Available online at www.HighTechJournal.org

HighTech and Innovation Journal

ISSN: 2723-9535

Vol. 6, No. 3, September, 2025

Review Article

ML and DL Models for Stroke Prediction from Bio-Signals: A Systematic Review and Bibliometric Analysis

May Issa Aldossary [1], Fatemah H. Alghamedy [2*], Dina A. Alabbad [3], Renad A. Alnuaim [3], Maimonah S. Altaweel [3], Reem A. H. Alshami [3], Haya A. Alzahim [3], Shahad F. Alotaibi [3], Sumayh S. Aljameel [4], Areej Almalki [2], Sunday O. Olatunji [5]

[1] *Computer Information Systems Department, College of Computer Science and Information Technology, Imam Abdulrahman Bin Faisal University, Dammam, 34212, Saudi Arabia.*

[2] *Computer Science Department, Applied College, Imam Abdulrahman Bin Faisal University, Dammam, 34212, Saudi Arabia.*

[3] *Computer Engineering Department, College of Computer Science and Information Technology, Imam Abdulrahman Bin Faisal University, Dammam, 34212, Saudi Arabia.*

[4] *Aramco Saudi Accelerated Innovation Lab (aramcoSAIL), Saudi Aramco, Dhahran 31311, Saudi Arabia.*

[5] *Faculty of Computing, Adekunle Ajasin University, Akungba Akoko, Ondo State, Nigeria.*

Received 12 June 2025; Revised 11 August 2025; Accepted 22 August 2025; Published 01 September 2025

Abstract

Strokes continue to be a primary reason for disability and death around the globe. Annually, over 12.2 million new strokes occur, which necessitates the development of early detection and intervention tools to reduce the potential harm. This systematic review and bibliometric analysis aim to review and visualize recent advances in predicting stroke or post-stroke effects using bio-signals, either with machine learning (ML) or deep learning (DL). The included studies were published between 2016 and 2024. A comprehensive search of IEEE, PubMed, MDPI, and ScienceDirect databases was performed using keywords related to stroke prediction, machine learning, deep learning, and bio-signals. From an initial pool of 152 studies, 15 studies met the inclusion criteria through the screening process. South Korea contributed the most to publishing studies on stroke prediction using bio-signals. The results show that Electroencephalography (EEG) is the most used bio-signal in the reviewed studies. The sample size ranged from 3 to 4068. The top ten cited journals in the selected literature are high-ranked journals, which indicates the scientific validity of the concept and its potential for dissemination.

Keywords: Applied AI; Bio-Signals; Deep Learning; EEG; Machine Learning; Stroke Detection; Post-Stroke Effect; Bibliometric Analysis.

1. Introduction

Strokes continue to be a primary reason for disability and death around the globe. Every year, over 12.2 million new strokes occur. In addition, above the age of 25, one in four individuals will experience a stroke in their lifetime, which necessitates the development of early detection and intervention tools [1]. Brain stroke, as a cerebrovascular accident (CVA), is a medical disorder that happens when the blood supply to the brain is suddenly disrupted. This disturbance can be produced by a blood vessel obstruction (ischemic stroke) or a blood vessel rupture (hemorrhagic stroke). In either

* Corresponding author: falghamedy@iau.edu.sa

http://dx.doi.org/10.28991/HIJ-2025-06-03-019

case, strokes can cause brain cells to lose oxygen and nutrients, resulting in brain damage and a variety of neurological symptoms such as paralysis, speech difficulty, and cognitive impairment [2]. Immediate medical intervention is required to reduce the potential harm from a stroke.

Traditionally, strokes are diagnosed by brain scans and physical examinations, such as Magnetic Resonance Imaging (MRI) and Computed tomography (CT) scans [3]. Despite these effective techniques, they are time-consuming and cannot be used continuously since they may increase cancer risk [4]. In contrast, in the last few years, the interest in exploring the use of bio-signals and machine learning (ML) as potential predictors of stroke occurrence has increased. Bio-signals, referred to as physiological signals, indicate the measurable electrical or chemical activities produced by the human body. For example, Electroencephalography (EEG) measures electrical activity in the brain, detecting neural patterns and diagnosing disorders [5], Electrocardiography (ECG) measures the electrical activity of the heart, aiding in the diagnosis of cardiac disease and arrhythmias [6], Electromyography (EMG) examines muscle electrical activity and assists in identifying neuromuscular disorders [7], and Photoplethysmography (PPG) measures variations in blood volume using fingertip sensors to monitor heart rate and detect blood flow irregularities [8]. These non-invasive methods are essential for diagnosing and monitoring a variety of medical disorders. In the context of stroke detection, bio-signals are used to identify specific patterns or changes that may indicate an increased risk of stroke [9].

Despite the growing interest in applying ML/DL techniques to bio signal-based stroke detection, existing published studies remain limited in different aspects. Most of the studies were centred on a specific region, with a small clinical dataset size, raising the need for a larger, more diverse dataset. Furthermore, none of the studies covered stroke detection in the hospital workflow. Additionally, none of the reviewed studies incorporated advanced model performance or enhancement techniques such as ensemble learning, explainable AI, wavelet transforms, or Fourier transforms. These gaps emphasize the need for systematic review and bibliometric analysis that focus on the studies that produced stroke detection or post-stroke effects detection based on ML and DL models using bio-signal data.

This systematic review and bibliometric analysis aims to identify gaps in the literature related to stroke detection or post-stroke effects detection using bio-signal data with ML and DL models. In addition, it provides a foundation for developing detection algorithms in the stroke field. To our knowledge, this is the first systematic review and bibliometric analysis that studies proposed methods applying ML or DL in stroke detection or post-stroke effects detection using bio-signal data.

Section 2 reveals the previous surveys and reviews utilizing ML and DL models for predicting strokes and post-stroke effects via bio-signals. Section 3 illustrates the systematic review methodology, including research questions, search strategy, inclusion and exclusion criteria, study selection, reporting quality assessment, and data extraction. In Section 4, the results of the systematic review methodology are illustrated, including the PRISMA flowchart, the AI models that have been used in the literature, and the sample size according to the algorithms used. Section 5 visualizes the bibliometric analysis of the selected literature based on author keyword co-occurrence and co-citation. Section 6 discusses the findings of the systematic review and bibliometric analysis. Section 7 discusses the limitations of the reviewed research. Finally, Section 8 concludes our systematic review and provides suggestions for future researchers.

2. Motivation and Related Surveys

Bio-signals are used for many purposes in medical fields, including monitoring conditions, detecting illness, limiting its effects, and accelerating recovery. Our motivation is to enrich the medical and AI fields by investigating the existing studies that use bio-signals to detect early-stage strokes or post-stroke effects by utilizing ML or DL. In addition, we hope this systematic review and bibliometric analysis will motivate researchers to leverage bio-signal data for stroke detection.

This section presents the previous surveys and reviews utilizing ML and DL models for predicting strokes and post-stroke effects via bio-signals. The following syntax was used to search for existing surveys and reviews:

(("Machine Learning" OR "Deep Learning" OR "Classification" OR "Supervised Learning" OR "Neural Networks") AND ("Stroke prediction" OR "Predicting Stroke") AND ("bio-signals" OR "ECG" OR "EMG" OR "PPG" OR "EEG") AND ("Review" OR "Survey")).

