



IJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** V **Month of publication:** May 2026

DOI: <https://doi.org/10.22214/ijraset.2026.82431>

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com

Analysis and Interpretation of Electroencephalogram (EEG) Signals for an Automated Diagnosis of Earlier Onset of Stroke using a Convolutional Neural Network (CNN)

Romain Atangana, Danwe Tsobsala, Amstrong Emini Me Zenanga, Daniel Gams Massi, Daniel Tchiotsop, Godpromesse Kenne

Abstract: *Electroencephalogram (EEG) recording is relatively safe for the patients who are in stroke disease, so it is often used to detect the occurrence of stroke in clinical practice. The objective of this paper is to apply Deep Learning method to EEG signal analysis in order to confirm the occurrence of stroke due to injuries attack. A novel approach using polynomial transform for spectral analysis of EEG signal in the Tchebychev basis is implemented to obtain spectral coefficients. Statistical metrics were extracted from these polynomial coefficients to constitute the input vector to a convolutional neural network. The model performs an accuracy of 98 % showing that the method can evaluate the condition of brain attack patients and can be a reliable tool of quasi stroke diagnosis.*

Keywords: *Electroencephalogram (EEG), Discrete Tchebychev Transform (DChT), Stroke Disease, Convolutional Neural Network (CNN), ROC curves.*

I. INTRODUCTION

A stroke is the sudden death of brain cells in a localized area which happens when the blood flow to an area of the brain is interrupted by either a blood clot or a broken blood vessel. A stroke is a medical emergency that kills many brain cells per minute and causes permanent brain damage. Depending on the region of the brain affected, a stroke may cause paralysis, speech impairment, loss of memory and reasoning ability, coma or death. A stroke is also sometimes called a brain attack or cerebrovascular accident (CVA). There are two main types of stroke: ischemic, due to lack of blood flow, and hemorrhagic, due to bleeding. About four out of every five strokes are ischemic. About one in every five stroke is hemorrhagic. After dementia, strokes are the second leading cause of disability. Disability may include loss of vision and/or speech, paralysis and confusion. Once the damage becomes clinically or radiographically apparent, proper neurological exams and imaging are very useful for detecting delayed cerebral ischemia stroke. Thus, EEG can be a useful way to detect and subsequently treat ischemia before the injury becomes irreversible. EEG is also very useful for identifying the stroke. In the operating room, EEG has an established role in identifying ischemia prior to the development of infarction during carotid endarterectomy (Foreman and Claassen 2012) [1]

Stroke is defined as a neurological deficit attributed to an acute focal injury of the central nervous system (i.e., brain, retina, or spinal cord), including ischemic stroke and hemorrhagic stroke (Sacco et al. 2013) [2]. It is the second leading cause of death and third leading cause of disability in the world (Kisa et al. 2019) [3]. Cognitive impairment is a common consequence of stroke (Jokinen et al. 2015) [4]. With excellent temporal resolution, EEG (Electroencephalogram) can reflect real-time neural electrical activity of brain. It is possible to assess stroke rehabilitation degree and associated neuro-cortical activity by EEG biomarkers (Petrovic et al. 2017; Aydin 2021) [5, 6].

Brain function is represented on an EEG by oscillations of certain frequencies, slower frequencies (typically delta (0.5-4 Hz) or theta (4-8 Hz)) are generated by thalamus and by cells in layer [II-V] of the cortex. Faster frequencies (or alpha typically (8-12 Hz)) derived from cells in layers IV and V of the cortex. (Amzica et al. 2010). All frequencies are modulated by the reticular activating system, which corresponds to the observation of reactivity on the EEG. Pyramidal neurons found in layers III, V and VI are exquisitely sensitive to conditions of low oxygen, such as ischemia, thus leading to many of the abnormal changes in the patterns seen on an EEG (Ordan 2004).

The common — observed oscillatory — waves recorded by an electroencephalogram (EEG) in normal adult human can be grouped into five main categories according to the frequency and amplitude namely: δ (0.5-4 Hz, 20-100 μ V); θ (4-8 Hz, 10 μ V); α (8-12 Hz, 20-200 μ V); β (12-30 Hz, 5-10 μ V) and γ (30-80 Hz, 2-4 μ V). Changes of brain oscillation patterns have long been implicated in the diseases of the central nervous system including ischemic stroke.

II. RELATED WORKS

The diagnosis of stroke relies predominantly on the use of neuroimaging. Early identification of stroke using Electroencephalogram (EEG) in the clinical assessment of stroke has been underutilized. Mesfer AIDuhayyimandal. (2023) «An Ensemble Machine Learning Technique for Stroke Prognosis» Computer System Science and Engineering In this work Authors proposed an ensemble voting model based on three Machine Learning (ML) algorithms.

Random Forest (RF), Extreme Gradient Boosting (XGBoost), and Light Gradient Boosting Machine (LGBM). They apply data preprocessing to manage the outliers and useless instances in the dataset. Furthermore, to address the problem of imbalanced data, they enhance the minority class's representation using the Synthetic Minority Over-Sampling Technique (SMOTE), allowing it to engage in the learning process actively. Results reveal that the suggested model outperforms existing studies and other classifiers with 96% accuracy, 97% precision, 97% Recall and 96% F1-Score. The experiment demonstrates that the proposed ensemble voting model outperforms state-of-the-art and other traditional approaches [?].

Sakai Y. and al. (2024) « Validity and Reliability of the Japanese Version of the Frontal Assessment Battery in Patients with Stroke » This study aimed to investigate the validity and reliability of the Japanese version of the Frontal Assessment Battery (FAB) in patients with stroke. The Japanese version of the FAB for dementia was modified and evaluated in 52 patients with stroke. FAB measurements were obtained twice over a 10-day period. Convergent validity was assessed using the Stroop Color Word Test (SCWT) and the Trial Making Test (TMT) part B. Internal consistency was measured using Cronbach's alpha (α). Test-retest evaluations were performed using intra-class coefficient (ICC[2.5]) measurements and limits of agreement (LOA) were calculated using the total FAB score. Results show that the total FAB score was correlated with SCWT scores for part I through IV ($r=0.70$ to 0.77) and the TMT score for part B ($p=0.53$). The LOA were -1.7 to 2.9 points. Conclusion, the Japanese version of the FAB had high validity and reliability in patients with stroke [?].

Michele Lauriola and al. (2024) « Risk of Stroke or Heart Attack in Mild Cognitive Impairment and Subjective Cognitive Impairments » Neurol. This study aimed to identify Mild Cognitive Impairment (MCI) as an alert clinical manifestation of increased probability of major acute vascular events (MVEs), such as Ischemic Stroke and Heart attack: in the longitudinal study 181 (M=81, F=100, mean age of 75.8 years) patients were enrolled and divided into three groups based on diagnosis. Subjective Cognitive Impairment (SCI), amnestic MCI Single Domain (aMCI-SD), and amnestic MCI More Domain (aMCI-MD), clinical assessment and the presence of vascular risk factors were collected. The results showed that the distribution of MVEs showed a higher incidence in the first two years of follow-up of 7.4% in SCI, 12.17% in aMCI-SD and 8.5% in aMCI-MD. Acute Myocardial infarction showed a major incidence in one year of follow-up (41%) and in two years of follow-up (29%). Also, Ischemic Stroke showed a major incidence in one year of follow-up (30%) and in two years of follow-up (40% significant difference in the progression to dementia was shown (SCI 3,75%; aMCI-SD 10,4%; aMCI-MD 37% ; p-value). Conclusion : MCI is considered an expression of the systematic activation of mechanisms of endothelial damage, representing a diagnosis predictive of increased risk of MVEs.

Pui Kit Tam and al. (2024) « Prevalence and Outcomes of Orthostatic Hypotension in Hemorrhagic Stroke Patients During Hospitalization ». This study aims to examine the prevalence of orthostatic hypotension (OH), its risk factors and potential impact in patients who were hospitalized due to hemorrhagic stroke. A retrospective analysis of in-patients records between 1 January 2021 and 30 April 2023 was conducted for patients with stroke due to intracerebral hemorrhage (ICH) or subarachnoid hemorrhage (SAH) who were referred to rehabilitation at a tertiary hospital in Singapore. OH was defined as a drop in systolic blood pressure of 20 mm Hg or diastolic blood pressure of 10 mm Hg during the sit-up test as part of the rehabilitation assessment. Additional data collection included demographic information, length of stay, anti-hypertensive medications used at the time of assessment, comorbidities, and discharge functional outcomes as measured by a modified Rankin scale. Results show that a total of 77 patients (65 [84,8%] with ICH and 12 [15,6%] with SAH) were included in the analysis. The prevalence of OH was 37,7%, a history of surgical intervention was identified as the major risk factor for the development of OH (odds ratio 4,28, 95% confidence interval 1,37 to 13,35, $p = 0,009$). There was no difference in hospital length of stay or discharge modified Rankin Score between the two groups.

Mattias Rosberg and al. 2024. In this work the goal was to examine the feasibility of detecting stroke using only brain activity data electroencephalography (EEG) machine. EEG-datasets from patients with and without stroke were collected, standardized and preprocessed using different methods. The different variation of this data was then used to train machine learning models and the ability of each trained model to identify EEG recording of stroke patients were tested. Using differences processing methods and with a sufficient amount of data, the model was able to correctly differentiate stroke and non-stroke EEG recording in the dataset with 100% accuracy [?].

Li-Min-Weng and al. 2024 [7] used EEG signal to explore the EEG characteristics of patients with post-stroke epilepsy and improve the detection rate of inter-seizure epileptiform discharges. Post-stroke epilepsy (PSE) refers to seizures that occur within a certain time after stroke in patients without a prior history of epilepsy or brain and systemic diseases. The study reveals that the epileptic discharges detected by electroencephalogram (EEG) are consistent with the lesion site of the stroke.

