearable devices to localize signal classification. However, the computational and memory requirements of machine learning algorithms challenge their implementation on embedded microcontrollers. This study demonstrates the technical feasibility and energy efficiency of an ML-based embedded EEG analysis for autism disorder spectrum detection. The technical feasibility of the proposed approach will be proven by showing, on one hand, the capability of the proposed scheme to accurately classifying ASD subjects using segments of EEG signals, and on the other, we will show the adequacy of processing the proposed algorithm in a limited resources wearable sensor.

The main contribution is an innovative EEG-based ASD detection scheme, intended to detect ASD in early-age subjects. It performs on-node EEG signal transformation, features extraction and classification as illustrated in Figure 2.



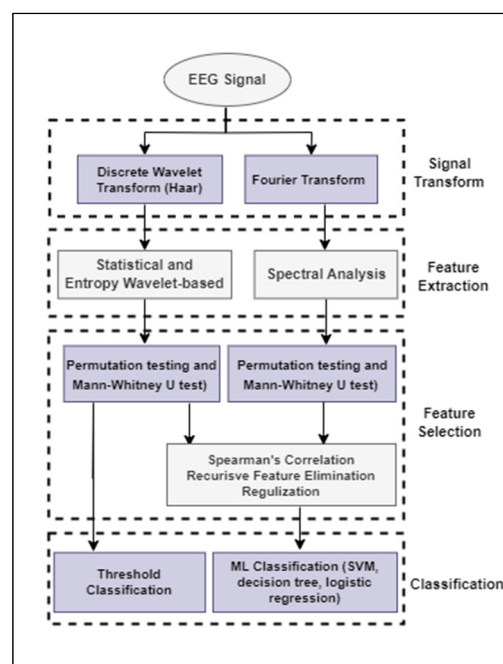
**Figure 2.** On-node EEG-based ASD detection scheme.

This paper is structured as follows. We first describe the specification and the design of the ASD detection scheme as well as the processing approach of the EEG signal. We next present the performance evaluation of this scheme and its accuracy to detect ASD cases in early-age subjects. The energy efficiency is addressed in the last section of this paper to prove the adequacy of the proposed scheme for the wearable sensor. At the end of the paper, we conclude and highlight the extension of this work in the future.

## 2. Design of the ASD Detection Scheme

### 2.1. The Proposed Approach for ASD Detection

Using wearable EEG sensors can improve ASD detection as they eliminate the need for extensive wiring and allow physicians to monitor brain activity in a convenient non-intimidating environment. Figure 3 shows the modular structure of the proposed embedded EEG analysis for ASD detection. First, the EEG signal is transformed into a set of sub-signals at different frequency bands. Signal decomposition is a prerequisite for useful feature extraction that occurs in the subsequent step. The Welch's approach of spectral analysis is applied over pre-processed overlapping windows of EEG segments. A digital wavelet decomposition is also applied to extract the wavelet statistical and information dynamics features relevant to ASD diagnosis. Based on the extracted features, an embedded classifier classifies the EEG segments. The remote backend will be notified if an ASD case has been detected. The proposed scheme requires low-complexity tasks to be processed at the sensor level due to limited resources. The classifiers that we are studying in this paper include a simple threshold classifier and embedded machine learning models, including support vector machine (SVM), logistic regression, and decision tree.



**Figure 3.** General adopted methodology for the EEG-based scheme design.

## 2.2. Signal Analysis and Feature Extraction

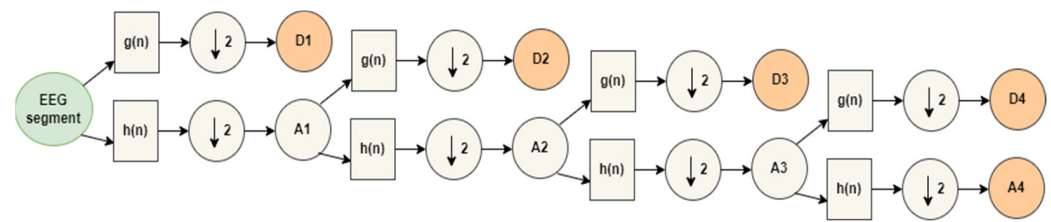
### 2.2.1. Signal Transform

For efficient and relevant features extraction, the EEG signal is often decomposed using a transform method to select the appropriate frequency sub-bands. EEG signals are decomposed into a set of sub-signals at the following frequency bands: delta sub-band (0–4 Hz), theta sub-band (4–8 Hz), alpha sub-band (8–12 Hz), beta sub-band (12–30 Hz), and gamma (30–100 Hz) [34]. Each frequency band has a biological significance and reflects a distinct distribution of rhythmic activity throughout the scalp [35]. We have applied two signal transformations: wavelet transform and Fourier transform.

Wavelet transforms [36–38] have the ability to compress time-varying EEG signals into a small number of parameters using variable-size sliding windows that localize the EEG signal in both frequency and time domains. Discrete wavelet transform (DWT) [10,39] employs discrete scaling parameters (dilation and translation) to one single function called a mother wavelet that acts as a reference in decomposing the original signal. The dilation parameter indicates the frequency and length of the wavelet, while the translation parameter represents the shifting position. Haar wavelets [40] are the simplest and have the lowest computational complexity, making them suitable for on-sensor implementation [41].

Every level of decomposition consists of two digital filters,  $g(n)$  and  $h(n)$ , and two down-samplers, as shown in Figure 4. The high pass filters,  $g(n)$ , produce high-frequency components, while low pass filters,  $h(n)$ , produce low-frequency components. The output of each level of decomposition is a set of details (D) and approximate (A) coefficients. Low-frequency components can be decomposed recursively according to the desired number of decomposition levels [36,42]. The number of levels for wavelet decomposition should be chosen so that the resulting frequencies closely resemble those of typical EEG sub-bands.

A four-level decomposition was used to decompose the EEG signals into detailed coefficients (D1–D4) and approximation coefficients A4. Table 1 summarizes the correspondence between wavelet coefficients and EEG frequency bands.



**Figure 4.** An example of a four-level decomposition of DWT.

**Table 1.** Frequency sub-bands for each wavelet coefficient.

Wavelet Coefficient and Its Frequency	EEG Approximate Label
D1 (25 Hz–50 Hz)	Gamma
D2 (12 Hz–25 Hz)	Beta
D3 (6 Hz–12 Hz)	Alpha
D4 (3 Hz–6 Hz)	Theta
A4 (0.1 Hz–3 Hz)	Delta

The Fourier transform [43,44], a mathematical procedure that decomposes any waveform into a sum of sine waves with varying frequencies, amplitudes, and phases, provides the foundation for EEG spectral analysis. It transforms the signal from the time domain into the frequency domain, allowing the analysis of the power spectrum at different frequencies.

### 2.2.2. Feature Extraction

Different domains such as time, frequency, time-frequency, and nonlinear domains can be used for signal transformation and features extraction [45]. Compared to other techniques, statistical feature extraction and entropy-based techniques yield higher classification accuracies and are, therefore, more prevalent in EEG-based ASD detection [8]. In addition, spectral analysis is one of the dominant features proposed in the literature [5,22,46–50].

For the EEG spectral density, the EEG signal is not stationary over extended periods [43], which challenges the accuracy of power spectrum analysis. An improved power spectral density estimator, the Welch method [35], has been widely used in literature for EEG analysis. It involves averaging the spectral power collected over short window segments, allowing for a significant reduction in power variance.

The Welch’s approach is applied over 100 ms EEG segments with 50% overlapping after applying the Hanning window [51]. The power spectrum was computed as two values for gamma, beta, alpha, theta, and delta sub-bands: the absolute power (i.e., power in a specific frequency band) and relative power (i.e., the ratio of frequency band power to the total power over all frequency bands).