Book chapters were excluded from the search due to their specific focus, which was not aligned with the research objectives. Scopus searches in titles, keywords, and abstracts, while the MDPI search focuses on titles and keywords. PubMed and Google Scholar are used for searches that concentrate on titles and abstracts. IEEE Xplore uses general settings. The search results are shown in Table 1.

As shown in Table 1, the search results totaled five articles. Google Scholar and IEEE Xplore have no articles that match our search query. We scrutinize each research to ensure that it meets our search keywords. None of the five studies conducted a systematic review and bibliometric analysis on utilizing ML and DL Models for predicting strokes and post-stroke effects via bio-signals, including ECG, EMG, PPG, and EEG. We aim for this paper to contribute significantly to applying AI in the medical field to predict stroke early by utilizing bio-signals.

Table 1. Related surveys and reviews

Ref.	Type	Database	Year	Stroke	ML	DL	ECG	EMG	PPG	EEG	Bibliometric	Note
[10]	Conference Paper	Scopus	2024	✓	✓	✗	✓	✗	✓	✗	✗	-
[11]	Review	Scopus PubMed	2020	✗	✓	✗	✓	✗	✗	✗	✗	Atrial Fibrillation (AF)
[12]	Review	MDPI	2022	✗	✗	✗	✗	✗	✗	✓	✓	Rehabilitation
[13]	Review	MDPI	2022	✓	✓	✗	✗	✗	✗	✗	✗	-
[14]	Review	MDPI	2021	✗	✗	✓	✓	✗	✗	✗	✗	Atrial Fibrillation (AF)
Our Study	Systematic Review	-	2024	✓	✓	✓	✓	✓	✓	✓	✓	-

3. Systematic Reviews Methods

This review uses Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA). To specify the research expectation, it is crucial to clearly define research questions, search strategy, and selection criteria.

3.1. Research Questions (RQs)

To define the key components of the research questions, this review utilized the PICO framework.

- RQ1: What bio-signals can be effectively utilized for the early detection of strokes?
- RQ2: What are the most ML and DL classifiers used with bio-signal data to detect strokes?
- RQ3: What are the most promising ML and DL models regarding result accuracy?
- RQ4: Which countries contribute the most to enriching research in this field?

3.2. Search Strategy

The reviewed studies were collected from IEEE, PubMed, MDPI, ScienceDirect, and Nature. The following keywords were used: ("Machine Learning," OR "Deep Learning" OR "Classification" OR "Supervised Learning" OR "Neural Networks") AND ("Stroke prediction" OR "Predicting Stroke") AND ("bio-signals" OR "ECG" OR "EMG" OR "PPG" OR "EEG"). After conducting an extensive search across multiple databases, a comprehensive protocol based on specific inclusion criteria was established to identify publications that match the requirements of the review. The qualified studies that met the following inclusion criteria were considered: (a) publication in English; (b) publication in high-ranking journals or conferences and excluding reviews; (c) publication between the years 2016 and 2024; (d) contains experiment and result sections; (e) Focused on predicting stroke or poststroke effects detection using bio-signal data; and (f) Availability of the complete study rather than just abstracts or notes. This systematic review aims to review recent advances in supervised ML and DL models for stroke detection or post-stroke effects detection.

3.3. Inclusion and Exclusion Criteria

An article is considered in this review when it meets the inclusion criteria as follows:

- Written English language;
- Published in high-ranking journals or conferences;
- Published between the years 2016 and 2024;
- Focused on predicting stroke or post-stroke effects detection;
- Using bio-signal data;
- Availability of the complete study rather than just abstracts or notes.

On the other hand, an article is not considered when it fits in one of the exclusion criteria as follows:

- Utilizing clinical evaluations, imaging (CT, MRI, MRA, etc.), blood tests, or any non-signal-based data;
- The source (journal or conference) is not peer-reviewed;
- Review, survey, chapter book, thesis, or dissertation articles;
- Missing experiment and result;
- Missing popular ML/DL metric measurements, e.g., accuracy;
- Published prior to 2016;
- Written in a language other than English;
- Medical-based methods to predict stroke or post-stroke effects detection.

Due to the sensitivity of the topic in the medical field, our strict criteria may unintentionally filter out innovative or non-traditional approaches that are published within the scope that are not covered by our inclusion criteria. However, including such excluded research could extend to other types of research, such as evidence-based, case studies, and early-stage innovations. In addition, including research from a non-peer-reviewed high-ranking journal may reduce bias in presenting positive results, thereby capturing a broader picture of the techniques. However, it may introduce more challenges in the assessment.

3.4. Study Selection

To evaluate the appropriateness of the studies obtained from the searches by examining the titles and abstracts of all articles. In case of any disagreement, extensive discussion was employed. For all studies considered relevant, their full text was thoroughly reviewed. The studies were considered eligible if they met the inclusion criteria.

3.5. Reporting Quality Assessment

A customized checklist items, as shown in Table A1, was created to evaluate the risk of bias in the selected studies that developed ML and DL Prediction Models for bio-signal data. Studies are assessed through their title, abstract, introduction, methods and results, and other information.

3.6. Data Extraction

A detailed form was created to collect data in an organized manner, which helps us extract the study characteristics (authors, publication year, study objective), methods (techniques and models), data (source of data, type of data, sample size), participants (participants' condition), and results (reported performance measure, code availability).

3.7. Citation of Tables and Figures

All tables and figures included in this systematic review paper are clearly cited and referenced within the main text. Additionally, supplementary materials are included in Appendix I, where Table A1 summarizes the customized checklist for the research sections criteria, Table A2 represents the extracted items from the reviewed papers, Table A3 illustrates the quality assessment data for each study, and lastly Table A4 includes the study author, objective, source of data, type of data, sample size, techniques, outcome, region under study, and published year.

4. PRISMA Results

We identified a total of 153 studies, from which 11 studies were from IEEE, 12 studies from MDPI, nine studies from PubMed, 115 studies from ScienceDirect, and six studies from Nature. After the removal of duplicates, as well as abstract and title screening, 57 studies were considered potentially relevant, 31 of which were not accessible/not available. After screening the full articles of accessible articles, 15 studies were identified for information extraction. The process is illustrated in Figure 1. All studies were published as peer-reviewed publications in reliable and well-known journals and conferences. All included studies were published after 2015, with more than half (9 studies) published after 2020, from which two studies were published in 2021 [15, 16], five studies were published in 2022 [17-21], two studies were published in 2023 [22, 23], and one study was published in 2024 [24]. In terms of regions under study, South Korea (7) [15-17, 20, 24, 25-26] and USA (2) [21, 22] make up more than half of the sample. Pakistan [27], Canada [28], and India [19] had one study each. Whereas the rest of the studies (3) [18, 23, 29] did not mention the region under study.

All included studies focused on stroke detection, except for two studies [21, 29] which focused on post-stroke effects detection. Out of the fifteen studies, more than half of the studies (8) [15-18, 21, 27-29] used EEG signal data, whereas three studies [19, 22, 25] used ECG signal data, three studies [26, 23, 24] used EMG signal data, and one study [20] used a combination of ECG and PPG signal data. For sources of data, more than half of the studies (11) [15-17, 19-21, 24-28] collected data from hospitals, while some studies (3) [18, 22, 23] utilized data from online databases, one study [29] did not mention clear information about how the utilized dataset was collected. More than half of the studies (9) [15-17, 19-21, 24-26] utilized datasets that include more than 100 samples, while some of the studies (5) [22-23, 27-29] used datasets that include less than 100 samples, and one study [18] did not mention clearly the sample size.