Benghanmetal.2024[8]assessedtheprognosticsignificanceandprevalenceofearlyelectroencephalogram(EEG)abnormalitiesinadultstrokepatients, receiving mechanical ventilation. The experiment included adult stroke patients requiring invasive mechanical ventilation, who underwent at least one intermittent EEG examination during their intensive care unit (ICU) stay. An unreactive EEG background to auditory and pain stimulations (OR 602.95% CL 2.27-15.99) was independently associated with unfavorable outcomes. An unreactive EEG predicted unfavorable outcome with a specificity of 48%.

Jin Yandal.2024[9] have identify brain damage areas for stroke rehabilitation. The EEG signals were obtained from stroke patients and Healthy subjects.

These participants divided into right-sided brain injury group, left-sided brain injury group, Bilateral brain injury group and Healthy Controls. First, the wavelet packet transform was used to perform a time-frequency analysis of the EEG signal and extracted a set of features, then to explore the nonlinear phase coupling information of the EEG signal, phase-locked value (PLV) and partial directed correlation (PDC) were extracted from the brain network and the brain network produced a second set of features noted as functional connectivity (FC) features. They fused the extracted multiple features and used the ResNet50 convolutional neural network to classify the fused multi-modal (WPT + FC) features. The classification accuracy of the model was up to 99.75%, showing that the multi-modal frequency features can be used as a potential indicator to distinguish regions of brain injury in stroke patients.

A. Recording EEG in the resting state

Yu Xia Huandal.2023[10] investigated whether resting-state EEG indicators could improve stroke rehabilitation evaluation. The participants were divided into three groups: severe, moderate and mild, then quantitative electroencephalography (QEEG) and five nonlinear parameters of resting state EEG were calculated for further analysis. Statistical tests were performed and the genetic algorithm support vector machine was used to select the best feature combination for classification. The QEEG parameters show significant differences in Delta, Alpha 1, Alpha 2 among the three groups. Regarding nonlinear parameters, ApEn, SampEn, LZ and C showed significant differences. The optimal feature classification combination rate reached 85.3%. (Sutcliffe et al. 2022)[11] explained that Surface electroencephalography (EEG) shows promise for stroke identification and outcome prediction. Authors gave answers to the questions:

- 1) Can EEG during acute clinical assessment identify Stroke versus non-stroke mimicking conditions? Ischemic versus hemorrhagic stroke? Ischemic stroke due to LVO?
- 2) Can these states be identified if EEG is applied < 6 hrs in onset?
- 3) Does EEG during acute assessment predict clinical recovery following confirmed stroke?

Although studies report important associations with EEG biomarker, further technological development and adequately powered real world studies are required before recommendations can be made regarding application during acute stroke assessment. (Ag. Lamatal.2023)[12] This study was conducted to determine the relevance of

EEG and its predictors with the clinical and stroke features. According to the findings, the type of the stroke and imaging characteristics are associated with EEG anomalies. Predictors of focal EEG slowing are NIHSS score and anterior circulation stroke. The study emphasized that EEG is a simple yet feasible investigation tool, and further plans for advancing stroke evaluation should consider the inclusion of this functional modality. (Yoom A., et al. 2021) designed and implemented a health monitoring system that can predict the precursors of stroke disease in the elderly in real time during daily walking.

First, raw electroencephalography (EEG) data from six channels were preprocessed via Fast Fourier Transform (FFT). The raw EEG power values were then extracted from the raw spectra: alpha (α), beta (β), gamma (γ)

, delta (δ) and theta (θ) as well as the low β , high β and θ to β ratio, respectively. The experiments confirm that the important features of EEG biometric signals alone during walking can accurately determine stroke precursors and occurrence in the elderly with more than 90% accuracy. (Foreman and Claassen 2012)[12].

Keser (2022)[13] summarized the use of EEG alone in post-stroke, then they demonstrated that EEG-based techniques are accessible and valuable non-invasive clinical tools in stroke neurology. Studies have also confirmed that the quantitative electroencephalographic (QEEG) parameters of stroke patients exhibit increased delta and alpha bands (Jordan 2003; Hirsch et al. 2012; Finnigan and Putten 2013)[14, 16]. In addition, Finnigan et al. (Finnigan et al. 2016; Sfa and Awb 2020)[10, 17] achieved a high classification accuracy rate between stroke patients and control groups with QEEG parameters.

Nonlinear dynamic analysis (NDA) have been widely used in the analysis of EEG signals (Stam 2005). It can better represent brain activity and provide more characteristic information in various neurological diseases. Based on single-channel EEG complexity and hemispheric dependency/connectivity between two EEG recordings/channels, Jeong (2004) reviewed EEG nonlinear changes for estimating quantitative markers of Alzheimer's disease, focused on the reduction of EEG complexity.

There are many nonlinear methods such as approximate entropy (ApEn), permutation entropy (Per-mEn), and Lempel–Ziv complexity (LZC) to measure the complexity of EEG (Aydin and al. 2018) [18].

Several studies have reported the appearance of certain attribute values in stroke patients with EEG-based stroke analysis [Finnigan, and al. 2016; Schneider et al. 2005; Varela et al. 2007; Ip, Z et al. 2019] [14,17,19,20]. For example, Simon et al. (Finnigan, Setal 2019) [21] confirmed that the main properties of EEG with respect to stroke include the generation of abnormal slow signals generated at the delta wave (δ) wave frequency range (14 Hz) with the simultaneous reduction of normal and fast activities at the alpha (α) wave frequency range (8–12 Hz). Through these experiments, we can confirm that relative delta power, delta and alpha power ratio, and the addition of delta and theta wave against alpha and beta wave ratios can be used to detect and predict stroke.

(Schneider et al. 2005) studied EEG frequency analysis and topographic maps and found an increase in large delta waves and a decrease in alpha wave activity in 17 out of 20 mild stroke patients. (Panayiotis et al. 2017) confirmed that a rhythmic and high amplitude theta wave - sus delta wave appeared in a patient with epilepsy during a stroke. (Ip et al. 2019) confirmed that brain waves in stroke patients that were measured in the cerebral cortex affected the activity and stability of the theta wave 4 of 18 and the delta wave, while the delta wave (0.1–3 Hz), alpha (7–13 Hz), beta wave (13–30 Hz), and high gamma wave (62–200 Hz) increased rapidly in the right hemisphere. Based on these studies, we believe that EEG research can help minimize medical costs and enable the early detection of stroke diseases in the elderly during their daily activities. A quick literature review found a few studies using various machine-learning techniques, including artificial neural networks (ANN), for stroke diagnosis or prediction (Shanthi, D. et al. 2009; Nwosu et al. 2019; Bentley et al. 2014; Hanifa, et al. 2010; Yu, et al. 2020). For example, (Shanthi et al. 2009) reported that an individual's risk rate for stroke can be detected using ANN based on stroke patient data. Specifically, they used the backpropagation algorithm for learning, and showed improvements in consistency and diagnostic accuracy for the prediction. (Nwosu et al. 2019) studied the analysis and prediction of risk factors associated with the onset of stroke using data mining techniques and individual patient electronic health records. According to the experimental results, the prediction accuracy of decision tree (DT) was 74.31%, Random Forest was 74.53%, and ANN algorithm was 75.02%. (Bentley et al. 2014) reported a prediction method considering CT information and clinical variables in the treatment of ischemic stroke. Based on computed tomography (CT) images of 16 ischemic stroke patients, they successfully found 9 out of 16 patients with hemorrhage symptoms using support vector machines (SVMs). (Hanifa et al. 2010) predicted and verified the risk factors of stroke by adjusting the parameter values of the SVM prediction model using various kernel functions. (Yu et al. 2020) published a study detailing a priori detection and prediction methodology for stroke diseases with machine-learning and deep-learning methodologies by collecting electromyography (EMG) biological signals from thighs and calves in real time. More specifically, they measured and collected EMG data from the left and right thighs and the calves at 1500 Hz from the healthcare device. Using those data, they achieved over 90% stroke disease prediction accuracy. (Yu et al. 2019; 2020) published an analytical study based on the decision tree methodology, which is a representative classification model of machine learning in data mining. In addition, (Yu et al. 2020) attempted to implement automatic classification and interpretation of the severity of NIHSS based on the C4.5 decision tree algorithm. By analyzing the rules on the additional operating principles provided by C4.5 decision trees, they were able to develop a novel attempt at these manual interpretations of stroke severity. However, decision trees are a predictive model algorithm by nature, which only provide partial interpretations, thus requiring in-depth analysis inherent in the data. In addition, (Amini et al. 2023) [22] conducted a study to predict stroke outbreaks based on abundant medical data on a wider range of diseases. However, such research methodologies, like prior studies, are not suitable for use in early prediction models of stroke symptoms in real time in everyday life. Because these strokes interact with various risk factors rather than with one factor, studies of stroke disease prediction using various statistical methods and machine-learning methods are needed, and they are actively underway.

III. MATERIALS AND METHOD

A. Hardware and software environment

All experiments and simulations were carried out using various tools such as the ANACONDA IDE (version 3.5) deployed on a laptop running an Intel (R) Core (TM) i5-6300U with a Processor of 2.40 GHz, 8.00 GB of RAM and Windows 10 PRO.

B. Methodology used

This work proposes a methodology for the analysis of EEG data in order to automatically diagnose stroke. As presented in the figure below, this methodology includes 04 main steps which are data acquisition, extraction of informative parameters, selection of relevant parameter, dimensionality reduction, classification and diagnosis in order to determine if a patient is affected by the stroke or not as illustrated in the following figure:

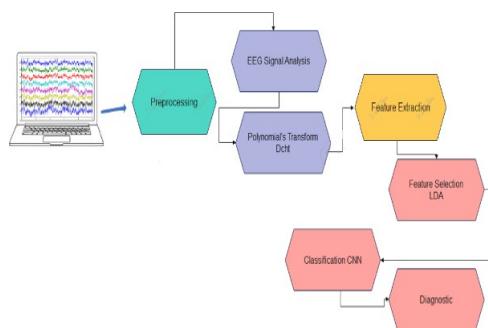


Fig.1.Synopticofthemethodologyproposedtodiagnose Stroke

C. Dataacquisitionandpre-processing

1) ParticipantsWhilerecordingtheEEGsignals

- Thepatientsitscomfortablyinachair;
- Weattachtheelectrodestodifferentplaceson the head:
 - Astickypasteisusedtoholdtheelectrodes in place;
 - Thepatientwearsacaptowhichtheelec- trodesarealreadyattached;
- The electrodes are connected by wires to the de-vicewhichamplifiesandrecordstheelectrical activity inside the brain
- Thedevice recordselectricalactivityaswavy linesproducedbyrows oftracersonamoving sheetofpaperorasanimagethatisdisplayed on a computer screen.