Wavelet statistical features [8,52], which represent the distribution of wavelet coefficients, are often used in EEG-based diagnosis. The wavelet statistical features used in the proposed scheme are root mean square (RMS), variance, and coefficient of variation (CV). The following Equations (1)–(3) represent the selected statistical features.

$$RMS = \sqrt{\frac{\sum_{i=1}^N x_i^2}{N}} \quad (1)$$

$$variance = \frac{1}{N-1} \sum_{i=1}^N |x_i - \mu|^2, \mu = \frac{\sum_{i=1}^N x_i}{N} \quad (2)$$

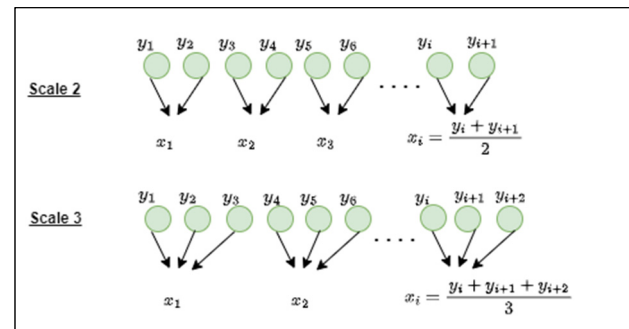
$$CV = \frac{\sqrt{variance}}{\mu} \quad (3)$$

The use of nonlinear signal processing techniques to quantify the temporal dynamics of brain activity is a novel approach [53]. Previous research has shown the significance



of combining time-frequency analysis and nonlinear dynamic features for ASD detection [10]. A nonlinear feature such as entropy, a measure of uncertainty or irregularity of a system [8,53], can be used to indicate functional changes or irregularities in the brain system [53].

In this paper, we adapted the multiscale entropy (MSE) technique introduced by Costa et al. in [53] to measure the complexity of brain functions at multiple time scales. Digital Haar wavelet decomposition is mathematically identical to multiple time scales coarse-graining approach for computing multiscale entropy [7] at a scale of the power of 2, as illustrated in Figure 5.



**Figure 5.** Illustration of the coarse-grained procedure.

For a given time-series samples  $Y = (y_1, y_2, \dots, y_n)$ , the consecutive coarse-grained procedure is performed to have  $\{x^\tau\}$ , at scale vector  $\tau$  using (4):

$$x_j^\tau = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} y_i, \quad 1 \leq j \leq \frac{n}{\tau} \quad (4)$$

We applied four types of multiscale entropies: Shannon, approximate, sample, and modified sample entropies. Multiscale approximate, sample, and modified sample entropies were applied to two versions: raw features and normalized features. We have applied min-max normalization (5) that guarantees all features with the same scale.

$$\text{normalize}(\text{feature}_i) = \frac{\text{feature}_i - \min(\text{feature})}{\max(\text{feature}) - \min(\text{feature})} \quad (5)$$

Approximate entropy (ApEn) [45] represents a statistical measure of a signal's regularity and variability over time. ApEn finds the fluctuation by comparing the signal with its delayed version [10]. Given  $m$  and  $r$ , with  $r$  being the tolerance value and  $m$  being the length of consecutive data points,  $u_m(i) = (x_{1+i}, x_{2+i}, \dots, x_{m+i})$ , approximate entropy is the probability of finding the similarity of a sequence with length  $m$  with the sequence of length  $(m+1)$  [54], as in (6).

$$\text{ApEn}(m, r, N) = \frac{1}{N-m} \sum_{i=1}^{N-m} \frac{\ln n_i^m}{\ln n_i^{m+1}} \quad (6)$$

$n_i^m$  stands for the number of vectors that satisfy the Euclidian distance  $d_{ij}^m$  between  $u_m(i)$  and  $u_m(j)$ , less than or equal to the threshold  $r$ , as expressed in Equation (7).

$$d_{ij}^m = \max \left\{ |x_{i+k} - x_{j+k}|, 0 \leq k \leq m-1, 1 \leq j \leq N-m \right\} \quad (7)$$

We used the default values of  $m$  and  $r$  [54],  $m = 2$ ,  $r = 0.15 \times \text{standard deviation of } x^\tau$ .

Sample Entropy (SamEn) is obtained from approximation entropy, as shown in Equation (8). It is suitable for short data sequences with low noise [45].

$$SamEn(m, r, N) = \ln \frac{\sum_{i=1}^{N-m} n_i^m}{\sum_{i=1}^{N-m} n_i^{m+1}} \quad (8)$$

In sample entropy, the similarities  $d_{ij}$  are computed as 0 or 1, which leads to a strict cut-off in computing similarities. The modified version of sample entropy (mSamEn) [55] computes the similarity of two segments of time series using a sigmoidal function Equation (9). The sigmoid function is the continuous and smoothed version of the 0/1 similarity function in the sample entropy.

$$D_{ij}^m = \frac{1}{1 + \exp \left[ \frac{d_{ij}^m - 0.5}{r} \right]} \quad (9)$$

Shannon entropy quantifies the average degree of signal uncertainty [38]. Shannon's entropy is calculated by Equation (10).

$$Shannon \text{ Entropy} = - \sum_{i=1}^k p(x_i) \log p(x_i) \quad (10)$$

where  $k$  is the number of unique values of  $X$  and  $p$  is the probability of these values.

For each EEG segment, 12 features are extracted to capture the irregularity of EEG and distinguish between ASD subjects and typically developing subjects.

### 2.3. Feature Selection

In our experiment, we have ten channels and 12 features, which yields 10 channels  $\times$  12 features = 120 features for each frequency band and 120 features  $\times$  5 frequency sub-bands = 600 features for each subject. Therefore, we need to minimize the feature dimension by applying the feature selection process [8,45].

To select the most significant features relevant for classification, two non-parametric statistical tests were used: Permutation testing [56] and the Mann–Whitney U-test [57], with a two-tail 95% significant interval ( $p$ -value  $< 0.05$ ). In this approach, the feature will be considered for classification if it is significant in both statistical tests. The 600 features are, thus, reduced to 203 features for each subject.

Further reduction of features was achieved by applying supervised feature selection: filter, wrapper, and embedded methods [58–60]. Table 2 highlights the significant reduction in the number of features for each sub-band.

**Table 2.** The number of features after applying feature selection methods.

Frequency Sub-Band	Number of Features	Number of Features after Permutation and Mann Whitney	Number of Features after Spearman Correlation
gamma	120	25	4
beta	120	56	10
alpha	120	56	8
theta	120	44	10
delta	120	22	4

As a filter model, we have applied a non-parametric spearman correlation [57] for each frequency sub-band. A feature is removed if the correlation coefficient between two features is greater than or equal to 0.8. We have applied the recursive feature elimination (RFE) algorithm as a wrapper method. This method performs model training on a set of gradually smaller features. Every time the feature importance's or coefficients are calculated, the lowest-scoring features are eliminated. As this method trains a model repeatedly, we must



instantiate an estimator. We have used four estimators: logistic regression, perceptron, decision tree, and support vector machine. We applied recursive feature elimination (RFE) for each possible number of features. As an embedded approach, we have used regularization with goal functions that reduce fitting errors while forcing coefficients to be either small or zero.

#### 2.4. Classification and ASD Detection

In computational biology, machine learning (ML) technologies have brought about a new paradigm shift [61–63]. Evidently, the ML-driven approach applied to clinical diagnosis has the potential to supplement traditional methods based on symptoms and external observations, intending to advance the individualized treatment plan [64]. ML approaches are fast expanding fields with applications in computational neuroscience as a result of improved neural data analysis efficiency and decoding brain function [17,61,64–68]. In neuroscience, the issue substantially restricts the extent and depth to which neural signatures can be functionally associated with human behaviour. These deficiencies can be addressed and solved with ML techniques [69].

In the context of ASD detection, ML algorithms demonstrated reliable and robust detection accuracy [70]. As proven by Liao et al. [68], several studies indicated that machine learning is more efficient and objective than conventional ASD diagnostic scales. In this paper, two classifiers are proposed: a simple threshold classifier and an ML classifier.

The threshold classifier is based on statistical and entropy wavelet-based features. Each feature in the test set is compared against thresholds learned from the training set. Thresholds are the mean, minimum, maximum, median, and mode of features from the ASD training set. Each feature in the testing set is classified as ASD or typically developing subjects when it crosses the threshold.

Conventional ML classification algorithms construct classification models with great precision [68]. The supervised learning model identifies the patterns and predicts the class of input data based on prior knowledge. The classification of each test data is determined by combining the features and identifying patterns in the training data. Classification consists of two stages: (1) A classification method is used for the training dataset. (2) The model generated from the training dataset is verified against a test dataset to assess the model's performance and accuracy [59].

We have used three supervised ML classifiers: support vector machine (SVM), logistic regression, and decision tree. These classifiers have been selected due to their simplicity, high interpretability [71,72], and demonstrated accuracy in EEG-based ASD detection [7,21,22,73]. The ML classifier used a combination of spectral analysis features and statistical and entropy wavelet features.

We have adopted hyperparameter tuning to find the best model architecture. This involved creating a model for each possible combination of the specified hyperparameter values, evaluating each model and choosing the architecture that yields the best results. Table 3 shows the tuned hyperparameters and their values for each classifier.