For predictive models' development, the most used ML methods were RF (6) [16, 20, 24-26, 28], SVM (3) [24, 25, 27], KNN (2) [24-25], LR (2) [24, 25], DT (2) [20, 24]. Whereas Adaboost [17], XGBoost [17], LightGBM [17], NB [25], LDA [29], and RDA+KDE classifier [21] were used by one study each. Meanwhile, the most used DL methods were CNN (3) [18-19, 22] and LSTM (2) [19, 26] and ensemble of CNN and LSTM [15, 20]. Whereas a single study CNN and bidirectional LSTM in one model [15], another study used RNN [19], and one study used Stacked CNN with LSTM and GMDH [23]. Upon model development and evaluation, the highest accuracy among the studies was obtained by [24], which developed a RF model that scored a remarkable accuracy of 100%. The number of studies published based on the algorithms used each year is depicted in Figure 2.

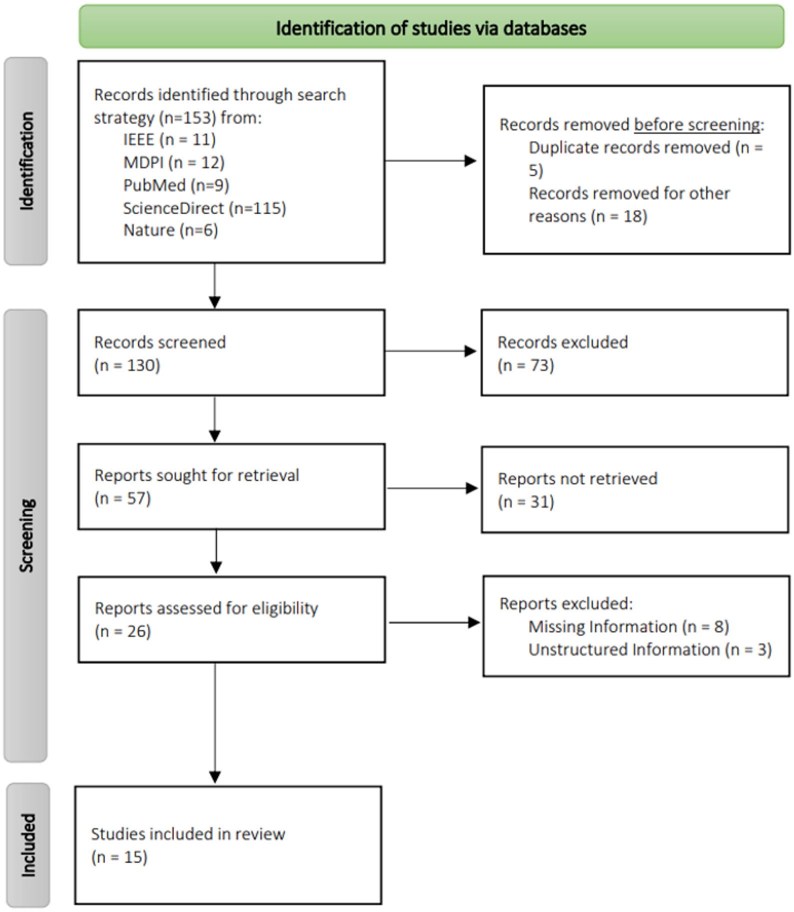

Figure 1. PRISMA Flowchart

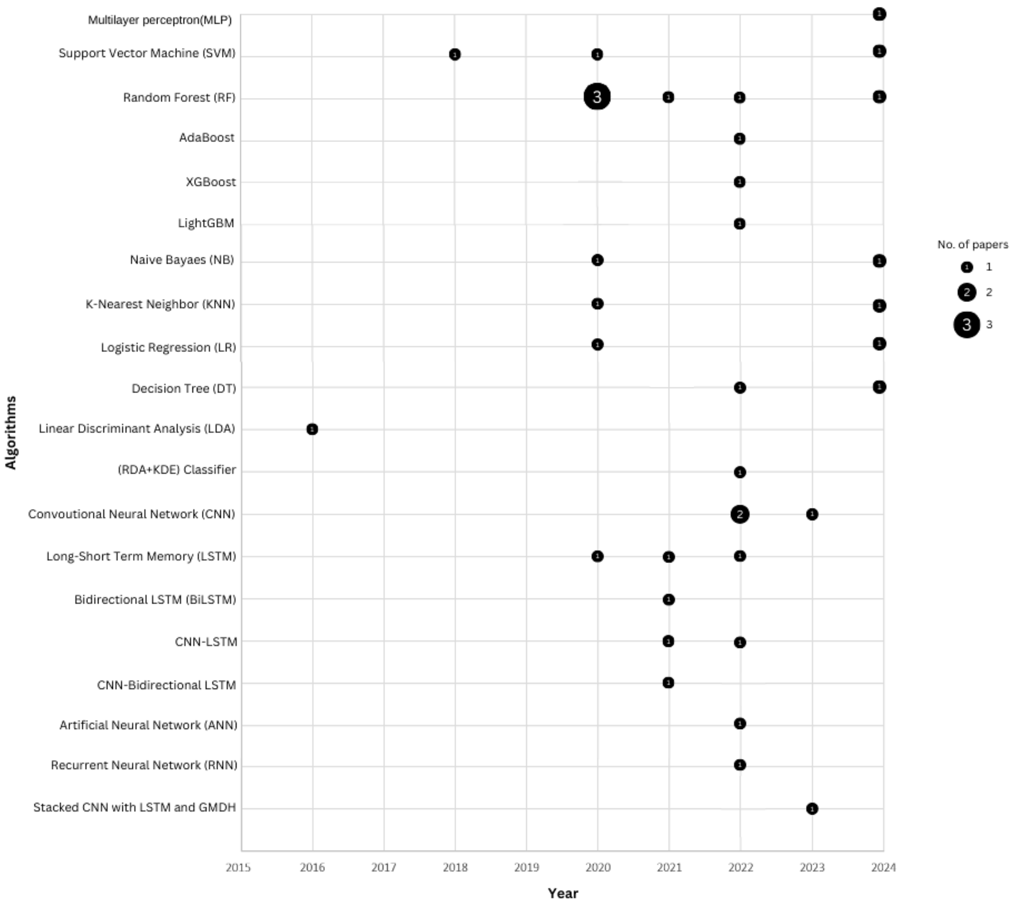

Figure 2. Number of studies published according to the algorithms used each year

We considered eleven customized checklist items for each study, as depicted in Figure 3. Two items (1) and (2) are about the title and abstract sections. One item (3) is about the introduction section. Seven items (4a)(4b)(5)(6)(7)(8a)(8b) are about the methods and results sections, and one item (9) is about the study code. Seven items (1)(2)(3)(4b)(7)(8a)(8b) are reported by all studies. Three items (4a)(5)(6) are reported by the majority of studies. However, item (9) is reported by one study [28] only as depicted in Figure 4. The relationship between the algorithms used and the sample size is depicted in Figure 5. ML algorithms were used with smaller sizes when compared to DL algorithms. For example, the mean and median of the RF algorithm are 312.4 and 273, respectively. On the other hand, the mean and median of the CNN algorithm are 2069.5. All sample sizes were less than 600 except for study [19], which has a sample size equal to 4068 and used CNN, RNN, and LSTM algorithms. In addition, Table 2 presents a summary of the studies.