Fortheparticipantswererecruitedparticipantamong our students and over one year clinical study ac- cross the department of neurology of the recent Re- gionalHospitalCenterofBertouaCameroon.Par- ticipants were divided into those suffering of stroke and healthy subjects. The first group consisted of 50participantswhowereidiagnosedbyexperiencedneurologistsassufferingofstroke.Thesecondgroup was constituted of 109 healthy controls. The age groupwasrangingwithin25to80years.

Allparticipantsinthisstudyprovidedinformedcon- sentforthe collectionanduseoftheirEEGdata.Datahave been anonymized to ensure the confidentiality ofparticipants.Theyarestoredonsecureservers,with accesslimitedtoauthorizedpersonnel.

- 2) EEG recordingdataforthestrokediagnosiswerecol- lected during 5 minutes resting awake period. The ac- quisitionssystemismadeupof6- channelsdeviceas shown in figure 2 below.



Fig.2. Patientdata collection deviceinstalled

The device uses a sampling frequency of 200Hz. The electrode layout is presented in figure 3.

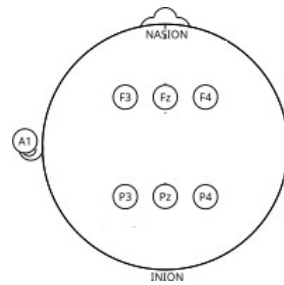


Fig.3. EEG electrodes placement for the stroke study

Participants were recruited among 109 volunteers students or Healthy controls of the Department of Computer Science of ENS Bertoua performing different imaging motor task [2] and 50 [subject 1-subject 50] patients with acute ischemic stroke aged from 25 to 80 year. These patients were diagnosed by experienced neurologists as suffering of a stroke over one year clinical study across the Department of neurology of the recent Regional Hospital Center of Bertoua-Cameroon. Participants while recording EEG signals were seated comfortably in a chair, wearing the electrodes at different places on the head with reduced 6-channels. A sticky paste is used to hold the electrodes in place and patients wore a cap to which the electrodes are already attached. These electrodes are connected by wires to the device which amplifies and records the electrical activity inside the brain.

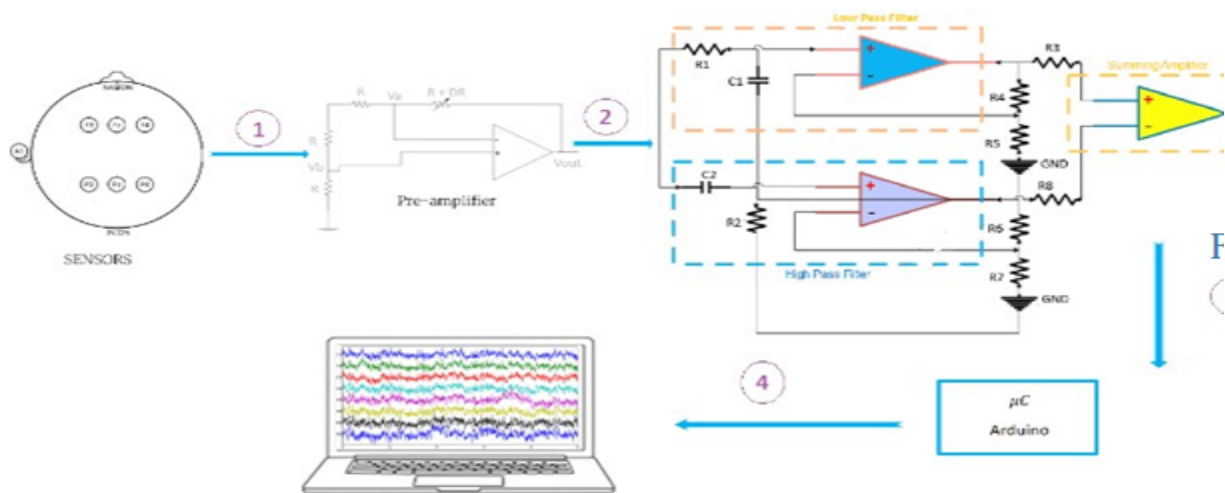


Fig.4. Synoptic scheme of the proposed automated EEG-based stroke signals acquisition

A sensor transforms a physical quantity into an electrical signal which can be interpreted by a controlled monitoring device. During this study we used cupule electrodes as sensors. These electrodes are filled with conductive battery and are stuck to the scalp by a paste and are connected to the recording device by means of simple insulated wires.

A filter is an electronic circuit that performs a signal processing operation. In other words it attenuates certain components of a signal and lets others pass.

The low amplitudes of EEG signals require amplification of the latter through a chain of pre-amplifiers and amplifiers with high gain. In the pre-amplifier part, we used the AD8221 with a gain of 10. Subsequently we carried out an amplification with a high gain $G = 40$.

• Microcontroller

Arduino is an open source programmed electronics platform which is based on μC (microcontroller) board and Software which is a true integrated development environment for writing, compiling, and transferring the program to the microcontroller board.

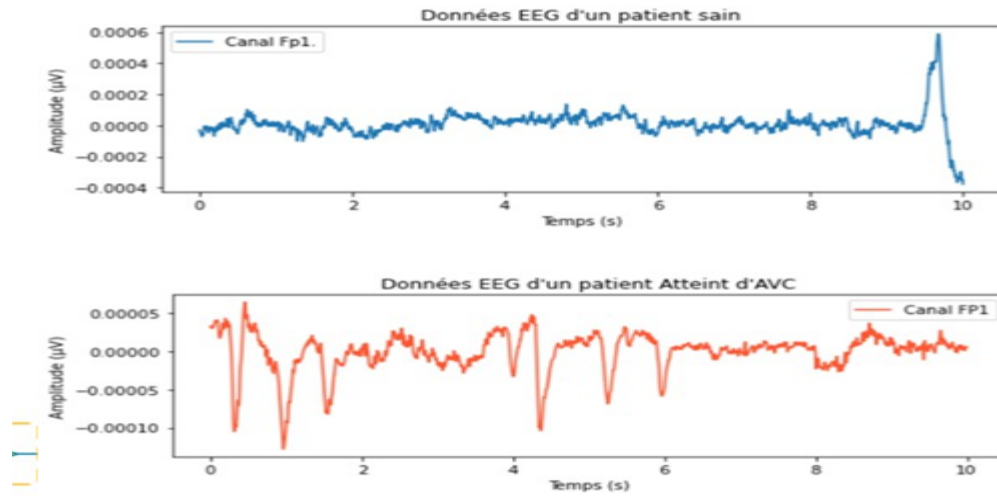


Fig.5.ExampleofEEGsignalsfromcanalFP1 for Healthypatient(blue)andastrokepatient:

Mathematical Principles Spectral Analysis

Theintegraloftheproductoffunctionsonabounded intervalisthesimplestscalarproductoffunctions:

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx \tag{1}$$

Generally,itispossibletointegratea"weightfunction" $W(x)$ intotheintegral(ontheintegrationinterval] a,b], $W(x)$ musthavefiniteandstrictlypositive values,andtheintegral oftheproductoftheweightfunctionbyapoly- mialmustbefinite;thelimits] a,b [canbeinfinite):

$$\langle f, g \rangle = \int_a^b f(x)g(x)W(x)dx \tag{2}$$

In this definition of the dot product, two functions areorthogonaltoeach otheriftheir dot productiszero (justastwovectorsareorthogonal(perpendicular)iftheir dot productiszero).Wethenintroducetheassociatednorm.

- Jacobi Polynomials $P_k^{(\alpha, \beta)}(x)$

$$a = -1, b = 1, \omega(x) = (1-x)^\alpha(1+x)^\beta, \quad \text{with } \alpha > -1 \text{ and } \beta > -1 \tag{3}$$

- Laguerre Polynomials $L_k^\alpha(x)$

$$a = 0, b = +\infty, \omega(x) = e^{-x}x^\alpha, \quad \text{avec } \alpha > -1 \tag{4}$$

- Hermite Polynomials $H_k(x)$

$$\int_{-\infty}^{+\infty} f(x)e^{-x^2}dx \tag{5}$$

For this study, we focus on Chebyshev polynomials, which represent a specific case of Jacobi polynomials [26].