**Table 3.** Hyperparameters tuning.

Classifier	Hyperparameter	Values
Support Vector Machine	Kernel	linear
	regularization parameter, C	0.01, 0.1, 1, 10, 100
Logistic Regression	Solver	liblinear, newton-cg, lbfgs
	regularization parameter, C	0.01, 0.1, 1, 10, 100
Decision Tree	Criterion	gini, entropy
	Splitter	random, best

To implement the ML classifier in the sensor, we extracted the SVM and logistic regression weights/coefficients offline to form the decision classification Equations (11) and (12).

For the decision tree, we trained the model offline. Then, the resulting if-then rules were included in the sensor.

#### 2.4.1. Support Vector Machine

Support vector machine (SVM) is a classifier that separates the two data classes using a hyperplane. The set of data points with the shortest distance from the hyperplane is known as the support vector. Using support vectors, the hyperplane is positioned to maximize the margin, which is a metric that indicates the distance between two classes [74,75].

The SVM technique has been selected due to its high performance with small data sets [76]. We have used a linear kernel, as shown in Equation (11). In Python's scikit-learn module [77], the weights/coefficients are assigned to the features and can be extracted only if the kernel is linear.

$$w^T x + b = 0 \quad (11)$$

where  $w^T$  is the weight/coefficient vector for the feature vector  $x$ , and  $b$  is the bias [78].

#### 2.4.2. Logistic Regression

In medicine and biology, binary classification problems are frequently solved using logistic regression [71]. Logistic regression describes the link between one dependent binary variable and one or more independent variables using a logistic function to predict the probability of a categorical outcome [13,61]. The logistic regression process is shown in Equation (12).

$$\text{Sigmoid}(z) = \frac{1}{1 + e^{-z}}, \quad z = w x + b \quad (12)$$

where  $w$  is the weight/coefficient vector for the feature vector  $x$ , and  $b$  is the bias [13].

#### 2.4.3. Decision Tree

Decision tree is among the most used classifiers in machine learning [79]. The decision tree can reduce complex decision processes into a succession of simpler decisions. The decision tree is a tree that is governed by if-then rules. The tree nodes are questions about the features, representing each answer as a child node. The tree leaves are the classification label [74,78]. The dataset is repeatedly subdivided for binary classification. Optimal partitioning points must be chosen during this procedure [79]. A criterion minimizes the probability of misclassification, including entropy and Gini index, as shown in Equations (13) and (14), where  $p_j$  is the probability of classifying.

$$\text{Entropy} = - \sum_j p_j \log_2 p_j \quad (13)$$

$$\text{Gini} = 1 - \sum_j p_j^2 \quad (14)$$

### 3. Performances Evaluation of the Proposed Scheme

#### 3.1. Classification of ASD Cases

For classification, we used EEG signals provided from a publicly accessible dataset from Catarino et al. [80] study. This dataset includes 15 ASD subjects (mean = 31 months, standard deviation = 6) and 15 typically developing subjects (mean = 29 months, standard deviation = 4). All subjects were right-handed males. Clinical psychologists diagnosed individuals with ASD using worldwide diagnostic criteria. Post-visual stimuli data were obtained using a 32-channel system corresponding to the international 10-20 system [81] and a reference electrode at the tip of the nose. The data were sampled with a bandpass filter between 0.1 and 50 Hz at a sampling rate of 1000 Hz. The number of epochs included in the study [80] was for ASD patients (mean = 81, SD = 8) and for typically developing subjects (mean = 83, SD = 7). The length of the signal is 400 ms post-stimulus period. Ten channels were included in the [80] dataset: P8, TP8, T8, P7, FT8, TP7, F8, T7, FT8, and F7.

For the performance evaluation, the proposed scheme was implemented using MATLAB and Python. A cross-validation approach was adopted in our case because of the small number of subjects. We used the k-fold cross-validation method ( $k = 5$ ), where the dataset is randomly divided into  $k$  ( $k = 5$ ) partitions of equal size, with one partition used for testing and the rest for training for each of the  $k$  iterations [8]. The classification performance of the ASD subject based on the extracted features is determined by averaging the five-fold performance findings. We were interested in evaluating the scheme's accuracy [21], sensitivity [5,7], specificity [5,7], positive predictive value [5,7], negative predictive value [5], and F1-score [79]. The mathematical equations for the performance metrics are given by the following Equations (15)–(20):

- Accuracy (*Acc*): It provides the correct prediction of the classifier.

$$Acc = \frac{TN + TP}{total} \quad (15)$$

- Sensitivity or recall (*Sen*): It expresses the ability of the scheme to identify subjects who have ASD correctly.

$$Sen = \frac{TP}{TP + FN} \quad (16)$$

- Specificity (*Spec*): It shows the scheme's ability to identify typical developing subjects correctly.

$$Spec = \frac{TN}{TN + FP} \quad (17)$$

- Positive Predictive Value (*PPV*) or Precision: It provides the probability of how likely it is that the subject has ASD.

$$PPV = \frac{TP}{TP + FP} \quad (18)$$

- Negative Predictive Value (*NPV*): It expresses the probability of how likely it is that the subject is a typical developing subject.

$$NPV = \frac{TN}{TN + FN} \quad (19)$$

- F1-score (*F1*): It combines both sensitivity and PPV in a single metric.

$$F1 = \frac{2 * Sen * PPV}{Sen + PPV} \quad (20)$$

where *TP*, *FP*, *TN*, *FN*, respectively, represent true positive, false positive, true negative, and false negative.

As explained in the previous section, we studied two scenarios to evaluate the capability of the proposed scheme for ASD recognition. The first scenario is based on the idea of using wavelet-based features with the application of the threshold classification approach. In the second approach, we extended the features to have the spectral analysis features, and the classification was performed based on ML classifiers.

Table 4 summarises the classification results for the first studied scenario based on the threshold classifier [82]. This table shows that the classification based on multiscale approximation entropy was capable of achieving high accuracy in only the beta and alpha sub-bands. The highest accuracy of 86% was obtained in the alpha sub-band using channel P7.

**Table 4.** The best performance measures of threshold classifier.

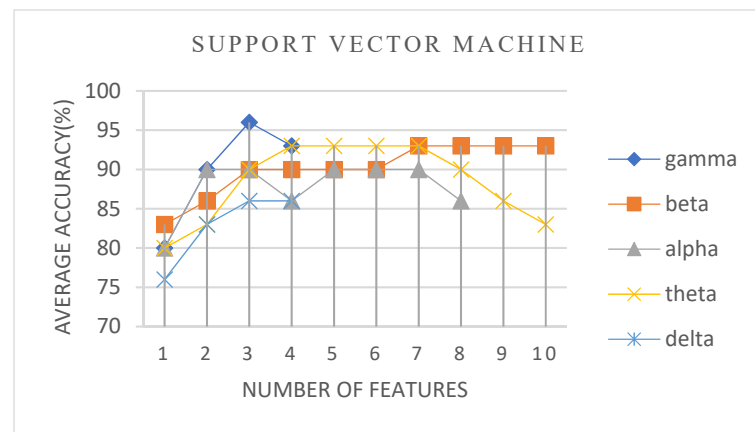
Number of Features	Sub-Band	Acc	Sen	Spec	PPV	NPV	F1
1	Beta	83.33	93.33	73.33	77.78	91.67	84.85
1	Beta	83.33	93.33	73.33	77.78	91.67	84.85
1	Alpha	83.33	80	86.67	85.71	81.25	82.76
1	Alpha	86.67	86.67	86.67	86.67	86.67	86.67
2	alpha/gamma	93.3	86.67	100	100	88.2	92.86

To improve the accuracy in the alpha sub-band, we combined this feature with other features that have 100% NPV or PPV. We used a set of if-then-else rules for the classification decision that allows first to consider the feature that provided 100% NPV or PPV; then, the Multiscale approximation entropy in the alpha-band is applied. We noted that the accuracy was enhanced to 93% when we combined the multiscale approximate entropy in the alpha sub-band in channel P7 with multiscale approximate entropy in channel P8 in the gamma sub-band. However, we still have low sensitivity (86.6%), as shown in Table 4.