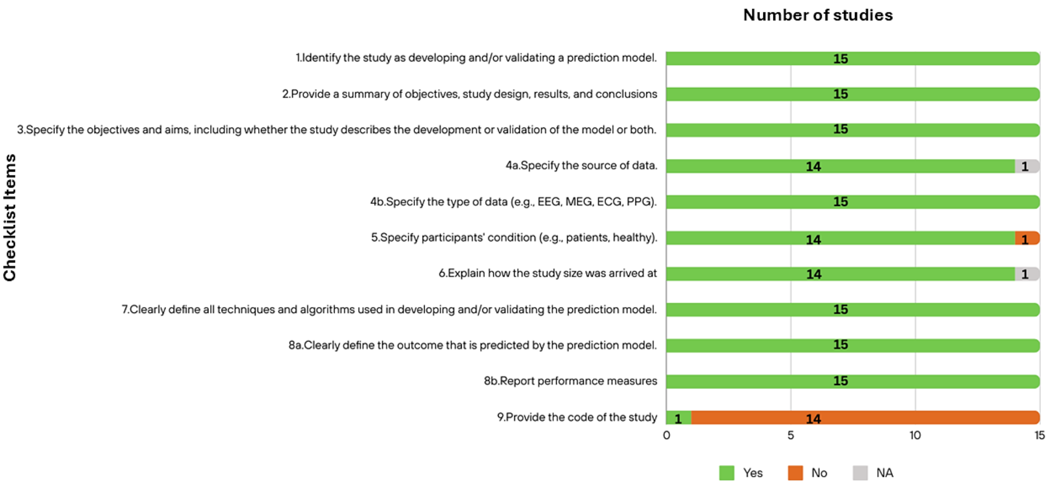

Figure 3. Number of studies reported for each checklist item

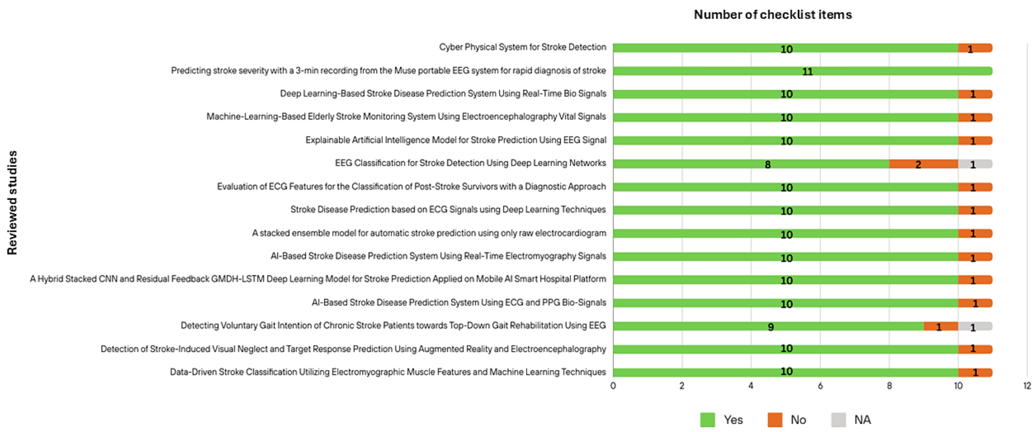

Figure 4. Number of checklist items reported in each study

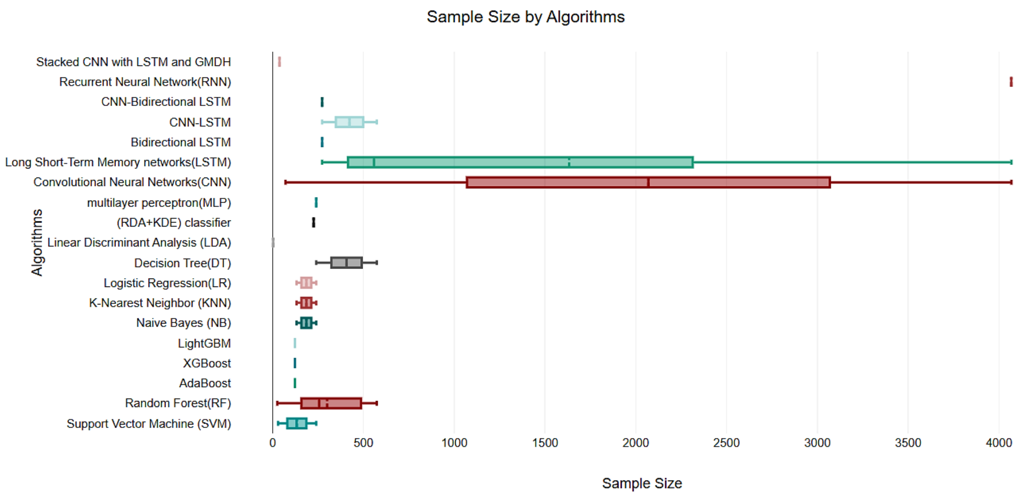

Figure 5. Boxplots showing the distribution of sample size according to algorithms used