Generalities on Chebyshev Polynomials

The application $\theta \rightarrow \cos \theta$ is a continuous and bijective function from $[0, \pi]$ on $[-1, 1]$ which means that to each continuous function F on $[0, \pi]$ is uniquely associated the continuous function f on $[-1, 1]$ by the relation:

$$F(\theta) = f(x) \text{ where } x = \cos \theta \tag{6}$$

Then,

$$\int_0^\pi |F(\theta)|^2 d\theta = \int_{-1}^1 \frac{|F(\theta)|^2}{\sqrt{1-x^2}} dx \text{ with } x = \cos \theta \tag{7}$$

We deduce from this that the application which from a function F of $[0, \pi]$, corresponds to the function f of $[-1, 1]$ extends into an isometric bijection from the Hilbert space $L^2([0, \pi], d\theta)$ on the Hilbert space $L^2([-1, 1], \omega(x) dx)$ similar square functions on $[-1, 1]$ for the density measure

$\omega(x) dx$ compared to the following measure:

$$\|f\|_\omega^2 = \int_{-1}^1 \frac{|f(x)|^2}{\sqrt{1-x^2}} dx < \infty \tag{8}$$

The scalar product is noted $\langle f, g \rangle_\omega = \int_{-1}^1 f(x)g(x)\omega(x) dx$:

$$\langle f, g \rangle = \int_{-1}^1 \frac{f(x)g(x)}{\sqrt{1-x^2}} dx \tag{9}$$

$f, g \in L^2([-1, 1], \omega(x) dx)$ We are in a position of applying the results of paragraph 1. The family $1, x, \dots, x^n$ is total in the space $L^2([-1, 1], \omega(x) dx)$ and the orthogonalization process of Gram-Schmidt makes it possible to produce a Hilbertian basis [27]. However, the isomorphism highlighted above allows us to quickly find this Hilbertian base, in fact the functions defined by: $\Phi_n(\cos \theta) = \cos(n\theta)$ form an orthogonal basis of $L^2([0, \pi], d\theta)$ and we have:

$$\int_0^\pi \Phi_n(\theta)\Phi_m(\theta) d\theta = \begin{cases} \frac{\pi}{2} \delta_{nm}, & \sin n, m \neq 0 \\ \pi \delta_{00}, & \sin n = 0 \end{cases}$$

It follows that the functions $T_n(x) = \cos(n \cdot \arccos(x))$ form a basis of $L^2([-1, 1], \omega(x) dx)$

$$\langle T_n, T_m \rangle_\omega = \begin{cases} \frac{\pi}{2} \delta_{nm}, & \sin n, m \neq 0 \\ \pi \delta_{00}, & \sin n \geq 0 \end{cases} \tag{11}$$

We know that $\cos(n\theta)$ is expressed by a polynomial in $\cos(\theta)$, which results in the fact that $T_n(x)$ is a polynomial of degree n in x , called the Chebyshev polynomial. We know that $\cos(n\theta)$ is expressed by a polynomial in $\cos(\theta)$, which results in the fact that $T_n(x)$ is a polynomial of degree n in x , called the Chebyshev polynomial.

$$\lim_{n \rightarrow \infty} \frac{1}{\omega} \left| \langle f, T_0 \rangle \right| + \sum_{k=1}^{\infty} \frac{1}{\omega} \left| \langle f, T_k \rangle \right| = 0 \quad (12)$$

and:

$$\|f\|_{\omega}^2 = \frac{1}{\pi} \left| \langle f, T_0 \rangle \right|^2 + \sum_{k=1}^{\infty} \frac{1}{\pi} \left| \langle f, T_k \rangle \right|^2 = 0 \quad (13)$$

We can deduce the following minimization property:

- Parity. For all $n \in \mathbb{N}$, we have:

$$T_n(1) = 1 \quad (14)$$

$$T_n(-x) = (-1)^n T_n(x) \quad (15)$$

- Recurrence relation and trigonometric identity:

$$\cos((n+1)\theta) + \cos((n-1)\theta) = 2\cos(\theta)\cos(n\theta) \quad (16)$$

We deduce that the Chebyshev polynomials satisfy the following recurrence relation:

$$T_{n+1}(x) + T_{n-1}(x) = 2xT_n(x) \quad (17)$$

The recurrence relation for the Chebyshev polynomials is $T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$, with initial conditions $T_0(x) = 1$ and $T_1(x) = x$. This relationship shows that the coefficient of x^n in $T_n(x)$ is 2^{n-1} . It also allows us to calculate T_n by recurrence.

- Generating function. We are looking for a function of two variables (x, z) in such that:

$$G(x, z) = \sum_{n=0}^{\infty} T_n(x) z^n \quad (18)$$

Knowing that $x = \cos(\theta)$, we have:

$$G(x, z) = \sum_{n=0}^{\infty} \frac{e^{in\theta} + e^{-in\theta}}{2} z^n, \quad |z| < 1 \quad (19)$$

- Differential equation. For all $n \in \mathbb{N}$, the Chebyshev polynomials T_n satisfy the following second order differential equation: $(1-x^2)y'' - xy' + n^2y = 0$, which results in the fact that $T_n(x)$ is a polynomial of degree n in x , called the Chebyshev polynomial.

$$(1-x) \frac{2d^2 T_n(x)}{dx^2} - x \frac{dT_n(x)}{dx} + n^2 T_n(x) = 0 \quad (20)$$

- Roots and extrema of T_n . The roots of T_n . Are easily calculated . Indeed , since $\cos(nx)=0$ if $x = \frac{(2k-1)\pi}{2n}$

$$x_k = \cos \frac{(2k-1)\pi}{2n} \quad \text{with } 1 \leq k \leq n \quad (21)$$

Are the roots of T_n . They are all real, simple, distinct and belong to the interval $[-1;1]$. On the other hand, the relationship

$$T_n'(x) = \sqrt{\frac{n}{1-x^2}} \sin(n \arccos(x)) \quad (22)$$

Show that $T_n'(x) = 0$ if $x = x_k' = \cos \frac{k\pi}{n}$ where $-1 \leq x_k' \leq 1$. One deduces rapidly that T_n reaches its extrema at $[-1;1]$, at the points x_k' and that at these points

$$T_n(x_k') = \cos(k\pi) = (-1)^k, \quad 1 \leq k \leq n \quad (23)$$

Therefore,

$$T_n(x_0') = T_n(1) = 1 \quad (24)$$

$$T_n(x_n') = T_n(-1) = (-1)^n \quad (25)$$

We can notice that the roots of T_n are symmetrical with respect to 0 and that for the interval $[-1;1]$. These roots are often used as interpolation points for continuous functions on $[-1;1]$

- Minimization properties Chebyshev polynomials play an important role in the- Where the x_k are the $L+1$ zeros of $T(L+1)(x)$. Thus, the discrete Chebyshev transform (DChT) is given by: ory of the approximation. This is because, as Chebyshev showed, these are the polynomials which deviate the least from zero on the segment $[1;1]$. In other words, if we designate by P_n the set of polynomials of degree x^n we have:

$$\forall q \in P_n \quad \sup_{|x| \leq 1} |q(x)| \geq 2^{n-1} \quad \sup_{|x| \geq 1} T_n(x) = 2^n - 1 \quad (26)$$

The expansion of the signal in the Chebyshev basis

The polynomial expansion of a signal $S(x)$ of order L in the orthogonal basis of Chebyshev is given by the equation (which represents the DChT inverse (I-DChT)):

$$S(x) = \sum_{k=0}^L a_k T_k(x) \quad (27)$$

This sequence of coefficients a_k for $k=0, 1, \dots, L$ constitutes the spectral coefficients of decomposition or projection of the signal:

$$a_k = \frac{\langle S(x), T_k(x) \rangle}{\langle T_k(x), T_k(x) \rangle} = \frac{\int_{-1}^1 S(x) T_k(x) \sqrt{\frac{dx}{1-x^2}}}{\int_{-1}^1 T_k(x) \sqrt{\frac{dx}{1-x^2}}} = \frac{1}{2} \int_{-1}^1 S(x) T_k(x) \sqrt{\frac{dx}{1-x^2}} \quad (28)$$

Obtaining numerical values of coefficients during polynomial expansion requires the evaluation of at least one integral. In order to better evaluate this integral, several methods have been proposed. Among these methods, the Gaussian quadrature method turns out to be very popular [28].

Therefore, the evaluation of the decomposition coefficients a_k is carried out by exploiting the Gauss-Lobatto quadrature method with a quadrature of order $L+1$. The Gauss-Lobatto method concerns the use of non-regularly spaced abscissa of the integration domain. In short, the Gauss-Lobatto method stipulates that for a quadrature of order L and a family of orthogonal Chebyshev polynomials $T_k(x)$ in the interval $[-1; 1]$, in respect to the weight function $\omega(x)$, the following approximation is made.

$$\int_{-1}^1 \frac{S(x) T_k(x) dx}{\sqrt{1-x^2}} = \frac{\pi}{M+1} \sum_{j=0}^M S(x_j) T_k(x_j) \quad (29)$$

Where the x_j are the $L+1$ zeros of $T(L+1)(x)$. Thus, the discrete Chebyshev transform (DChT) is given by:

$$a_0 = \frac{1}{M+1} \sum_{j=1}^{M+1} \psi_j = \frac{1}{M+1} \sum_{j=1}^{M+1} \frac{(2i-1)\pi}{2(M+1)} \quad (30)$$

$$a_k = \frac{1}{M+1} \sum_{j=1}^{M+1} \psi_j = \frac{1}{2(M+1)} \sum_{j=1}^{M+1} \frac{(2i-1)\pi}{2(M+1)} \quad (31)$$

With $k=1, 2, \dots, M$ and this sequence of coefficients a_k constitutes the coefficients spectral decomposition of the signal S in the Chebyshev basis.

- Extraction of informative parameters

The extraction of informative parameters or characteristic features is an important step in the process of classifying EEG signals. Feature extraction consists of these set of techniques processing EEG signal to obtain quantitative characteristics (features) that are proven or potential indicators of the presence or absence of a given phenomenon in the cerebral activity [7]. The goal of extraction of parameters is to transform larger raw data into a feature vector. The extraction of statistical and discriminative parameters was carried out on 04 different signals, namely:

the signal of the spectral coefficients, its absolute value, the original signal and the reconstructed signal. The following presents some parameters extracted from these different signals (the functions represented here are functions

- Minimum returns the minimum number of elements in an array and is calculated by the following function.
- Maximum returns the maximum elements of an array and is calculated by the function L.
- The mean returns the average of the elements of A along the first dimension of the array whose size is not equal to 1.
- The standard deviation (Std) returns the standard deviation of elements of A along the dimension of the array whose size is not 1.
- By default, the standard deviation is normalized by N-1, where N is the number of observations. The variance (var) which takes into account the dispersion of all the values in a set of data.
- The Kurtosis is the length of the tail of a signal distribution, or equivalently, how prone the signal is to outlier. The development of default scan can increase the number of outliers, and therefore increase the value of the Kurtosis metric. Its formula is:
- The number of zero crossings (nzc) is the number of times the signal crosses zero, furthermore, it is the number of negative variables in a set.
- The RMS returns the root mean square (RMS) value of the input X.
- The magnitude peak is the ratio of peak amplitude / RMS. Returns the ratio of the largest absolute value in X to the effective value (RMS) of X. Calculate the peak-magnitude-RMS ratio of X, dimension.
- The impulse factor compares the height of a peak to the average level of the signal.

$$X_{jeF} = \frac{X_{pe}}{N \sum_{j=1}^N |X_{je}|} \tag{33}$$

- The crest factor is the maximum value divided by the RMS. Defaults often first manifest themselves as changes in the peak of a signal before manifesting themselves in the energy represented by the signal's root mean square. Crest factor can provide early warning for faults when they are developed for the first time :

$$X_{ieF} = \frac{X_{pe}}{N \sum_{j=1}^N |X_{je}|} \tag{34}$$

- The form factor is the RMS divided by the average of the absolute value. The form factor depends on the shape of the signal while being independent of the dimension of the signal.

$$X_{ieF} = \frac{X_{rms}}{N \sum_{j=1}^N |X_{je}|} \tag{35}$$

- Higher order statistics provide insight into the system behavior through the fourth moment (Kurtosis) and third moment (asymmetry) of the vibrations signal.
- PRD (percent root square difference) is somehow translate the percentage of the normalized relative error into energy and makes it possible to quantify the quality of the reconstructed signal:

$$\text{PRD} = \frac{\sum_n^s (S_n - \tilde{S}_n)^2}{\sum_m S^2} \quad (36)$$

The number of zero crossing (n_{zc}) is the number of times the signal crosses zero. Furthermore, it is the number of negative variable in a set.