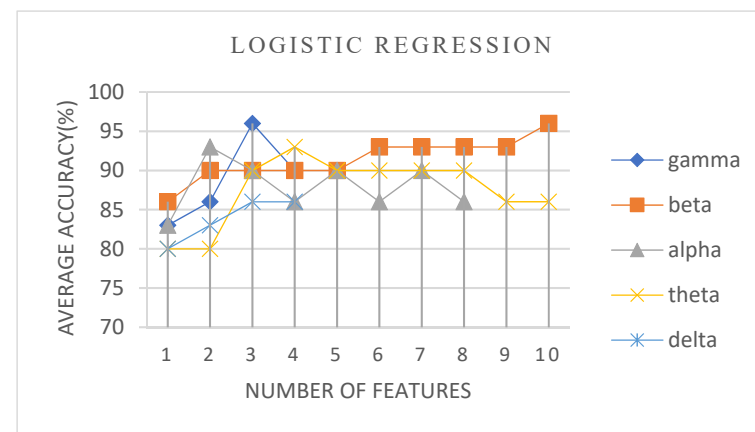
In the second scenario, the training experiments were conducted using K80, T4, and P100 GPU with 52 GB RAM and 8 cores Intel(R) Xeon(R) CPU @ 2.2 GHz. We studied the classification accuracy of the different EEG subjects in the data sets using each possible combination of features resulting from the Recursive Feature Elimination (RFE) algorithm for all combinations of machine learning (ML) hyperparameters values. Figure 6 presents the best overall accuracy values obtained with the classification based on the deployment of different ML algorithms with hyperparameters tuning. From this figure, we can note that the classification with all ML algorithms achieved an accuracy score of 96% in classifying ASD cases. This accuracy is higher than the performance achieved with the threshold classifier approach. This reflects the capability of employing ML models to recognize a pattern that could achieve reliable and accurate diagnostics.

In Figure 6a,b, we can see that the use of SVM and logistic regression algorithms achieved the highest accuracy score with three selected features in the gamma sub-band: absolute Welch, multiscale approximate entropy, and normalized multiscale approximate entropy. The logistic regression algorithm has also achieved an accuracy of 96%, using ten features in the beta sub-band. For a wearable body sensor, adopting fewer features is adequate for low-processing capabilities. With the use of the decision tree classifier, Figure 6c shows that the highest accuracy, 96%, was obtained with the adoption of four selected features in the alpha sub-band: absolute Welch, relative Welch, normalized multiscale approximate entropy, and variance. The results validate the role of feature selection in choosing a subset of highly discriminating features capable of distinguishing samples from distinct classes. Moreover, results prove that too many irrelevant or redundant features in the data can reduce the accuracy of the ML models [59,60].

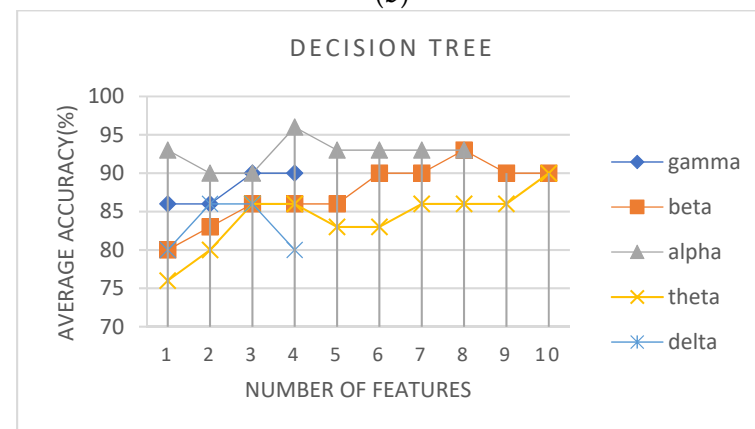
Based on the previous discussion of the results, the best classification accuracies were obtained in the gamma and alpha sub-bands. Gamma frequency oscillations have been linked to various brain activities, such as attention and visual perception, including object perception [83]. On the other hand, Orekova et al. demonstrated that alpha rhythm is associated with attention activities such as visual stimuli [18,19,82]. Additionally, the alpha rhythm is less susceptible to muscle and movement artifacts. Moreover, individual differences in emotional and cognitive involvement have a reduced effect on alpha activity. We, therefore, anticipated that inter-individual differences in these uncontrolled parameters during passive stimulus viewing would contribute less to the alpha sub-band [19].



(a)



(b)



(c)

**Figure 6.** Averaged accuracy variations with the number of features selected by RFE for the ML classifiers for each sub-band. (a) support vector machine, (b) logistic regression, (c) decision tree.

The best selection set of hyperparameters applied to ML algorithms and features used for the accurate classification results are presented in Table 5. From this table, we can observe that multiscale approximate entropy and spectral power (Welch) are the best features for accurate classification to achieve the best discrimination. In depth, approximate entropy is efficient when deployed to calculate the complexity of time-series data, even in the presence of artifacts [84]. It is also suitable for short data, as in our case [85]. The

spectral power (Welch) improves the precision of traditional spectral analysis. Because of the EEG nonstationary propriety, Welch's approach, which involves averaging the spectral power collected over short window segments, reduces this variance significantly.

**Table 5.** Features resulting from RFE and model hyperparameter values for the best performances of ML algorithms.

Classifier	Sub-Band	Number of Features	Channel, Features Resulting from RFE	Hyperparameter	Hyperparameter Value
SVM	Gamma	3	TP7, absolute Welch P8, ApEn F7, ApEn Normalized	Kernel C	Linear 10
Logistic Regression	Gamma	3	TP7, absolute Welch P8, ApEn F7, ApEn Normalized	Solver C	liblinear, newton-cg, lbfgs 10, 100
Decision Tree	Alpha	4	P7, Variance T8, absolute Welch TP7, relative Welch TP8, ApEn Normalized	criterion splitter	entropy random

Despite major differences in how machine learning algorithms are operating, we note from Table 6 that they were all capable of classifying the different EEG signals with high accuracy, sensitivity, and F1 scores. Compared to the threshold classifier, we can say that both the accuracy and sensitivity metrics were significantly enhanced. In depth, the accuracy is increased from 93% to 96%, while the sensitivity has been elevated from 86% to 100%. This result attests the technical feasibility of the proposed approach for efficient detection of ASD cases with the deployment of the described processing techniques.

**Table 6.** The performance of our scheme compared with the benchmarks.

Study	Classifier	Acc	Sen	Spec	PPV	NPV	F1
Our Scheme	Threshold	93.3	86.67	100	100	88.2	92.86
	SVM	96.67	100	95	93.33	100	96.55
	Logistic Regression	96.67	100	95	93.33	100	96.55
	Decision Tree	96.67	100	96	90	100	94.74
Bosl et al. [7]	SVM	63.33	95	35	59	90	72.79
Gabard-Durnam et al. [22]	Logistic Regression	73.33	73.33	75	58.33	80	64.98
Zhao et al. [21]	SVM	73.33	79.33	56.67	67.67	68.33	73.04

Compared to similar EEG-based ASD detection studies, we can note that our proposed scheme outperforms all stated similar schemes reported in the literature, as presented in Table 6. To conduct this comparison, we implemented the feature extraction and classification phases of [5,7,21] with the same dataset that we used. Furthermore, we evaluated the same adopted performance metrics with five-fold cross-validation.

Bosl et al. in [7] used the Daubechies (DB4) wavelet for multiscale decomposition. They extracted nonlinear features from each frequency band: recurrence quantitative analysis, detrended fluctuation analysis, and sample entropy. They used the SVM algorithm for classification with default values of hyperparameters.

Gabard-Durnam et al. [5] used the power spectral of EEG signal with logistic regression classifier. Zhao et al. [21] employed singular spectrum analysis (SSA) to extract the desired alpha rhythm and fed individual alpha peak frequency and individual alpha absolute power features into linear SVM. Bosl et al. [7] achieved a classification accuracy up to 63%, while [5,21] had 73%. The F1 score was in the range of 64–72%. The low performance



of these studies may relate to the variation of studies subjects' ages, experiment designs, extracted features, and/or classifiers.

The highest detection performance of our EEG-based ASD detection scheme clearly indicates that early ASD biomarkers can be extracted from EEG. Time-frequency EEG decomposition, nonlinear features, and spectral power (Welch approach) are promising automated assistive tools for ASD detection that can reduce the bias of the behavioural-based EEG diagnosis and optimize the time and effort of neurologists.

### 3.2. Energy-Consumption Estimation of the Proposed Scheme

The energy consumption of the proposed scheme was performed using Contiki-NG [86,87], an open-source Internet of Things operating system. It is intended for low-power microcontroller emulation. It is integrated with Cooja that allows the emulation of some motes such as Zolertia Z1 platform, which was adopted in our study. The sensor Z1 platform is based on a low-power MSP430 microcontroller with IEEE 802.15.4 radio modules [87].

Contiki-NG uses the Energest module that is capable of estimating the energy and the time related to the processing of a given task. It also estimates the energy related to the radio activities for the transmission and reception of data. Using this information along with the hardware power consumption model according to the mote datasheet, the developer can estimate the system's energy usage.