Table 1. A summary of the reviewed studies

Ref.	Objective	Data source	Data type	Sample size	Techniques	Outcome
[27]	Reveal the occurrence of stroke in patients who have previously survived a stroke or who have a high risk of experiencing stroke.	From Shaheed Mohtarma Benazir Bhutto Medical University, Larkana, Pakistan.	EEG Signals.	30	ML: SVM.	SVM: Precision: 100% Recall: 99.16%
[28]	Utilizing an inexpensive portable EEG device as a method for prehospital stroke and examining whether EEG data can be used to detect changes in stroke intensity.	University of Alberta Hospital.	EEG Signals.	25	ML: RF - Muse by InteraXon Inc., a device to record EEG signals.	RF: Accuracy: 76% Sensitivity: 63% Specificity: 86%
[15]	Presented a new way for applying deep learning models to raw EEG data without relying on the frequency features of EEG.	Emergency Medical Center of Chungnam National University Hospital.	EEG Signals.	273	DL: LSTM, Bidirectional LSTM, CNN-LSTM, CNN-Bidirectional LSTM.	Raw values showed the best accuracy. LSTM: Accuracy: 70.1% Bidirectional LSTM: Accuracy: 91.8% CNN-LSTM: Accuracy: 93.7% CNN-Bidirectional LSTM: Accuracy: 94.0%
[16]	Developing a health monitoring system that can anticipate the symptoms of stroke diseases in old people in real-time while they walk on a regular basis.	Emergency Medical Center and Rehabilitation Department of Chungnam National University Hospital.	EEG Signals.	273	ML: RF.	RF: Accuracy 92.51%
[17]	Classifying stroke and healthy control groups for stroke prediction in active situations.	Korea Research Institute of Standards and Science.	EEG Signals.	123	ML: AdaBoost, XGBoost, and LightGBM.	AdaBoost: Accuracy: 80%
[18]	Develop models to categorize EEG signals as strokes or non-strokes.	Normal and abnormal EEG activity from PhysioNet.	EEG Signals.	Unknown	Deep neural network architecture, RESNET-50, and VGG-16.	RESNET-50: Accuracy: 90% Sensitivity: 100% VGG-16: Accuracy: 90% Specificity: 100% Precision: 100%
[25]	Develop a classification model using machine learning and ECG signals for diagnosing stroke disease.	Chungnam National Hospital, Daejeon, South Korea.	ECG Signals.	132	ML: SVM, RF, Naïve Bayes, KNN, and LR.	KNN: Accuracy: 96.6% RF: Accuracy: 94.4% SVM: Accuracy: 85.4% Naïve Bayes: Accuracy: 72.7% LR: Accuracy: 66.9%
[19]	Proposing a medical framework to detect abnormalities in the ECG associated with stroke disease.	Indian hospitals.	ECG Signals.	4068	DL: LSTM, CNN, and RNN.	LSTM: Accuracy: 93.78% CNN: Accuracy: 89.25% RNN: Accuracy: 86.19%
[22]	Developing a classification model based on ECG signals for stroke diagnosis.	The cerebral vasoregulation Dataset.	ECG Signals.	71	Stacking ensemble model of CNN models.	CNN: Accuracy: 99.7% F1: 99.69% Recall: 99.71% Precision: 99.67%
[26]	Developing a stroke prediction system with the use of real-time EMG signals.	Emergency Medical Center and the Department of Rehabilitation Medicine at Chungnam National University Hospital	EMG Signals.	558	ML: RF. DL: LSTM.	RF: Accuracy: 90.38% LSTM: Accuracy: 98.96%
[23]	Proposing a telemedicine system that predicts heart, and brain stroke.	EMG Lower Limb Dataset mHealth Dataset EMG Physical Action Dataset.	EMG Signals.	38	DL: Stacked CNN + LSTM + GMDH. Explainable AI (XAI).	Stacked CNN + LSTM + GMDH: Accuracy: 99%
[20]	Develop multi-models based on ML, ECG, and PPG signals.	Emergency medical center and department of rehabilitation medicine at Chungnam National University Hospital, Republic of Korea.	ECG and PPG Signals.	574	DL: an ensemble structure that combines CNN and LSTM. ML: Decision Tree, RF.	Decision Tree: Accuracy: 91.56% RF: Accuracy: 97.51% CNN-LSTM: Accuracy: 99.15%
[29]	Decoding stroke patients' gait intentions using EEG signals.	Unknown	EEG Signals.	3	Linear Discriminant Analysis (LDA).	LDA: Accuracy: 73.2% Delay is 0.13 s
[21]	Proposing system combines EEG data and Augmented Reality (AR) to identify the presence of Visual-Spatial Neglect (SN) in stroke patients.	University of Pittsburgh Medical Center Inpatient Rehabilitation.	EEG Signals.	226	(RDA+KDE) Classifier.	(RDA+KDE): Average train AUC: 0.788 Average test AUC: 0.760
[24]	Examine the impact of the statistical features of muscle activity of the major leg muscles during gait as predictive factors across various models to differentiate between stroke patients and healthy individuals.	Multiple medical institutions across South Korea.	EMG Signals.	240	DT, RF, LR, MLP, SVC, K-NN, NB.	RF: Accuracy: 100% LR: Accuracy: 96% DT: Accuracy: 94% MLP: Accuracy: 99% SVM: Accuracy: 94% NB: Accuracy: 77% KNN: Accuracy 85%

Table 3 illustrates a customized structured bias matrix employed across five key dimensions: (D1) dataset clarity, (D2) model description, (D3) evaluation metrics, (D4) validation approach, and (D5) reproducibility. Table 4. comprehensively explain each dimension definition and evaluation guidance. The results of the risk bias matrix demonstrated low bias across (D1-D2), indicating that most studies provided transparent information related to utilized dataset, models, evaluation metrics and validation procedure. However, (D5) reproducibility showed high bias among studies, due to limited access to code or data sharing, which prevented replication. The use of risk bias matrix ensured comparability of results across diverse methodologies and robustness of performance metric extraction from heterogeneous sources.

Table 2. Customized risk of bias matrix

Ref.	D1: Clarity of dataset	D2: Model description	D3: Evaluation metric	D4: Validation (CV)	D5: Reproducibility
Cyber Physical System for Stroke Detection [27]	Low bias	Low bias	Low bias	Low bias	High bias
Predicting stroke severity with a 3-min recording from the Muse portable EEG system for rapid diagnosis of stroke [28]	Low bias	Low bias	Low bias	Low bias	Low bias
Deep Learning-Based Stroke Disease Prediction System Using Real-Time Bio Signals [15]	Low bias	Low bias	Low bias	Low bias	High bias
Machine-Learning-Based Elderly Stroke Monitoring System Using Electroencephalography Vital Signals [16]	Low bias	Low bias	Low bias	Low bias	High bias
Explainable Artificial Intelligence Model for Stroke Prediction Using EEG Signal [17]	Low bias	Low bias	Low bias	Low bias	High bias
EEG Classification for Stroke Detection Using Deep Learning Networks [18]	High bias	Low bias	Low bias	Low bias	High bias
Evaluation of ECG Features for the Classification of Post-Stroke Survivors with a Diagnostic Approach [25]	Low bias	Low bias	Low bias	Low bias	High bias
Stroke Disease Prediction based on ECG Signals using Deep Learning Techniques [19]	Low bias	Low bias	Low bias	Low bias	High bias
A Stacked Ensemble Model for Automatic Stroke Prediction Using only Raw Electrocardiogram [22]	Low bias	Low bias	Low bias	Low bias	High bias
AI-Based Stroke Disease Prediction System Using Real-Time Electromyography Signals [26]	Low bias	Low bias	Low bias	Low bias	High bias
A Hybrid Stacked CNN and Residual Feedback GMDH-LSTM Deep Learning Model for Stroke Prediction Applied on Mobile AI Smart Hospital Platform [23]	Low bias	Low bias	Low bias	Low bias	High bias
AI-Based Stroke Disease Prediction System Using ECG and PPG Bio-Signals [20]	Low bias	Low bias	Low bias	Low bias	High bias
Detecting Voluntary Gait Intention of Chronic Stroke Patients Towards Top-Down Gait Rehabilitation Using EEG [29]	High bias	Low bias	Low bias	Low bias	High bias
Detection of Stroke-Induced Visual Neglect and Target Response Prediction Using Augmented Reality and Electroencephalography [21]	Low bias	Low bias	Low bias	Low bias	High bias
Data-Driven Stroke Classification Utilizing Electromyographic Muscle Features and Machine Learning Techniques [24]	Low bias	Low bias	Low bias	Low bias	High bias

Table 3. Criteria definitions for risk of bias assessment

Code	Description	Evaluation guidance
D1	Clarity of dataset	Rate whether dataset source, size, and characteristics are clearly described. *High bias*: unclear/missing dataset info. *Low bias*: fully described.
D2	Model description	Is the algorithm/architecture and key settings described? *Low bias*: algorithm, parameters, and rationale provided. *High bias*: named but lacks necessary detail.
D3	Evaluation metric used	*Low bias*: metrics (e.g., Accuracy, F1, AUC, Sensitivity/Specificity) appropriate and stated. *High bias*: unsuitable or unreported metrics.
D4	Validation (cross-validation)	Is validation proper (e.g., holdout, CV, external test) with no leakage? *High bias*: train/test not separated, or leakage suspected. *Low bias*: appropriate CV/holdout described.
D5	Reproducability	Can results be reproduced (code/data availability, sufficient procedural detail)? *High bias*: no access and insufficient detail. *Low bias*: code/data or full protocol provided.

5. Bibliometric Analysis

In this section, bibliometric analysis is conducted to visualize the literature in Table 2 using VOSviewer. The bibliometric analysis aims to discover trending topics and ML/DL methods for using bio signals in stroke detection. In addition, we aspire to assess the trustworthiness of the knowledge basis in selected studies based on the source ranking.