- Dimensionality reduction

Feature extraction reduces the amount of data by using a vector of features.

The aim of dimensionality reduction in machine learning is to reduce the dimension of a d -dimensional dataset by projecting it onto a k -dimensional subspace (where $k < d$) in order to eliminate redundant information while retaining relevant information about the original data. With this in mind, linear and non-linear transformations were performed using linear discriminant analysis (LDA). According to Subasi [29] LDA, also known as Fisher's discriminant analysis, is a supervised algorithm that calculates the linear discriminant by maximizing the distance between classes and reducing the distance within classes.

In contrast, LDA attempts to explicitly represent the distinction between class labels in the data. Once the procedure is complete, each class will have a normal distribution of discriminant parameters. The projection can be represented using a single matrix equation, such as:

$$Y = X * W \tag{37}$$

Where X and Y are the new and old features respectively, and W is the projection matrix of size $d \times k$ formed by the vectors $S_w^{-1} S_b S_b$ where S_w and S_b are respectively the intra-class and inter-class dispersion matrices, and are calculated as [30].

$$S_w = \sum_{i=1}^c S_i = \sum_{i=1}^c \sum_{x_k \in C_i} (x_k - \mu_i)(x_k - \mu_i)^T \tag{38}$$

$$S_b = \sum_{i=1}^c n_i (\mu_i - \mu)(\mu_i - \mu)^T \tag{39}$$

The Fisher criterion that LDA aims to maximize is:

$$\frac{w^T S_b w}{w^T S_w w} \tag{40}$$

- Classification using a Convolution Neural Network

The classification of data into feature spaces is frequently necessary in biomedical signal processing, particularly when using EEG signals [31].

Classification is a method that offers the possibility of acquiring knowledge using the model or transfer function between feature vectors and labels. Given the interest in subject diagnosis, once labels have been obtained for each feature vector, it is necessary to combine the labels associated with a specific subject to provide a subject label. Classifiers can be classified into two categories depending on the type of output provided: informative or discriminative. According to González [32] informative classifiers acquire knowledge about class models. Thus, in order to classify a vector of characteristic features, it is necessary to calculate the probability of each class and choose the most probable.

The aim of this study is to set up a convolutional neural network that will enable EEG data from stroke patients and healthy patients to be classified. To do this, we use one-dimensional convolution layers because they are suitable for time series analysis, natural language processing (NLP) for sequences of words or characters, detection of patterns in sequences, etc. We also use one-dimensional convolution layers for time series analysis [33]. In terms of operation, it applies a kernel (filter) to a sliding window along the one-dimensional dimension of the sequence to extract local features.

Convolutional Neural Network architecture: The neural network used to develop this binary classification system follows the architecture shown in the figure below.

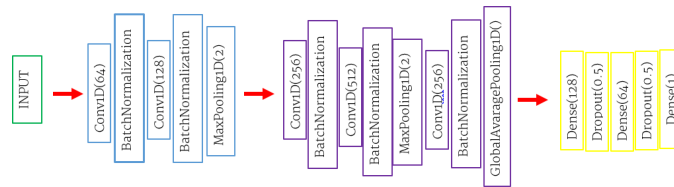


Fig.6.General architecture of the Convolutional Neural Network.

This model is made up of 17 hidden layers divided between the one-dimensional convolution layer, the Batch Normalization, Max Pooling, Global Average Pooling, Dense and Dropout layers.

Description of the Architecture The proposed network is based on a four (04) part architecture. Each part can be made up of several types of neurons, with the exception of the input layer, which contains only the input matrix (the data).

- **The Input Layer:** Defines the form of the input image. Here is the length of these sequence or the number of timesteps, and the number of features at each time step.
- **1D Convolutional Layer:** Apply a large number of convolutional filters (64 filters in the case of the first convolutional layer) to the input image with a kernel size of 3. This allows local pixels to be detected in the image. In all the rest of the neural network, for layers of the same type, we used a kernel size of 3, an activation function relu and a padding value defined as same ;
- **Batch Normalization layers:** A Batch Normalization layer normalizes the activations of a previous layer by recentering and resizing them. More specifically, for each training batch, the Batch Normalization layer
 - Normalises the activations so that they have a mean of zero and a variance of one.
 - Affine Transformation: After normalisation, the layer applies a linear transformation (affine) to allow the model to readjust the data if necessary.
- **1D Pooling Layer:** Reduces the dimensionality of data by taking the maximum from a window of size 2. This helps to reduce the size of features while retaining the most important information.
- **Dense layers:** A dense layer is a fully connected layer. Each neuron is connected to all the neurons in the previous layer. It uses the ReLU function [13] function for non-linearity.
- **Global Average Pooling layers:** A Global Average Pooling layer takes as input a feature map produced by a convolutional layer and calculates the average of each feature map. Unlike a dense layer, which treats each element of the feature map individually, Global Average Pooling reduces the entire feature map to a single number.
- **Dropout layers:** These are a regularisation technique used in neural networks to reduce the risk of overfitting (overfitting). During training, the Dropout layer randomly "deactivates" (or "zeroes out") a certain percentage of neurons in a given layer at each iteration. This means that the 'deactivated' neurons do not participate in the training phase for that specific iteration.
- **The Output Layer:** This layer has a single neuron with a sigmoid activation function, which is appropriate for a binary classification problem. The sigmoid function transforms the output into a probability between 0 and 1, indicating the probability of the positive class (sick or not sick).

• **Error reduction: Learning by retropropagation**

Gradient descent learning:

Output layer: Error-guided correction between desired and actual output

Hidden layers: Sensitivity correction (influence of the neuron on the error output layer).

Neuron output values: data x

$$y_j^t = f \left(\sum_{i=1}^R \omega_{ij} x_i^t + \omega_{j0} \right) \quad (41)$$

Where is the neuron activation function:

$a_i^t = \sum_{j=1}^R \omega_{ij} v_j^t + \omega_{i,0}$ is the cumulative sum of the input neurons;

$\omega_{i,j}$ weight of the link connecting neuron j to neuron i in the previous layer;

$\omega_{i,0}$ the bias of neuron i ;

v_j^t output of neuron j in the previous layer for input data X ;

;

R number of neurons on the previous layer.

3. Output error to output layer: $A = \{ \{x^t, r^t\} \}^N$ $t=1$ with

$$r^t = \begin{matrix} r_1^t \\ r_2^t \\ r_3^t \\ \vdots \\ r_k^t \end{matrix} = \begin{matrix} x_1^t & \cdots & x_n^t \end{matrix}^T$$

$r_i^t = 1$ six $t \in C_j$ and $r_i^t = 0$

Observed error for data x^t on neuron j of the output layer

Observed quadratic error for the data x^t on all K neurons in the output layer (one neuron per class):

$$E^t = \frac{1}{2} \sum_{j=1}^k (e_j^t)^2 \tag{42}$$

The average error observed for the dataset N is:

$$E = \frac{1}{N} \sum_{t=1}^N E^t \tag{43}$$

4. Correction of the error for the output layer:

Weight correction by gradient descent of the root mean square error:

$$\Delta \omega_{j,i} = - \eta \frac{\partial E}{\partial \omega_{j,i}} = - \sum_{t=1}^N \frac{\partial E^t}{\partial \omega_{j,i}} \tag{44}$$

$$\Delta \omega_{j,i} = - \frac{\eta}{N} \sum_{t=1}^N \frac{\partial E^t}{\partial \omega_{j,i}} \tag{45}$$

The error of neuron j depends on the neurons in the previous layer.