The energy for each Energest state is expressed by Equations (21)–(23). We have computed energy consumption for the CPU state, the radio transmitting state, and the whole system. The whole system energy consumption is evaluated by summing the values of all tracked states.

$$Current_{state}(mA) = \frac{ticks_{state} * current\_HW_{state}}{RTIMER\_ARCH\_SECOND * Execution\_time_{sec}} \quad (21)$$

$$Power_{state}(mW) = Current_{state} * voltage \quad (22)$$

$$Energy_{state}(mJ) = Power_{state} * Execution\_time_{sec} \quad (23)$$

where  $ticks_{state}$  is the number of clock cycles a system has spent in a state obtained from the Energest module. The  $current\_HW_{state}$  is the current state provided from the mote datasheet.  $RTIMER\_ARCH\_SECOND$  is a mote-specific number of ticks per second.

We assumed that the data segment is already acquired since the data acquisition has the same energy consumption for all scenarios. We studied the following scenarios for the energy evaluation:

- **On-node feature extraction and classification:** In this scenario, we evaluated the energy consumption related to the processing of the EEG signal and the extraction of the features and the classification at the wearable sensor. We implemented the process related to the extraction of the features that provided the highest accuracy, 96%, in our scheme (Table 6).
  - For the classification with SVM and logistic regression, the EEG signal was processed in the gamma sub-band. We evaluated the deployment of the best performance features (absolute Welch, ApEn, and ApEn normalized). For the classification, we added the decision classification Equation (11) for SVM and Equation (12) for logistic regression.
  - For the classification with the decision tree algorithm, the EEG signal was processed in the alpha sub-band. The energy consumption was evaluated for four features of the proposed scheme (absolute Welch, relative Welch, variance, and ApEn normalized). For the classification, we have added the if-else rules resulting from the decision tree model.
- **Streaming raw EEG signal segment:** This scenario is based on the idea of streaming raw EEG signal as in the traditional computerized scheme.

Table 7 shows the results of the execution time in CPU Energest module time results for each different feature proposed to be used for classification in the designed scheme. We can see that the same extracted wavelet features with the classification with the decision tree model generally require less processing time and consequently less energy consumption. In depth, with the application of the decision tree, the features are extracted in the alpha sub-band with fewer wavelet coefficients than the gamma sub-band in SVM or logistic regression classifiers because of the down-sampling process. For example, ApEn normalized feature in the gamma sub-band consumes 217,575 clock cycles at the CPU processing while it takes only 14,957 clock cycles in the alpha sub-band.

**Table 7.** The Energest module time results of each feature in our scheme.

ML Classification Algorithm	Logistic Regression/SVM				Decision Tree		
Sub-band, Extracted Features	gamma, absolute Welch	gamma, ApEn	gamma, ApEn Normalized	alpha, variance	alpha, absolute Welch	alpha, relative Welch	alpha, ApEn Normalized
CPU (ticks) <sup>1</sup>	322,124	217,377	217,575	851	322,096	322,108	14,957
Total Time (ticks) <sup>2</sup>	387,115	294,830	294,830	98,222	387,115	387,115	98,222

<sup>1</sup> The number of clock cycles a system has spent in CPU state. <sup>2</sup> The number of clock cycles a system has spent in all states.

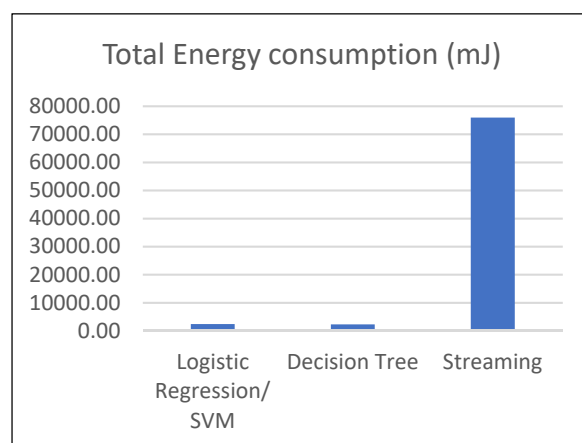
For the spectral analysis feature, the computational energy consumption is higher than the energy consumption of the extraction of the other features. The lowest computation energy consumption is in the variance feature.

Table 8 shows each ML model's energy consumption (transmit, CPU, total energy) in our scheme and streaming scenario. Figure 7 shows a comparison of the energy consumption between the two scenarios. We can see that streaming the whole EEG signal for classification in a remote server needs high energy compared to detecting the ASD disorder with the execution of the proposed scheme in the wearable sensor. A gain of around 97% of energy is ensured while executing the proposed scheme with the decision tree algorithm. This result attests to the proposed scheme's energy efficiency and its adequacy for on-node processing.

**Table 8.** The energy consumption of our proposed scheme and streaming scenario.

Scheme	On-Node Feature Extraction and Classification		Streaming
	Logistic Regression/SVM	Decision Tree	
Total Time (ticks) <sup>1</sup>	976,775	970,674	43,579,790
CPU (ticks) <sup>2</sup>	757,076	660,012	1,062,120
Radio Tx (ticks) <sup>3</sup>	102	102	189,570
Radio Rx (ticks) <sup>4</sup>	976,673	970,572	43,390,220
Transmit Energy consumption (mJ) <sup>5</sup>	0.16	0.16	301.99
CPU Energy consumption (mJ) <sup>6</sup>	693.12	604.26	972.40
Total Energy consumption (mJ) <sup>7</sup>	2374.33	2274.96	75,957.26

<sup>1</sup> The number of clock cycles a system has spent in all states. <sup>2</sup> The number of clock cycles a system has spent in CPU state. <sup>3</sup> The number of clock cycles a system has spent in radio transmitting state. <sup>4</sup> The number of clock cycles a system has spent in radio receiving state. <sup>5</sup> The system energy consumption evaluated for radio transmitting state. <sup>6</sup> The system energy consumption evaluated for CPU state. <sup>7</sup> The summed system energy consumption evaluated for all tracked states.



**Figure 7.** Total energy consumption results in our scheme compared with the streaming scenario.

From another side, we can also see that the decision tree classifier based on different features requires less energy consumption than the deployment of an SVM classifier or logistic regression algorithm. This difference in energy consumption is mainly related to the fact that the decision tree is extracting the features in the alpha sub-band, which requires less computation than the gamma sub-band, during signal transformation.

The different tasks implemented in the proposed scheme were selected to meet the requirement of low complexity for the adequacy of embedded processing. The results presented in Figure 7 and in Tables 7 and 8 demonstrate that the scheme can be efficiently processed in a wearable sensor with limited capabilities attesting about the feasibility of this solution.

#### 4. Conclusions

This paper advocates EEG signals as an objective diagnostic tool for ASD detection in early-age subjects. It demonstrates the technical feasibility of this approach by showing the adequacy of the proposed scheme to be processed in a wearable sensor with limited processing capabilities while maintaining an accurate level of detecting ASD cases. It also attests the energy efficiency of the proposed ML-based embedded EEG analysis for ASD detection with a high level of energy saving. Results have shown that the on-node feature extraction and classification scheme strike the balance of energy efficiency and high accuracy using a combination of nonlinear analysis, multiscale approximate entropies in the time-frequency domain, and spectral analysis (Welch) of EEG signals. The embedded implementation of SVM, logistic regression, and decision trees has reached an accuracy of 96% and has proven to be more energy efficient than typical streaming of non-processed EEG samples. The decision tree yields the highest energy savings, around 97%.

As to future works, we are interested in prototyping the proposed scheme as wearable wireless sensor for in-laboratory experimental deployment, which will help study the effect of classification under data collected by untrained people in uncontrolled environments. In addition, we also think that using larger datasets with an adequate Convolutional Neural Network architecture might contribute to design a scalable and highly accurate assistive tool in clinical decisions. While the use of deep neural networks might be a powerful tool for efficient classification [88], the feasibility of this idea requires performing optimization techniques that can significantly compress the overall classification model size [33,89]. Further optimizations are also required, such as a runtime optimization of the model [90].