5.1. Author Keyword Co-Occurrence

Author keyword co-occurrence analysis discloses the knowledge produced by selected studies. Clusters are formed based on the authors' keywords for citing papers that frequently appear together [30]. In Figure 6, the bibliometric analysis presents the co-occurrence analysis based on authors' keywords of studies in Table 2. The bibliometric data was extracted from Scopus. Data was preprocessed to unify the keywords regarding the abbreviation. We use index keywords of these articles [28, 29] because the authors' keywords are missing.

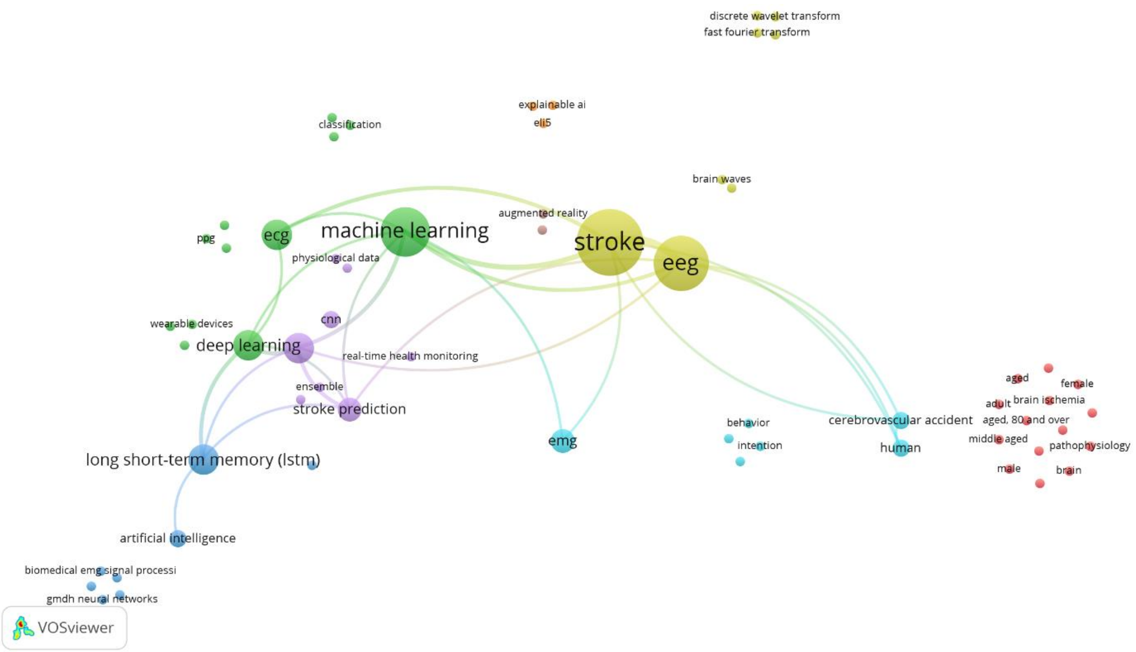

Figure 6. The authors' keywords co-occurrence network visualization

As we see, the center keywords are stroke, EEG, and machine learning. In the visualization map, items with more occurrences of keywords are shown more prominently than items with fewer occurrences. Therefore, stroke and EEG are the most common occurrences, followed by machine learning, deep learning, Long Short-Term Memory (LSTM), ECG, and stroke disease analysis, respectively, as shown in Table 5.

Table 4. The top author keywords occurrence

Keyword	Occurrence
Stroke	10
EEG	8
Machine Learning	7
Deep Learning	4
Long Short-Term Memory	4
ECG	4
Stroke Disease Analysis	4

The links that connect two keywords indicate that these keywords have been appearing in the same publication. The number of publications in which two keywords occur together increases the link strength. We set the minimum number of publications for which any two keywords appear together to two publications. The top keyword pairs that have the most occurrences in two or more publications are stroke with EEG and stroke with machine learning.

This indicates that using EEG in ML and DL models is the most common bio-signal data. In addition, more researchers have been applying machine learning, which leaves promising avenues for researchers to apply deep learning models to benefit from their capacity to handle complex data. In addition, there are eight clusters, each represented by a different color. The clusters were generated by VOSviewer using the association strength method proposed in [31]. The clusters form based on the association strength between the keywords, calculated using the number of co-occurrence links between keywords.

The largest cluster is the red cluster, which contains the brain, clinical trials, and different terms of human age and gender, such as male, female, and elderly. The common feature among the red cluster items is that they represent humans in different circumstances. The second largest is the green, blue, and yellow clusters, which include machine learning, Long Short-Term Memory (LSTM), CNN, prediction, analysis, model, etc. Its theme is AI terms. The authors' keywords, which co-occurred a few times, such as explainable AI, wavelet transform, and Fourier transform, indicate future opportunities for integrating emerging technologies of trend AI methods with the AI-based stroke detection system. In addition, Figure 7 shows the keywords over the years. The use of machine learning models started around 2021. On the other hand, deep learning models emerged as new methods to utilize bio-signal data later.