**5. Learning for the output layer:
Adjustment of output layer weights**

$$\Delta \omega_{j,i} = - \frac{\eta}{N} \sum_{t=1}^N \frac{\partial E^t}{\partial \omega_{j,i}} = - \frac{\eta}{N} \sum_{t=1}^N \frac{\partial E^t}{\partial e^t} \cdot \frac{\partial e^t}{\partial y_i^t} \cdot \frac{\partial y_i^t}{\partial a_i^t} \cdot \frac{\partial a_i^t}{\partial \omega_{j,i}} \quad (46)$$

$$\Delta \omega_{j,i} = \frac{\eta}{N} \sum_{t=1}^N e_i^t y_i^t (1 - y_i^t) y_i^t \quad (47)$$

$$\Delta \omega_{j,0} = \frac{\eta}{N} \sum_{t=1}^N \frac{\partial E^t}{\partial \omega_{j,0}} = \frac{\eta}{N} \sum_{t=1}^N \frac{\partial E^t}{\partial e^t} \cdot \frac{\partial e^t}{\partial y_i^t} \cdot \frac{\partial y_i^t}{\partial a_i^t} \cdot \frac{\partial a_i^t}{\partial \omega_{j,0}} = \frac{\eta}{N} \sum_{t=1}^N e_i^t y_i^t (1 - y_i^t) \quad (48)$$

6. Deltarule:

Let δ_i^t which correspond to the local gradient of neuron j for the input data X.

$$\delta_j^t = e_j^t \cdot y_j^t \cdot (1 - y_j^t) \cdot y_j^t \quad (49)$$

So,

$$\Delta \omega_{(j,i)} = \frac{\eta}{N} \sum_{t=1}^N \delta_j^t y_i^t \quad (50)$$

and,

$$\omega_{(i,0)} = \frac{1}{N} \sum_{j=1}^N \delta_j E = \frac{1}{2} \sum_{j=1}^2 (e_j) \quad (51)$$

7. Correction of the corresponding error layer :
The gradient of the hidden layer error is:

$$\frac{\partial E^t}{\partial \omega_{(i,i)}} = \frac{\partial E^t}{\partial y_i^t} \cdot \frac{\partial y_i^t}{\partial d_i^t} \cdot \frac{\partial d_i^t}{\partial \omega_{(i,i)}} \quad (52)$$

Only $\frac{\partial d_i^t}{\partial \omega_{(i,i)}}$ changes while $\frac{\partial y_i^t}{\partial d_i^t}$ and $\frac{\partial E^t}{\partial y_i^t}$ are the same as on the output layers. The error for a neuron in the hidden layer depends on the error of the K neurons in the next layer (backpropagation of the error).

$$\omega_{(i,i)} = \frac{1}{2} \sum_{j=1}^2 (e_j) \quad (53)$$

$$\frac{\partial E^t}{\partial y_i^t} = \frac{\partial}{\partial y_i^t} \left[\frac{1}{2} \sum_{j=1}^2 (e_j^2) \right] = \sum_k e_k^t \cdot \frac{\partial e_k^t}{\partial y_i^t} \quad (54)$$

$$\frac{\partial E^t}{\partial y_i^t} = \sum_k e_k^t \cdot \frac{\partial e_k^t}{\partial d_j^t} \cdot \frac{\partial d_j^t}{\partial y_i^t} = \sum_k e_k^t \cdot \frac{\partial [r^t - y^t]}{\partial d_j^t} \cdot \frac{\partial \sum_k \omega_{(k,i)} y^t + \omega_{(k,0)}}{\partial y_i^t} \quad (55)$$

$$\frac{\partial E^t}{\partial y_i^t} = \sum_k e_k^t \cdot [-y_k^t (1 - y_k^t)] \cdot \omega_{(k,i)} \quad (56)$$

$$\delta_k^t = e^t \cdot [-y^t (1 - y^t)] \quad (57)$$

$$\frac{\partial E^t}{\partial y_i^t} = - \sum_k \delta_k^t \cdot \omega_{(k,i)} \quad (58)$$

8. Correction of the corresponding error:

$$\frac{\partial E^t}{\partial \omega_{(i,i)}} = \frac{\partial E^t}{\partial y_i^t} \cdot \frac{\partial y_i^t}{\partial d_i^t} \cdot \frac{\partial d_i^t}{\partial \omega_{(i,i)}} = \sum_k \delta_k^t \cdot \frac{\partial \omega_{(k,i)}}{\partial \omega_{(i,i)}} \quad (59)$$

$$\frac{\partial d_i^t}{\partial \omega_{(i,i)}} = e^t (1 - y^t) \cdot \sum_j \frac{\partial y_j^t}{\partial \omega_{(i,i)}} \cdot \omega_{(j,i)} \quad (60)$$

$$\begin{aligned}
 (i,j) &= -\eta \cdot \frac{\partial L}{\partial (i,j)} = -\eta \sum_{t=1}^T \frac{\partial L}{\partial (i,j)} = \eta \sum_{t=1}^T \delta_i \cdot v_j \\
 & \quad (61)
 \end{aligned}$$

IV. RESULTS

1) Acquired signal

1.

Figure(4) shows the EEG signals recorded on the same

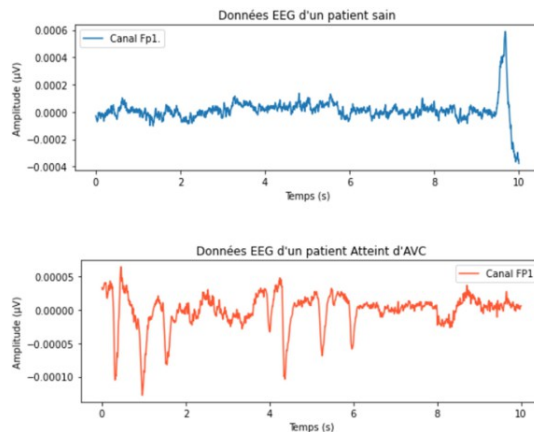


Fig.7.Signal pattern of healthy patients (blue) and stroke patients (red)

For these signals, we have a duration of 10 seconds each for the Fp1 channel to make it easier to see the signal. In reality, the length of the signal for processing will be 120 seconds on data from 25 EEG channels, which will provide enough information to help characterize a signal.

2) Spectral analysis results

This stage, which follows data acquisition and pre-processing, was implemented by interpolating the signals. Their frequency distributions are represented in terms of Chebyshev decomposition spectral coefficients.

3) Correction of the corresponding error:

Figures(5) and (6) show examples of signal decomposition for healthy and stroke patients respectively.

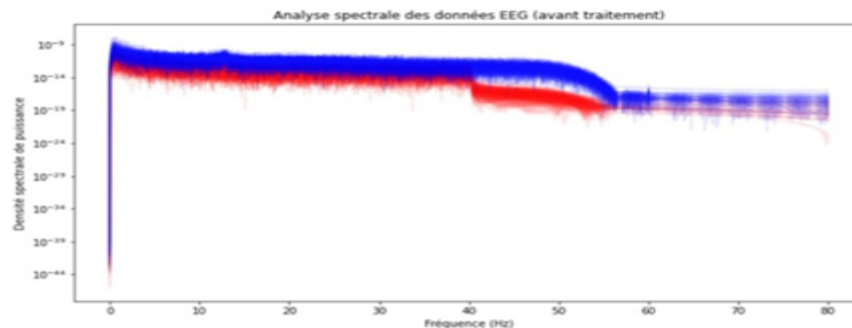


Fig.8.Representation of the original signal

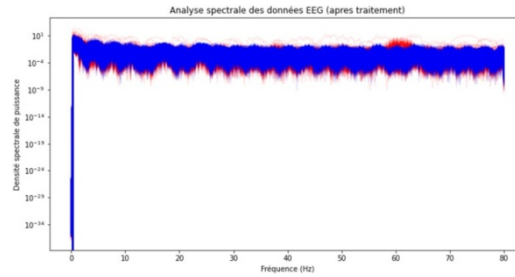


Fig.9.Representationofthereconstructed signal

The difference between the original signal and there- constructed signal is clearly visible. The blue and red curves represent the signal from healthy patients and the signal from stroke patients respectively. One observation can be made: the signal reconstruction error rate after using DChT is less than 0.0025, which proves that the reconstructed signal after interpolation is almost equal to the original signal.

We can also see that the amount of energy extracted in the reconstructed signal is virtually the same as that extracted in the original signal.

2D projection of the relevant extracted features.

To analyse a signal, one could first apply the Chebyshev transform to obtain a compact and smooth representation of the signal. Then, the resulting coefficients could be analysed by parameter extraction to extract statistical features that can be used for various signal analysis tasks such as classification, pattern detection, or anomaly detection.

The following figures show the 2D projection of the relevant extracted features from the 2 types of patients:

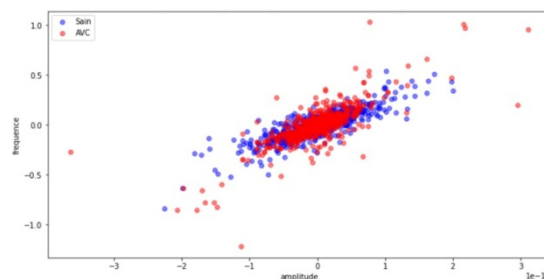


Fig. 10. View of extracted characteristics of healthy and stroke patients

The blue and red dots show the extracted characteristics of healthy and stroke patients respectively. The isolated blue and red dots show the respective characteristics of healthy and stroke patients. It is these characteristics that will enable us to facilitate decision-making. This observation shows that healthy and stroke patients are easily distinguishable, demonstrating that they are informative, representative and discriminative.

Learning outcomes

Once the features have been extracted, the data is ready to be fed into the neural network shown in Figure (3).

The neural network trained over 20 epochs produces an accuracy of 98.87 on the training data and 97.12 on the test data. As the following learning curves show, the model is reliable because it avoids overlearning.

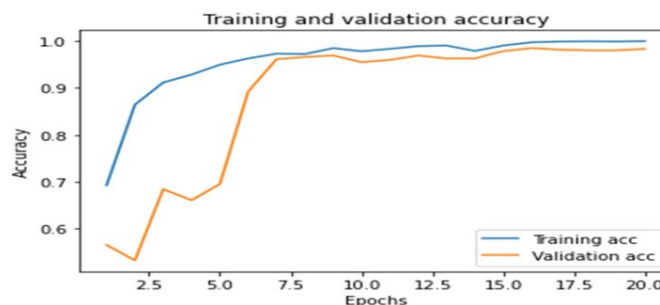


Fig. 11. Accuracy learning curves as a function of the number of

We can also see that the learning process went well, with decreasing error curves reaching values of less than 0.12.