**Author Contributions:** Supervision, validation, review, and editing were performed by A.S., M.A.; formal analysis, methodology, paper writing, and software were performed by S.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Acknowledgments:** The authors extend their appreciation to the Deputyship for Research & Innovation, Ministry of Education in Saudi Arabia for funding this research work through the project no. (IFKSURG-1078).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. CDC Data and Statistics on Autism Spectrum Disorder | CDC. Available online: <https://www.cdc.gov/ncbddd/autism/data.html> (accessed on 30 May 2022).
2. Jeste, S.S.; Frohlich, J.; Loo, S.K. Electrophysiological Biomarkers of Diagnosis and Outcome in Neurodevelopmental Disorders. *Curr. Opin. Neurol.* **2015**, *28*, 110–116. [CrossRef] [PubMed]
3. Autism and Autism Spectrum Disorders. Available online: <https://www.apa.org/topics/autism-spectrum-disorder> (accessed on 29 May 2022).
4. Elder, J.; Kreider, C.; Brasher, S.; Ansell, M. Clinical Impact of Early Diagnosis of Autism on the Prognosis and Parent-Child Relationships. *PRBM* **2017**, *10*, 283–292. [CrossRef]
5. Gabard-Durnam, L.J.; Wilkinson, C.; Kapur, K.; Tager-Flusberg, H.; Levin, A.R.; Nelson, C.A. Longitudinal EEG Power in the First Postnatal Year Differentiates Autism Outcomes. *Nat. Commun.* **2019**, *10*, 4188. [CrossRef] [PubMed]
6. Sheldrick, R.C.; Maye, M.P.; Carter, A.S. Age at First Identification of Autism Spectrum Disorder: An Analysis of Two US Surveys. *J. Am. Acad. Child Adolesc. Psychiatry* **2017**, *56*, 313–320. [CrossRef] [PubMed]
7. Bosl, W.J.; Tager-Flusberg, H.; Nelson, C.A. EEG Analytics for Early Detection of Autism Spectrum Disorder: A Data-Driven Approach. *Sci. Rep.* **2018**, *8*, 6828. [CrossRef]
8. Brihadiswaran, G.; Haputhanthri, D.; Gunathilaka, S.; Meedeniya, D.; Jayarathna, S. EEG-Based Processing and Classification Methodologies for Autism Spectrum Disorder: A Review. *J. Comput. Sci.* **2019**, *15*, 1161–1183. [CrossRef]
9. Gurau, O.; Bosl, W.J.; Newton, C.R. How Useful Is Electroencephalography in the Diagnosis of Autism Spectrum Disorders and the Delineation of Subtypes: A Systematic Review. *Front. Psychiatry* **2017**, *8*, 121. [CrossRef]
10. Bhat, S.; Acharya, U.R.; Adeli, H.; Bairy, G.M.; Adeli, A. Automated Diagnosis of Autism: In Search of a Mathematical Marker. *Rev. Neurosci.* **2014**, *25*, 851–861. [CrossRef]
11. Reiersen, A.M. Early Identification of Autism Spectrum Disorder: Is Diagnosis by Age 3 a Reasonable Goal? *J. Am. Acad. Child Adolesc. Psychiatry* **2017**, *56*, 284–285. [CrossRef]
12. Billeci, L.; Sicca, F.; Maharatna, K.; Apicella, F.; Narzisi, A.; Campatelli, G.; Calderoni, S.; Pioggia, G.; Muratori, F. On the Application of Quantitative EEG for Characterizing Autistic Brain: A Systematic Review. *Front. Hum. Neurosci.* **2013**, *7*, 442. [CrossRef]
13. Siuly, S.; Li, Y.; Zhang, Y. *EEG Signal Analysis and Classification*; Health Information Science; Springer International Publishing: Cham, Switzerland, 2016; ISBN 978-3-319-47652-0.
14. Joshi, V.; Nanavati, N. A Review of EEG Signal Analysis for Diagnosis of Neurological Disorders Using Machine Learning. *J.-BPE* **2021**, *7*, 040201. [CrossRef]
15. Heunis, T.; Aldrich, C.; Peters, J.M.; Jeste, S.S.; Sahin, M.; Scheffer, C.; de Vries, P.J. Recurrence Quantification Analysis of Resting State EEG Signals in Autism Spectrum Disorder—A Systematic Methodological Exploration of Technical and Demographic Confounders in the Search for Biomarkers. *BMC Med.* **2018**, *16*, 101. [CrossRef]
16. Ahmadlou, M.; Adeli, H. Electroencephalograms in Diagnosis of Autism. In *Comprehensive Guide to Autism*; Patel, V.B., Preedy, V.R., Martin, C.R., Eds.; Springer New York: New York, NY, USA, 2014; pp. 327–343. ISBN 978-1-4614-4787-0.
17. McPartland, J.C.; Lerner, M.D.; Bhat, A.; Clarkson, T.; Jack, A.; Koohsari, S.; Matuskey, D.; McQuaid, G.A.; Su, W.-C.; Trevisan, D.A. Looking Back at the Next 40 Years of ASD Neuroscience Research. *J. Autism. Dev. Disord.* **2021**, *51*, 4333–4353. [CrossRef]
18. Haartsen, R.; Jones, E.J.H.; Orekhova, E.V.; Charman, T.; Johnson, M.H. The BASIS team. Functional EEG Connectivity in Infants Associates with Later Restricted and Repetitive Behaviours in Autism; a Replication Study. *Transl. Psychiatry* **2019**, *9*, 66. [CrossRef]
19. Orekhova, E.V.; Elsabbagh, M.; Jones, E.J.; Dawson, G.; Charman, T.; Johnson, M.H.; The BASIS Team. EEG Hyper-Connectivity in High-Risk Infants Is Associated with Later Autism. *J. Neurodevel. Disord.* **2014**, *6*, 40. [CrossRef] [PubMed]
20. Dickinson, A.; Daniel, M.; Marin, A.; Gaonkar, B.; Dapretto, M.; McDonald, N.; Jeste, S. Multivariate Neural Connectivity Patterns in Early Infancy Predict Later Autism Symptoms. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **2020**, *6*, 59–69. [CrossRef] [PubMed]
21. Zhao, J.; Song, J.; Li, X.; Kang, J. A Study on EEG Feature Extraction and Classification in Autistic Children Based on Singular Spectrum Analysis Method. *Brain Behav.* **2020**, *10*, e01721. [CrossRef] [PubMed]
22. Wilkinson, C.L.; Levin, A.R.; Gabard-Durnam, L.J.; Tager-Flusberg, H.; Nelson, C.A. Reduced Frontal Gamma Power at 24 Months Is Associated with Better Expressive Language in Toddlers at Risk for Autism. *Autism Res.* **2019**, *12*, 1211–1224. [CrossRef]
23. Heunis, T.-M.; Aldrich, C.; de Vries, P.J. Recent Advances in Resting-State Electroencephalography Biomarkers for Autism Spectrum Disorder—A Review of Methodological and Clinical Challenges. *Pediatr. Neurol.* **2016**, *61*, 28–37. [CrossRef]
24. Lau-Zhu, A.; Lau, M.P.H.; McLoughlin, G. Mobile EEG in Research on Neurodevelopmental Disorders: Opportunities and Challenges. *Dev. Cogn. Neurosci.* **2019**, *36*, 100635. [CrossRef]