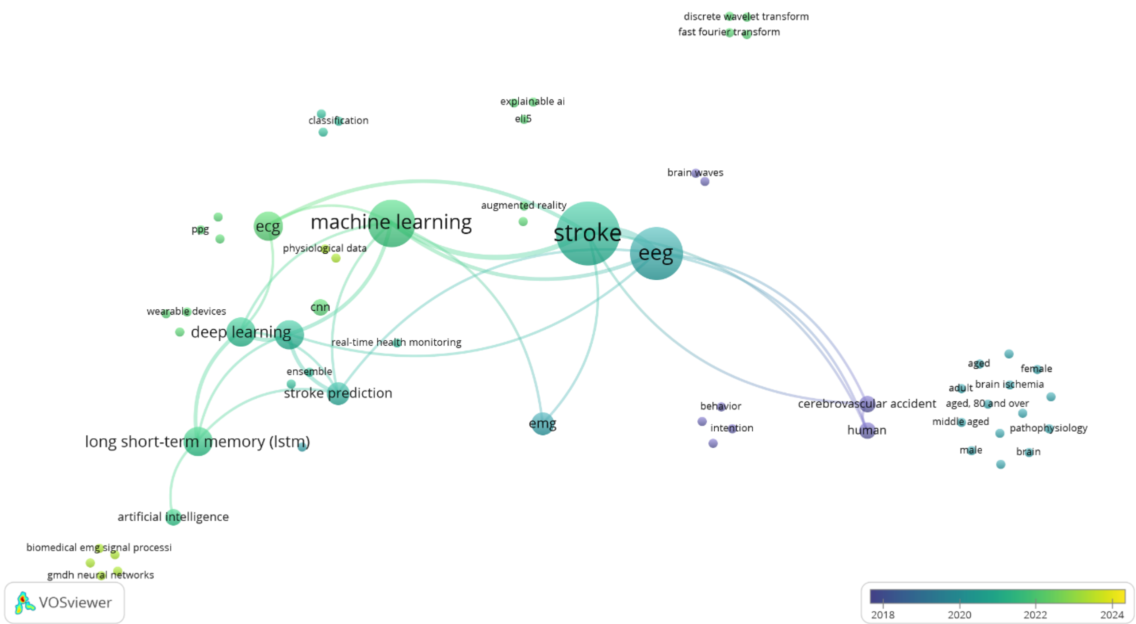

Figure 7. The distribution of keywords co-occurrence over years

5.2. Co-Citation

Co-citation analysis reveals the foundations of knowledge relied on by selected literature, in which clusters are formed based on the cited documents, which often occur together [30]. In this analysis, Table 6 presents the most cited sources that have been cited by five or more of the articles in Table 2. Figure 8 illustrates the co-citation network visualization among the top ten cited articles. The strongest links are between Stroke, Clinical Neurophysiology, and Sensors journals. We noticed that all the top ten cited journals are highly ranked journals based on SJR. In addition, Figure 8 reveals a new track of research and innovation, where neurology, wearable sensors, and AI are combined. This combination leverages the advantages of each, with neurology providing the clinical and physiological foundation, wearable sensors facilitating continuous and real-world data acquisition, and AI offering advanced analytical and predictive capabilities. Accordingly, neurological research and care from periodical, hospital-based assessments are shifting to continuous and personalized monitoring, which can be done remotely. The combination holds promise for early disease detection, not only stroke, but may extend to include long-term monitoring of neurodegenerative conditions, cognitive rehabilitation, and real-time mental health assessment. Even though these emerging communication fields introduce opportunities for innovation in digital health ecosystems, other challenges arise, such as data privacy and Explainable AI (XAI), as well as transparency.

Table 5. The top ten co-citation journals based on bibliometric analysis

Journal	Citation	Rank	Publisher
Stroke	33	Q1	American Heart Association
Clinical Neurophysiology	25	Q1	Elsevier
Sensors	19	Q1	MDPI
IEEE Access	17	Q1	IEEE
Journal of Stroke	11	Q1	Korean Stroke Society
Neuropsychologia	7	Q2	Elsevier
PloS One	7	Q1	Public Library of Science
Neurology ®	6	Q1	Wolters Kluwer
Cortex	5	Q1	Lippincott Williams and Wilkins
Applied Science	5	Q2	MDPI
Journal of Neuroscience Methods	5	Q2	Elsevier

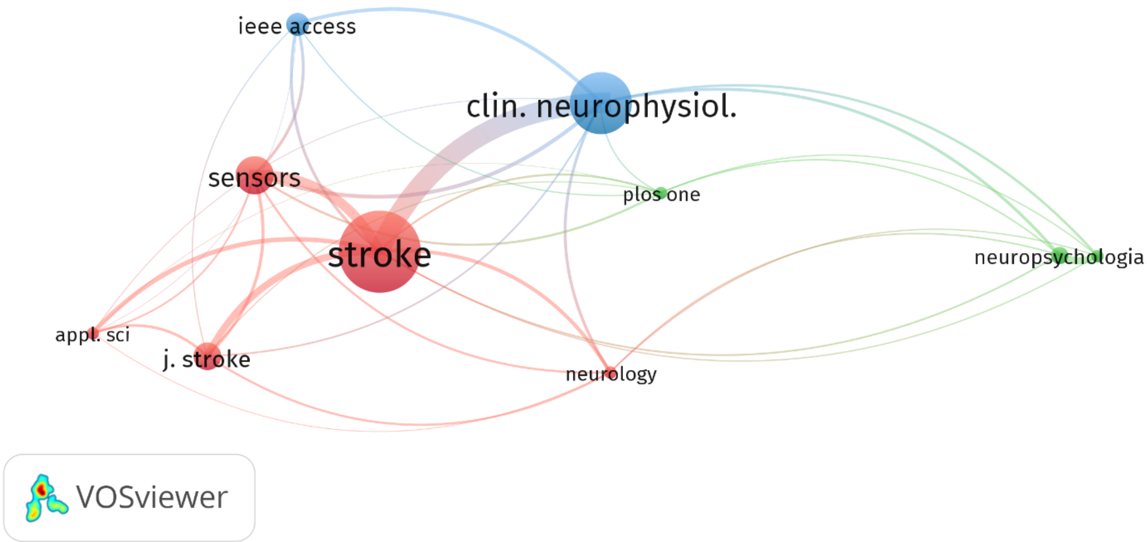

Figure 8. The top ten cited journals in co-citation network visualization

6. Discussion

This systematic review and bibliometric analysis paper reviews and visualizes DL and ML methods with bio-signals data to detect stroke and post-stroke effects. From 2016 to 2024, there was an interest in publishing studies; 2022 stands out as the year in which almost a third of the reviewed studies were published. Regarding where most reviewed studies were published, South Korea contributes the most in the stroke detection field. The most used bio-signal in the reviewed studies is EEG. Most of the reviewed studies gathered data from hospitals. Moreover, based on the sample size analysis, the largest sample size among the reviewed studies is 4068 [19]. On the other hand, the smallest sample size is 3 [29]. Notably, only one study did not mention the sample size clearly [18].

The most frequent ML methods used are RF and SVM. RF was used as the only method employed in both studies [16, 28]. In addition, the RF in studies [25, 26, 20] did not outperform other utilized methods in other studies. On the other hand, SVM was applied as a single method in the study [27]. However, it did not perform the best in the study [25]. Additionally, the most used DL methods are CNN and LSTM. CNN was exclusively applied in both studies [22, 18]. Nonetheless, it did not outperform other methods in the study [19]. Conversely, LSTM outperformed in both studies [19, 26].

DL and ML models were evaluated using metric measurements such as accuracy, precision, and recall. The studies depended on internal validation to ensure generalization ability. The internal validation techniques used were splitting the data into train and test, or cross-validation. Despite the differences in the datasets, the RF model outperforms all other models in terms of accuracy; either RF is used merely [28, 16], or RF was part of the proposed multimodel, such as [20, 24-26]. Studies such as [32, 33] show that the results of different ML/DL models may be artifacts of dataset size and preprocessing choices.

On the other hand, signal data suffers from complexity, nonstationarity, and high dimensionality [34]. Additionally, bio-signal datasets are considered time-series data. They are highly susceptible to interference from unrelated signals, such as eye blinks and muscle activity, which can serve as noise and yield high inter-individual variability [35, 36]. Nevertheless, some properties of ML/DL models can significantly enhance the results based on the characteristics of the bio-signal data. For example, the Random Forest (RF) model can handle the noise in signal datasets by aggregating the decisions across several sub-trees. Additionally, RF is considered a non-linear model, which enables it to work effectively with the signal dataset. Also, the RF model works well with small datasets, which is particularly applicable to datasets used in inclusion studies.

CNNs excel in spatial invariance, allowing them to detect patterns regardless of their location in the bio-signal, making them valuable for shift-invariant data. They generate a hierarchical representation of features, enabling the identification of complex patterns. CNNs eliminate the need for manual feature engineering, as they can learn and adapt to the unique qualities of the data. This automation simplifies signal data analysis, improves accuracy in tasks such as classification and regression, and enhances the power of CNNs for signal processing applications [37]. As bio-signal datasets are considered time-series data, LSTM is known as one of the DL models designed to learn dependencies from the data, yielding promising results with bio-signal datasets [19].

7. Conclusion

In conclusion, this systematic review and bibliometric analysis focused on the recent advancements in supervised machine learning (ML) and deep learning (DL) models for stroke detection and post-stroke effects detection using bio-signals. Most of the reviewed studies collected data from hospitals of varying sizes. The most frequently used ML methods were RF and SVM, while CNN and LSTM were the commonly employed DL methods. The model's performance was evaluated using metric measurements like accuracy, precision, and recall, with internal validation techniques such as data splitting and cross-validation. Future research should aim to overcome the limitations addressed in this systematic review by incorporating larger and more diverse datasets, conducting external validation on hospital experiments, and exploring more advanced AI techniques such as ensemble learning, explainable AI, and transfer learning. By addressing these critical aspects, the field of stroke detection and post-stroke effects detection will advance, and robust and reliable predictive models will be developed.