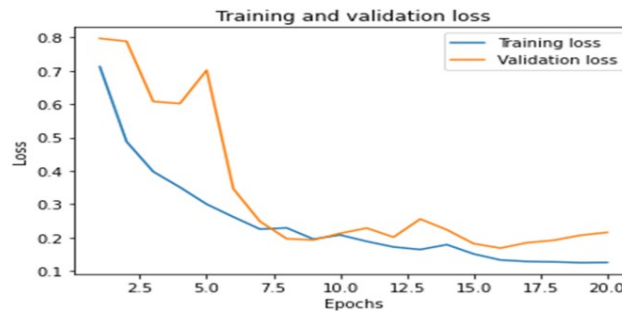


Fig. 12. Diagram of the loss learning curve as a function of the number of

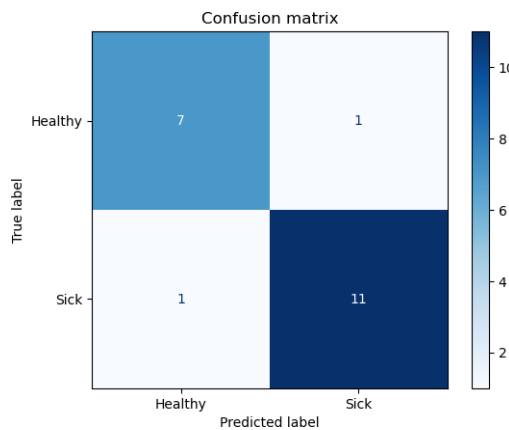
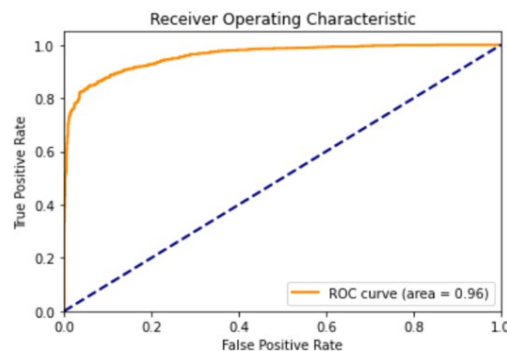


Fig. 13. Confusion matrix for prediction.

In the following Figure (10), it is easy to see that the AUC is well above 0.5 (reaching 0.96), which attests to the discriminating ability of the classifier.



r. Fig. 14. Representation of the ROC curve

versus healthy controls. The performance of classification between stroke patients and healthy controls demonstrated that spectral coefficients of the Tchebychev polynomials transform were effective features to classify stroke patients and healthy controls.

In comparison with previous studies on the diagnosis of stroke through EEG signals, there were some similar results. The table below shows the comparative study.

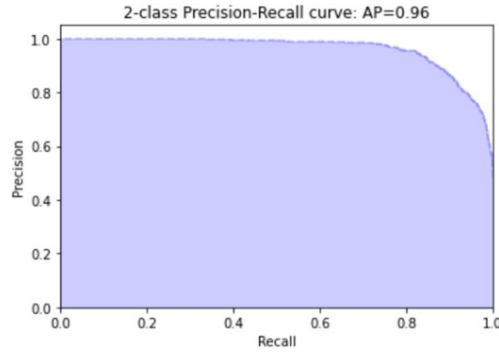


Fig. 15. Representation of the precision curve

V. DISCUSSION

Our results demonstrate that nonlinear properties of the EEG signals can be a useful way to detect modifications of brain state during stroke disease. The work in this study makes two contributions. The first contribution is the establishment of correlations between EEG signals with the clinical states of stroke patients through statistical metrics extracted from the spectral coefficients of Discrete Chebyshev Transform (DChT). The second contribution is the designing of a classifier for the detection of stroke patients. Through correlation analysis, it has been found that the occurrence of stroke has activity changes in the lesional and contralesional hemispheres, and analyses tended to late significant changes to stroke onset across different statistical metrics. EEG-based technologies are particularly promising as diagnostic or predictive biomarkers for individual stroke patients. In clinical diagnosis, a test with high sensitivity is necessary for ruling out disease. Sensitivity of 80.88% in this work refers to the ability of the analysis to correctly detect patients with stroke who do have, showing that the convolutional neural network (CNN) classifier used improved the detection of patients with stroke.

Table 1. Summary of analysis methods and results

| Authors | Method of Analysis and Extraction of patterns | Dimensionality reduction | Classifiers | Accuracy |
|-------------------------------|---|--------------------------|--|---------------|
| Zhang S. et al. (2022) | Temporal-Spectral Distance | | Weighted Convolutional Siamese Network | 66%-72.8% |
| Lin P. J. et al. (2022) | Biomarker and Scores | | CNN | 91% |
| Choi Ya et al. (2021) | EEG Data | | CNN | 94% |
| Al Arfaj A. A. et al. (2022) | | | Transfer Learning Technique | 96.5% |
| Xu. F. et al. (2023) | MST-CPS | | Clustering: K-means | 89% |
| Samar Bouazizi et al. (2024) | D-ESNLIME and ELIS | | Novel Diversified Echo-State Network | 95% |
| P. Nancy et al. (2023) | DKELM-AS | | FHT | 95.2% |
| Aktham Sawan et al. (2022) | BiGRU+HS+MVO | | CNN | 99.99% |
| E. W. Guntari et al. (2020) | Wavelet ^o CNR | | Genetic Algorithm | 72.22%-90.00% |
| Choi, Y. A. et al. (2021) | LSTM | | CNN | 94.00% |
| Choi, Y. A. et al. (2021) | FFT | | Random Forest | 92.51% |
| Mandeep Kaure et al. (2022) | LSTM, BiLSTM, GRU | | FFNN | 83%-95.6% |
| Selon Rohan Kalahastri et al. | Power Spectral Density | | ----- | 94.4%-100% |
| Li-Min-Weng (2024) | Spectral Analysis | | ----- | 92% |

| | | | | |
|-------------------------|---|-----|---|------------|
| Sutcliffe et al. (2022) | Spectral Analysis | | CNN | |
| Yu Xia Hu et al. (2023) | Spectral Analysis (QEEG parameters) | | CNN | 85.3% |
| Benghanem et al. (2024) | Modified Rankin Scales | | Multivariable Regression Analysis (MVR) | Spof 48% |
| Jin Y. et al. (2024) | Multi-Modal Analysis and Time-Frequency Analysis (WPT+FC) | | 10-Fold-Cross Validation | 99.75% |
| Yoon A. et al. (2024) | Spectral Analysis and Functional Connectivity (FFT+FC) | | Random Forest + Z-Score | 90%-92.52% |
| Our Proposed Solution | Spectral Analysis (DChT) | LDA | CNN | 98.87% |

VI. CONCLUSION

We propose a new health diagnosis system that detects the occurrence of stroke diseases with the attribute information of statistical metric extracted from Discrete Tchebychev Transform (DChT) of raw EEG signal collected from stroke patients and healthy controls. Spectral analysis, Extraction of discriminative parameters, Dimensionality reduction and Classification were performed through Tchebychev polynomials transform, Linear discriminant analysis and a one-dimensional Convolutional Neural Network (1D-CNN). Above all, the health diagnosis system in this study can detect and predict the occurrence of stroke, thus providing accurate prediction results in a system that can be implemented at low cost.

As a result, the system has great advantages as it can provide in-depth analysis information useful for the behavioral characteristics before and during stroke attack. In order to ameliorate this work it is necessary for further studies to integrate similar analysis with ECG and EMG signals.

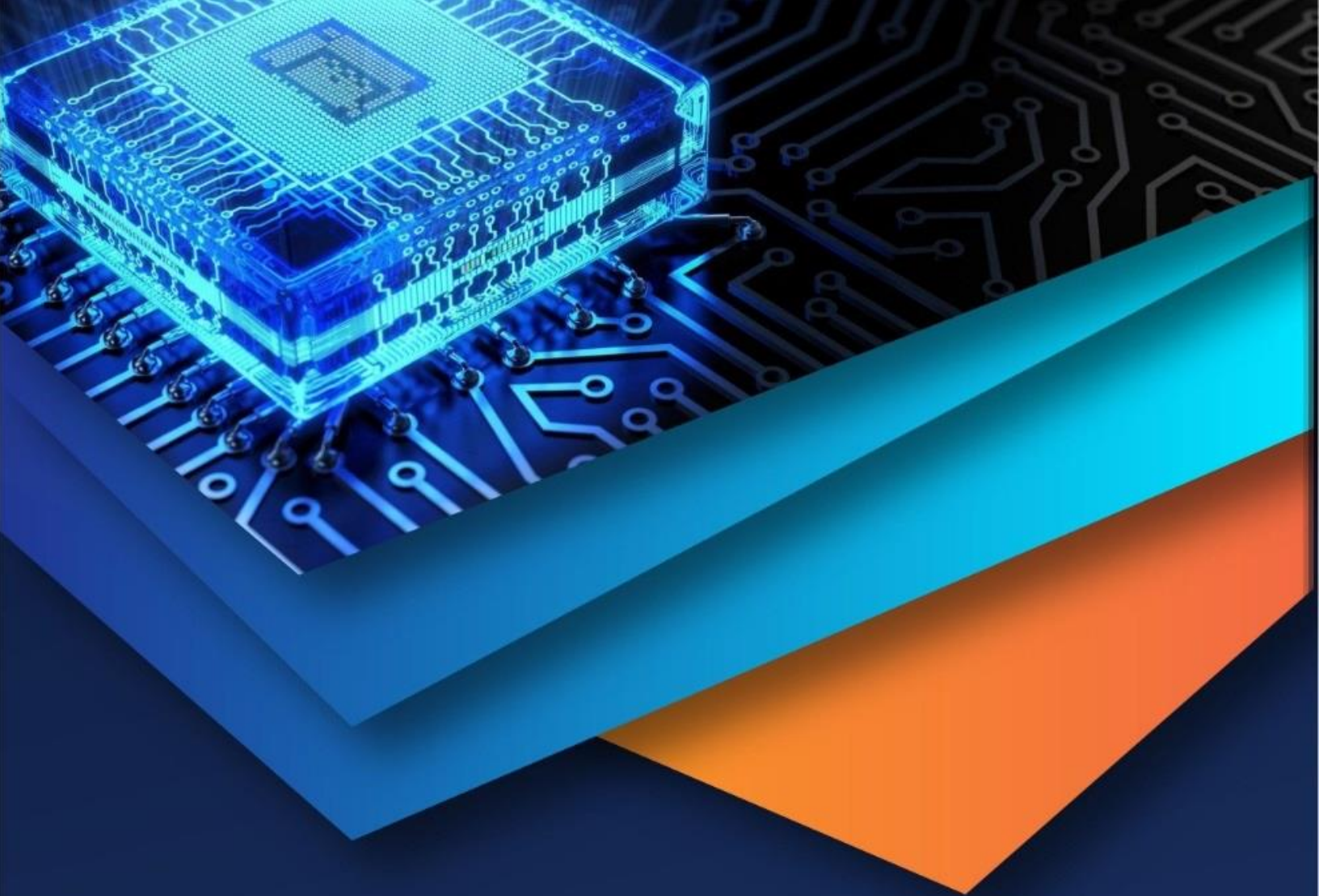
VII. CONFLICT OF INTEREST

The authors declare that they have no conflict of interest. The authors declare that there is no competing financial interest or personal relationship that could have appeared to influence the work reported in this paper.