25. Ratti, E.; Waninger, S.; Berka, C.; Ruffini, G.; Verma, A. Comparison of Medical and Consumer Wireless EEG Systems for Use in Clinical Trials. *Front. Hum. Neurosci.* **2017**, *11*, 398. [[CrossRef](#)] [[PubMed](#)]
26. Mihajlovic, V.; Grundlehner, B.; Vullers, R.; Penders, J. Wearable, Wireless EEG Solutions in Daily Life Applications: What Are We Missing? *IEEE J. Biomed. Health Inform.* **2015**, *19*, 6–21. [[CrossRef](#)] [[PubMed](#)]
27. Manickam, P.; Mariappan, S.A.; Murugesan, S.M.; Hansda, S.; Kaushik, A.; Shinde, R.; Thipperudraswamy, S.P. Artificial Intelligence (AI) and Internet of Medical Things (IoMT) Assisted Biomedical Systems for Intelligent Healthcare. *Biosensors* **2022**, *12*, 562. [[CrossRef](#)]
28. Johnson, K.T.; Picard, R.W. Advancing Neuroscience through Wearable Devices. *Neuron* **2020**, *108*, 8–12. [[CrossRef](#)] [[PubMed](#)]
29. Wan, J.; Al-awlaqi, M.A.A.H.; Li, M.; O'Grady, M.; Gu, X.; Wang, J.; Cao, N. Wearable IoT Enabled Real-Time Health Monitoring System. *J. Wirel. Com. Netw.* **2018**, *2018*, 298. [[CrossRef](#)]
30. Almusallam, M.; Soudani, A. Feature-Based ECG Sensing Scheme for Energy Efficiency in WBSN. In Proceedings of the 2017 International Conference on Informatics, Health & Technology (ICIHT), Riyadh, Saudi Arabia, 21–23 February 2017; IEEE: Piscataway, NJ, USA; pp. 1–6.
31. Soudani, A.; Almusallam, M. Atrial Fibrillation Detection Based on ECG-Features Extraction in WBSN. *Procedia Comput. Sci.* **2018**, *130*, 472–479. [[CrossRef](#)]
32. Dufort y Alvarez, G.; Favaro, F.; Lecumberry, F.; Martin, A.; Oliver, J.P.; Oreggioni, J.; Ramirez, I.; Seroussi, G.; Steinfeld, L. Wireless EEG System Achieving High Throughput and Reduced Energy Consumption Through Lossless and Near-Lossless Compression. *IEEE Trans. Biomed. Circuits Syst.* **2018**, *12*, 231–241. [[CrossRef](#)]
33. Ajani, T.S.; Imoize, A.L.; Atayero, A.A. An Overview of Machine Learning within Embedded and Mobile Devices—Optimizations and Applications. *Sensors* **2021**, *21*, 4412. [[CrossRef](#)]
34. Hashemian, M.; Pourghassem, H. Diagnosing Autism Spectrum Disorders Based on EEG Analysis: A Survey. *Neurophysiology* **2014**, *46*, 183–195. [[CrossRef](#)]
35. Hu, L.; Zhang, Z. (Eds.) *EEG Signal Processing and Feature Extraction*; Springer: Singapore, 2019; ISBN 9789811391132.
36. Al-Fahoum, A.S.; Al-Fraihat, A.A. Methods of EEG Signal Features Extraction Using Linear Analysis in Frequency and Time-Frequency Domains. *ISRN Neurosci.* **2014**, *2014*, 1–7. [[CrossRef](#)]
37. Iftikhar, M.; Khan, S.A.; Hassan, A. A Survey of Deep Learning and Traditional Approaches for EEG Signal Processing and Classification. In Proceedings of the 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, Canada, 1–3 November 2018; p. 6.
38. Djemal, R.; AlSharabi, K.; Ibrahim, S.; Alsuwailam, A. EEG-Based Computer Aided Diagnosis of Autism Spectrum Disorder Using Wavelet, Entropy, and ANN. *BioMed. Res. Int.* **2017**, *2017*, 1–9. [[CrossRef](#)] [[PubMed](#)]
39. Mallat, S.G. *A Wavelet Tour of Signal Processing: The Sparse Way*, 3rd ed.; Elsevier: Amsterdam, The Netherlands; Academic Press: Boston, MA, USA, 2009; ISBN 978-0-12-374370-1.
40. Van Drongelen, W. Wavelet Analysis. In *Signal Processing for Neuroscientists*; Elsevier: Amsterdam, The Netherlands, 2018; pp. 401–423. ISBN 978-0-12-810482-8.
41. Alhassan, S.; AlDammas, M.A.; Soudani, A. Energy-Efficient Sensor-Based EEG Features' Extraction for Epilepsy Detection. *Procedia Comput. Sci.* **2019**, *160*, 273–280. [[CrossRef](#)]
42. Jiang, X.; Bian, G.-B.; Tian, Z. Removal of Artifacts from EEG Signals: A Review. *Sensors* **2019**, *19*, 987. [[CrossRef](#)]
43. Walczak, T.S.; Chokroverty, S. Electroencephalography, Electromyography, and Electro-Oculography: General Principles and Basic Technology. In *Sleep Disorders Medicine*; Sudhansu, C., Ed.; Elsevier: Amsterdam, The Netherlands, 2009; pp. 157–181. ISBN 978-0-7506-7584-0.
44. Van Drongelen, W. Continuous, Discrete, and Fast Fourier Transform. In *Signal Processing for Neuroscientists*; Elsevier: Amsterdam, The Netherlands, 2018; pp. 103–118. ISBN 978-0-12-810482-8.
45. Sridevi, S.; ShinyDuela, D.J. A comprehensive study on eeg signal processing—methods, challenges and applications. *IT in Industry* **2021**, *9*, 3.
46. Gabard-Durnam, L.; Tierney, A.L.; Vogel-Farley, V.; Tager-Flusberg, H.; Nelson, C.A. Alpha Asymmetry in Infants at Risk for Autism Spectrum Disorders. *J. Autism. Dev. Disord.* **2015**, *45*, 473–480. [[CrossRef](#)]
47. Levin, A.R.; Varcin, K.J.; O'Leary, H.M.; Tager-Flusberg, H.; Nelson, C.A. EEG Power at 3 Months in Infants at High Familial Risk for Autism. *J. Neurodev. Disord.* **2017**, *9*, 34. [[CrossRef](#)]
48. Damiano-Goodwin, C.R.; Woynaroski, T.G.; Simon, D.M.; Ibañez, L.V.; Murias, M.; Kirby, A.; Newsom, C.R.; Wallace, M.T.; Stone, W.L.; Cascio, C.J. Developmental Sequelae and Neurophysiologic Substrates of Sensory Seeking in Infant Siblings of Children with Autism Spectrum Disorder. *Dev. Cogn. Neurosci.* **2018**, *29*, 41–53. [[CrossRef](#)]
49. Simon, D.M.; Damiano, C.R.; Woynaroski, T.G.; Ibañez, L.V.; Murias, M.; Stone, W.L.; Wallace, M.T.; Cascio, C.J. Neural Correlates of Sensory Hyporesponsiveness in Toddlers at High Risk for Autism Spectrum Disorder. *J. Autism. Dev. Disord.* **2017**, *47*, 2710–2722. [[CrossRef](#)]
50. Wang, J.; Barstein, J.; Ethridge, L.E.; Mosconi, M.W.; Takarae, Y.; Sweeney, J.A. Resting State EEG Abnormalities in Autism Spectrum Disorders. *J. Neurodev. Disord.* **2013**, *5*, 24. [[CrossRef](#)]
51. Van Drongelen, W. 1-D and 2-D Fourier Transform Applications. In *Signal Processing for Neuroscientists*; Elsevier: Amsterdam, The Netherlands, 2018; pp. 119–152. ISBN 978-0-12-810482-8.