REFERENCES

- [1] B. Foreman, J. Claassen, K. Abou Khaled, J. Jirsch, D. M. Alschuler, J. Wittman, R. G. Emerson, and L. J. Hirsch, "Generalized periodic discharges in the critically ill: a case-control study of 200 patients," *Neurology*, vol. 79, no. 19, pp. 1951–1960, 2012.
- [2] B. Ovbiagele, L. B. Goldstein, R. T. Higashida, V. J. Howard, S. C. Johnston, O. A. Khavjou, D. T. Lackland,
- [3] J. H. Lichtman, S. Mohl, R. L. Sacco *et al.*, "Forecasting the future of stroke in the united states: a policy statement from the American Heart Association and American Stroke Association," *Stroke*, vol. 44, no. 8, pp. 2361–2375, 2013.
- [4] C. O. Johnson, M. Nguyen, G. A. Roth, E. Nichols, T. Alam, D. Abate, F. Abd-Allah, A. Abdelalim, H. N. Abraha, N. M. Abu-Rmeileh *et al.*, "Global, regional, and national burden of stroke, 1990–2016: a systematic analysis for the global burden of disease study 2016," *The Lancet Neurology*, vol. 18, no. 5, pp. 439–458, 2019.
- [5] H. Jokinen, S. Melkas, R. Ylikoski, T. Pohjasvaara, M. Kaste, T. Erkinjuntti, and M. Hietanen, "Post-stroke cognitive impairment is common even after successful clinical recovery," *European journal of neurology*, vol. 22, no. 9, pp. 1288–1294, 2015.
- [6] J. Petrovic, V. Milosevic, M. Zivkovic, D. Stojanov, O. Milojkovic, A. Kalauzi, and J. Saponjic, "Slower eeg alpha generation, synchronization and "flow"—possible biomarkers of cognitive impairment and neuropathology of minor stroke," *PeerJ*, vol. 5, p. e3839, 2017.
- [7] T. Aydin, F. N. Kesiktas, M. M. Oren, T. Erdogan, Y. C. Ahisha, T. Kizilkurt, M. Corum, I. Karacan, S. Öztürk, and G. Bahat, "Sarcopenia in patients following stroke: An overlooked problem," *International Journal of Rehabilitation Research*, vol. 44, no. 3, pp. 269–275, 2021.
- [8] N. Wright, K. Newell, K. H. Chan, S. Gilbert, A. Hacker, Y. Lu, Y. Guo, P. Pei, C. Yu, J. Lv *et al.*, "Long-term ambient air pollution exposure and cardio-respiratory disease in china: findings from a prospective cohort study," *Environmental Health*, vol. 22, no. 1, p. 30, 2023.
- [9] S. Benghanem, N. Kubis, E. Gayat, A. Loiodice, E. Pruvost-Robieux, T. Sharshar, A. Foucrier, S. Figueiredo, V. Bouilleret, E. De Montmollin *et al.*, "Prognostic value of early eeg abnormalities in severe stroke patients requiring mechanical ventilation: a pre-planned analysis of the spice prospective multicenter study," *Critical Care*, vol. 28, no. 1, p. 173, 2024.
- [10] Y. Jin, J. Li, Z. Fan, X. Hua, T. Wang, S. Du, X. Xi, and L. Li, "Recognition of regions of stroke injury using multi-modal frequency features of electroencephalogram," *Frontiers in Neuroscience*, vol. 18, p. 1404816, 2024.
- [11] Y. Hu, Y. Wang, R. Zhang, Y. Hu, M. Fang, Z. Li, L. Shi, Y. Zhang, Z. Zhang, J. Gao *et al.*, "Assessing stroke rehabilitation degree based on quantitative eeg index and nonlinear parameters," *Cognitive Neurodynamics*, vol. 17, no. 3, pp. 661–669, 2023.
- [12] L. Sutcliffe, H. Lumley, L. Shaw, R. Francis, and C. I. Price, "Surface electroencephalography (eeg) during the acute phase of stroke to assist with diagnosis and prediction of prognosis: a scoping review," *BMC Emergency Medicine*, vol. 22, no. 1, p. 29, 2022.

- [13] M.S.N. AgLamat, M.S.H. AbdRahman, W. A. WanZaidi, W.N.N.W. Yahya, C.S. Khoo, R. Hod, and H.J. Tan, "Qualitative electroencephalogram and its predictors in the diagnosis of stroke," *Frontiers in Neurology*, vol. 14, p. 1118903, 2023.
- [14] Z. Keser, S. C. Buchl, N. A. Seven, M. Markota, H. M. Clark, D. T. Jones, G. Lanzino, R. D. Brown Jr, G. A. Worrell, and B. N. Lundstrom, "Electroencephalogram (eeg) with or without transcranial magnetic stimulation (tms) as biomarkers for post-stroke recovery: a narrative review," *Frontiers in neurology*, vol. 13, p. 827866, 2022.
- [15] A. L. Schneider and K. G. Jordan, "Regional attenuation without delta (rawod): a distinctive eeg pattern that can aid in the diagnosis and management of severe acute ischemic stroke," *American journal of electrophysiology*, vol. 45, no. 2, pp. 102–117, 2005.
- [16] B. Foreman and J. Claassen, "Quantitative eeg for the detection of brain ischemia," *Critical care*, vol. 16, pp. 1–9, 2012.
- [17] S. Finnigan and M. J. Van Putten, "Eeg in ischemic stroke: quantitative eeg can uniquely inform (sub-)acute prognosis and clinical management," *Clinical neurophysiology*, vol. 124, no. 1, pp. 10–19, 2013.
- [18] S. Finnigan, A. Wong, and S. Read, "Defining abnormal slow eeg activity in acute ischemic stroke: Delta/alpha ratio as an optimal qeeg index," *Clinical Neurophysiology*, vol. 127, no. 2, pp. 1452–1459, 2016.
- [19] R. A. Chowdhury, G. Pellegrino, Ü. Aydin, J.-M. Lina, F. Dubeau, E. Kobayashi, and C. Grova, "Reproducibility of eeg-meg fusion source analysis of interictal spikes: Relevance in presurgical evaluation of epilepsy," *Human brain mapping*, vol. 39, no. 2, pp. 880–901, 2018.
- [20] F. De Vico Fallani, L. Astolfi, F. Cincotti, D. Mattia, D. la Rocca, E. Maksuti, S. Salinari, F. Babiloni, B. Vegso, G. Kozmann *et al.*, "Evaluation of the brain network organization from eeg signals: a preliminary evidence in stroke patient," *The Anatomical Record: Advances in Integrative Anatomy and Evolutionary Biology: Advances in Integrative Anatomy and Evolutionary Biology*, vol. 292, no. 12, pp. 2023–2031, 2009.
- [21] Z. Ip, G. Rabiller, J. He, Z. Yao, Y. Akamatsu, Y. Nishi-jima, J. Liu, and A. Yazdan-Shahmorad, "Cortical stroke affects activity and stability of theta/delta states in remote hippocampal regions," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2019, pp. 5225–5228.
- [22] M. W. M. G. A. J. SP Finnigan, SE Rose, "Correlation of quantitative eeg in acute ischemic stroke with 30-day nihss score: comparison with diffusion and perfusion mri," p. 5, 2019.
- [23] J. J. C. H. C. P. H. L. YA Choi, S Park, "Machine-learning-based elderly stroke monitoring system using electroencephalography vital signals," p. 5, 2023.
- [24] X. L. H. Liu, "Eeg dataset of stroke patients," 2023.
- [25] G. M. D. H. T. B. N. W. J. Schalk, "Eeg motor movement," 2009.
- [26] D. d' Helsing Laennec, "Principes éthiques applicables aux recherches médicales sur des sujets humains," 2002.
- [27] D. Tchiotso, "Modélisation polynomiale des signaux eeg: application à la compression," 2007.
- [28] B. H. et A. LOUAIL, "Applications des polynômes orthogonaux pour la résolution des équations intégrales," 2021.
- [29] V. Rossetto, "Outils numériques," 2021.
- [30] A. Subasi and M. I. Gursoy, "Eeg signal classification using pca, ica, lda and support vector machines," *Expert systems with applications*, vol. 37, no. 12, pp. 8659–8666, 2010.
- [31] K. Fukunaga, *Introduction to statistical pattern recognition*. Elsevier, 2013.
- [32] L. C. D. Nkengfack, D. Tchiotso, R. Atangana, V. Louis-Door, and D. Wolf, "Eeg signals analysis for epileptic seizures detection using polynomial transforms, linear discriminant analysis and support vector machines," *Biomedical Signal Processing and Control*, vol. 62, p. 102141, 2020.
- [33] R. C. González, "Towards an automated portable electroencephalography-based system for alzheimer's disease diagnosis," Ph.D. dissertation, Institut National de la Recherche Scientifique (Canada), 2018.
- [34] R. Chetana, A. S. Rao, and K. Mahantesh, "Application of conv-1d and bi-lstm to classify and detect epilepsy eeg data," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 6, 2023.



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)