52. Bhuvaneswari, P.; Kumar, J.S. Influence of Linear Features in Nonlinear Electroencephalography (EEG) Signals. *Procedia Comput. Sci.* **2015**, *47*, 229–236. [\[CrossRef\]](#)
53. Maximo, J.O.; Nelson, C.M.; Kana, R.K. Unrest While Resting? Brain Entropy in Autism Spectrum Disorder. *Brain Res.* **2021**, *1762*, 147435. [\[CrossRef\]](#)
54. Pan, Y.-H.; Lin, W.-Y.; Wang, Y.-H.; Lee, K.-T. Computing multiscale entropy with orthogonal range search. *J. Mar. Sci. Technol.* **2011**, *19*, 7. [\[CrossRef\]](#)
55. Xie, H.-B.; He, W.-X.; Liu, H. Measuring Time Series Regularity Using Nonlinear Similarity-Based Sample Entropy. *Phys. Lett. A* **2008**, *372*, 7140–7146. [\[CrossRef\]](#)
56. Bonnini, S.; Corain, L.; Marozzi, M.; Salmaso, L. *Nonparametric Hypothesis Testing: Rank and Permutation Methods with Applications in R*; Wiley: Hoboken, NJ, USA, 2014; ISBN 978-1-119-95237-4.
57. Dodge, Y. *The Concise Encyclopedia of Statistics*, 1st ed.; Springer: New York, NY, USA, 2008; ISBN 978-0-387-31742-7.
58. García, S.; Luengo, J.; Herrera, F. *Data Preprocessing in Data Mining*; Intelligent Systems Reference Library; Springer International Publishing: Cham, Switzerland, 2015; Volume 72, ISBN 978-3-319-10246-7.
59. Aggarwal, C.C. (Ed.) *Data Classification Algorithms and Applications*, 1st ed.; Chapman and Hall/CRC: Boca Raton, FL, USA, 2014; ISBN 978-1-4665-8674-1.
60. Xue, B.; Zhang, M.; Browne, W.N.; Yao, X. A Survey on Evolutionary Computation Approaches to Feature Selection. *IEEE Trans. Evol. Comput.* **2016**, *20*, 606–626. [\[CrossRef\]](#)
61. Segato, A.; Marzullo, A.; Calimeri, F.; De Momi, E. Artificial Intelligence for Brain Diseases: A Systematic Review. *APL Bioeng.* **2020**, *4*, 041503. [\[CrossRef\]](#) [\[PubMed\]](#)
62. Koteluk, O.; Wartecki, A.; Mazurek, S.; Kołodziejczak, I.; Mackiewicz, A. How Do Machines Learn? Artificial Intelligence as a New Era in Medicine. *JPM* **2021**, *11*, 32. [\[CrossRef\]](#) [\[PubMed\]](#)
63. Mazlan, A.U.; Sahabudin, N.A.; Remli, M.A.; Ismail, N.S.N.; Mohamad, M.S.; Nies, H.W.; Abd Warif, N.B. A Review on Recent Progress in Machine Learning and Deep Learning Methods for Cancer Classification on Gene Expression Data. *Processes* **2021**, *9*, 1466. [\[CrossRef\]](#)
64. Gupta, C.; Chandrashekar, P.; Jin, T.; He, C.; Khullar, S.; Chang, Q.; Wang, D. Bringing Machine Learning to Research on Intellectual and Developmental Disabilities: Taking Inspiration from Neurological Diseases. *J. Neurodev. Disord.* **2022**, *14*, 28. [\[CrossRef\]](#)
65. Gemein, L.A.W.; Schirrmester, R.T.; Chrabaszcz, P.; Wilson, D.; Boedecker, J.; Schulze-Bonhage, A.; Hutter, F.; Ball, T. Machine-Learning-Based Diagnostics of EEG Pathology. *NeuroImage* **2020**, *220*, 117021. [\[CrossRef\]](#)
66. Dev, A.; Roy, N.; Islam, K.; Biswas, C.; Ahmed, H.U.; Amin, A.; Sarker, F.; Vaidyanathan, R.; Mamun, K.A. Exploration of EEG-Based Depression Biomarkers Identification Techniques and Their Applications: A Systematic Review. *IEEE Access* **2022**, *10*, 16756–16781. [\[CrossRef\]](#)
67. Noor, N.S.E.M.; Ibrahim, H. Machine Learning Algorithms and Quantitative Electroencephalography Predictors for Outcome Prediction in Traumatic Brain Injury: A Systematic Review. *IEEE Access* **2020**, *8*, 102075–102092. [\[CrossRef\]](#)
68. Saeidi, M.; Karwowski, W.; Farahani, F.V.; Fiok, K.; Taiar, R.; Hancock, P.A.; Al-Juaid, A. Neural Decoding of EEG Signals with Machine Learning: A Systematic Review. *Brain Sci.* **2021**, *11*, 1525. [\[CrossRef\]](#) [\[PubMed\]](#)
69. Vahid, A.; Mückschel, M.; Stober, S.; Stock, A.-K.; Beste, C. Applying Deep Learning to Single-Trial EEG Data Provides Evidence for Complementary Theories on Action Control. *Commun. Biol.* **2020**, *3*, 112. [\[CrossRef\]](#) [\[PubMed\]](#)
70. Bussu, G.; Jones, E.J.H.; Charman, T.; Johnson, M.H.; Buitelaar, J.K. Prediction of Autism at 3 Years from Behavioural and Developmental Measures in High-Risk Infants: A Longitudinal Cross-Domain Classifier Analysis. *J. Autism. Dev. Disord.* **2018**, *48*, 2418–2433. [\[CrossRef\]](#)
71. Musa, A.B. Comparative Study on Classification Performance between Support Vector Machine and Logistic Regression. *Int. J. Mach. Learn. Cyber.* **2013**, *4*, 13–24. [\[CrossRef\]](#)
72. Molnar, C. *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*, 2nd ed.; 2022; ISBN 979-8411463330.
73. Liao, M.; Duan, H.; Wang, G. Application of Machine Learning Techniques to Detect the Children with Autism Spectrum Disorder. *J. Healthc. Eng.* **2022**, *2022*, 1–10. [\[CrossRef\]](#)
74. Hosseini, M.-P.; Hosseini, A.; Ahi, K. A Review on Machine Learning for EEG Signal Processing in Bioengineering. *IEEE Rev. Biomed. Eng.* **2021**, *14*, 204–218. [\[CrossRef\]](#)
75. Russell, S.J.; Norvig, P. *Artificial Intelligence: A Modern Approach*, 3rd ed.; Pearson: New York, NY, USA, 2009; ISBN 978-0-13-604259-4.
76. Abdel Hameed, M.; Hassaballah, M.; Hosney, M.E.; Alqahtani, A. An AI-Enabled Internet of Things Based Autism Care System for Improving Cognitive Ability of Children with Autism Spectrum Disorders. *Comput. Intell. Neurosci.* **2022**, *2022*, 1–12. [\[CrossRef\]](#)
77. Sklearn.Svm.SVC. Available online: <https://scikit-learn/stable/modules/generated/sklearn.svm.SVC.html> (accessed on 19 August 2022).
78. Garg, A.; Mago, V. Role of Machine Learning in Medical Research: A Survey. *Comput. Sci. Rev.* **2021**, *40*, 100370. [\[CrossRef\]](#)
79. Baygin, M.; Dogan, S.; Tuncer, T.; Datta Barua, P.; Faust, O.; Arunkumar, N.; Abdulhay, E.W.; Emma Palmer, E.; Rajendra Acharya, U. Automated ASD Detection Using Hybrid Deep Lightweight Features Extracted from EEG Signals. *Comput. Biol. Med.* **2021**, *134*, 104548. [\[CrossRef\]](#)



80. Catarino, A.; Andrade, A.; Churches, O.; Wagner, A.P.; Baron-Cohen, S.; Ring, H. Task-Related Functional Connectivity in Autism Spectrum Conditions: An EEG Study Using Wavelet Transform Coherence. *Mol. Autism*. **2013**, *4*, 1. [CrossRef] [PubMed]
81. Britton, J.; Frey, L.; Hopp, J.; Korb, P.; Koubeissi, M.; Lievens, W.; Pestana-Knight, E.; Louis, E.K.S. *Electroencephalography (EEG): An Introductory Text and Atlas of Normal and Abnormal Findings in Adults, Children, and Infants*; St. Louis, E.K., Frey, L., Eds.; American Epilepsy Society: Chicago, IL, USA, 2016; ISBN 978-0-9979756-0-4.
82. Alhassan, S.; Soudani, A. Energy-Aware EEG-Based Scheme for Early-Age Autism Detection. In Proceedings of the 2022 2nd International Conference of Smart Systems and Emerging Technologies (SMARTTECH), Riyadh, Saudi Arabia, 9–11 May 2022; IEEE: Piscataway, NJ, USA; pp. 97–102.
83. Muthukumaraswamy, S.D.; Singh, K.D. Visual Gamma Oscillations: The Effects of Stimulus Type, Visual Field Coverage and Stimulus Motion on MEG and EEG Recordings. *NeuroImage* **2013**, *69*, 223–230. [CrossRef] [PubMed]
84. Oh, S.L.; Jahmunah, V.; Arunkumar, N.; Abdulhay, E.W.; Gururajan, R.; Adib, N.; Ciaccio, E.J.; Cheong, K.H.; Acharya, U.R. A Novel Automated Autism Spectrum Disorder Detection System. *Complex Intell. Syst.* **2021**, *7*, 2399–2413. [CrossRef]
85. Acharya, U.R.; Fujita, H.; Sudarshan, V.K.; Bhat, S.; Koh, J.E.W. Application of Entropies for Automated Diagnosis of Epilepsy Using EEG Signals: A Review. *Knowl. Based Syst.* **2015**, *88*, 85–96. [CrossRef]
86. Contiki-NG · GitHub. Available online: <https://github.com/contiki-ng> (accessed on 12 September 2022).
87. Kurniawan, A. *Practical Contiki-NG*; Apress: Berkeley, CA, USA, 2018; ISBN 978-1-4842-3407-5.
88. Amrani, G.; Adadi, A.; Berrada, M.; Souirti, Z.; Boujraf, S. EEG Signal Analysis Using Deep Learning: A Systematic Literature Review. In Proceedings of the 2021 Fifth International Conference On Intelligent Computing in Data Sciences (ICDS), Fez, Morocco, 20–22 October 2021; pp. 1–8.
89. Zaidi, S.A.R.; Hayajneh, A.M.; Hafeez, M.; Ahmed, Q.Z. Unlocking Edge Intelligence Through Tiny Machine Learning (TinyML). *IEEE Access* **2022**, *10*, 100867–100877. [CrossRef]
90. Shoaran, M.; Haghi, B.A.; Taghavi, M.; Farivar, M.; Emami-Neyestanek, A. Energy-Efficient Classification for Resource-Constrained Biomedical Applications. *IEEE J. Emerg. Sel. Top. Circuits Syst.* **2018**, *8*, 693–707. [CrossRef]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.