7.1. Reviewed Research Gaps

Our systematic review paper identifies several significant gaps in the existing literature that utilize ML and DL to detect strokes using bio-signals that future researchers could address.

First, the reviewed research highlights an important limitation in the geographic coverage aspect; most of them originated in South Korea. That limits the generalizability of findings to diverse populations. Hence, there is a crucial need for cross-cultural datasets and international collaborations. Second, although most of the reviewed studies gathered data from hospitals, no reviewed study reported the detection of stroke in real hospital workflows. That reveals a significant gap between experimental results and clinical applicability, suggesting the need to fill the absence of validation in real-world healthcare settings. Third, although the largest dataset size among all studies was around 4000, it is considered relatively small to train robust ML and DL models. Future research should consider a larger, high-quality clinical dataset with external validation to ensure reliability. Fourth, while multiple ML and DL models were developed, none of the included studies employed an ensemble learning technique that could introduce a promising detection result by combining robust models. Finally, the authors' keywords, which co-occurred a few times, such as explainable AI, wavelet transform, and Fourier transform, indicate future opportunities for integrating emerging technologies of trend AI methods with the AI-based stroke detection system.

Addressing these gaps will empower future studies and could deliver more reliable health care decisions toward brain stroke detection. This will require ensuring that they cover a larger, diverse clinical dataset with integration of real hospital workflows, and testing a variety of model enhancement techniques such as ensemble learning, explainable AI, wavelet transform, and Fourier transform.

8. Declarations

8.1. Author Contributions

Conceptualization, M.I.A.; methodology, M.I.A.; software, F.H.A., R.A.A., M.S.A., R.A.H.A., H.A.A., and S.F.A.; validation, D.A.A., S.S.A., and S.O.O.; formal analysis, M.I.A. and F.H.A.; investigation, R.A.A., M.S.A., R.A.H.A., H.A.A., and S.F.A.; resources, R.A.A., M.S.A., R.A.H.A., H.A.A., and S.F.A.; data curation, R.A.A., M.S.A., R.A.H.A., H.A.A., and S.F.A.; writing—original draft preparation, R.A.A., M.S.A., R.A.H.A., H.A.A., and S.F.A.; writing—review and editing, M.I.A., H.A.A., R.A.H.A., and F.H.A.; visualization, F.H.A. and A.A.; supervision, M.I.A.; project administration, M.I.A. All authors have read and agreed to the published version of the manuscript.

8.2. Data Availability Statement

Data sharing is not applicable to this article.

8.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

8.4. Institutional Review Board Statement

Not applicable.

8.5. Informed Consent Statement

Not applicable.

8.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

9. References

[1] Feigin, V. L., Brainin, M., Norrving, B., Martins, S., Sacco, R. L., Hacke, W., Fisher, M., Pandian, J., & Lindsay, P. (2022). World Stroke Organization (WSO): Global Stroke Fact Sheet 2022. International Journal of Stroke, 17(1), 18–29. doi:10.1177/17474930211065917.

[2] Upadhyay, S. K., Dan, S., & Ali, S. A. (2024). Evidence-Based Neurological Disorders: Symptoms, Causes, and Therapy. Jenny Stanford Publishing, New Yorkm, United States. doi:10.1201/9781003496342.

[3] Dehkharghani, S. (2021). Stroke. Exon Publications, Brisbane, Australia. doi:10.36255/exonpublications.stroke.2021.

[4] Stallard, J. (2019). CT Vs MRI: What's the Difference? and How Do Doctors Choose Which Imaging Method to Use? — Memorial Sloan Kettering Cancer Center, New York, United States. Available online: https://www.mskcc.org/news/ct-vs-mri-what-s-difference-and-how-do-doctors-choose-which-imaging-method-use (accessed on August 2025).

[5] Marsicano, G., Bertini, C., & Ronconi, L. (2024). Decoding cognition in neurodevelopmental, psychiatric and neurological conditions with multivariate pattern analysis of EEG data. Neuroscience and Biobehavioral Reviews, 164, 105795. doi:10.1016/j.neubiorev.2024.105795.

[6] Dykiert, I., Florek, K., Kraik, K., Gac, P., Poreba, R., & Poreba, M. (2024). Tpeak-Tend ECG Marker in Obesity and Cardiovascular Diseases: A Comprehensive Review. Scientifica, 4904508. doi:10.1155/2024/4904508.

[7] Taha, M. A., & Morren, J. A. (2024). The role of artificial intelligence in electrodiagnostic and neuromuscular medicine: Current state and future directions. Muscle and Nerve, 69(3), 260–272. doi:10.1002/mus.28023.

[8] Mehta, S., Kwatra, N., Jain, M., & McDuff, D. (2024). Examining the challenges of blood pressure estimation via photoplethysmogram. Scientific Reports, 14, 18318. doi:10.1038/s41598-024-68862-1.

[9] Katan, M., & Luft, A. (2018). Global Burden of Stroke. Seminars in Neurology, 38(2), 208–211. doi:10.1055/s-0038-1649503.

[10] Zen, H., Dasari, N. M., Latha, Y. M., & Rao, S. S. (2023). Soft Computing and Signal Processing. Proceedings of 6th ICSCSP, 2. doi:10.1007/978-981-99-8451-0.

[11] Siontis, K. C., Yao, X., Pirruccello, J. P., Philippakis, A. A., & Noseworthy, P. A. (2020). How Will Machine Learning Inform the Clinical Care of Atrial Fibrillation? Circulation Research, 127(1), 155–169. doi:10.1161/CIRCRESAHA.120.316401.

[12] Tsiamalou, A., Dardiotis, E., Paterakis, K., Fotakopoulos, G., Liampas, I., Sgantzos, M., Siokas, V., & Brotis, A. G. (2022). EEG in Neurorehabilitation: A Bibliometric Analysis and Content Review. Neurology International, 14(4), 1046–1061. doi:10.3390/neurolint14040084.

[13] Cerasa, A., Tartarisco, G., Bruschetta, R., Ciancarelli, I., Morone, G., Calabrò, R. S., Pioggia, G., Tonin, P., & Iosa, M. (2022). Predicting Outcome in Patients with Brain Injury: Differences between Machine Learning versus Conventional Statistics. Biomedicines, 10(9), 2267. doi:10.3390/biomedicines10092267.

[14] Murat, F., Sadak, F., Yildirim, O., Talo, M., Murat, E., Karabatak, M., Demir, Y., Tan, R. S., & Acharya, U. R. (2021). Review of deep learning-based atrial fibrillation detection studies. International Journal of Environmental Research and Public Health, 18(21), 11302. doi:10.3390/ijerph182111302.

[15] Choi, Y. A., Park, S. J., Jun, J. A., Pyo, C. S., Cho, K. H., Lee, H. S., & Yu, J. H. (2021). Deep learning-based stroke disease prediction system using real-time bio signals. Sensors, 21(13), 4269. doi:10.3390/s21134269.

[16] Choi, Y. A., Park, S., Jun, J. A., Ho, C. M. B., Pyo, C. S., Lee, H., & Yu, J. (2021). Machine-learning-based elderly stroke monitoring system using electroencephalography vital signals. Applied Sciences (Switzerland), 11(4), 1–18. doi:10.3390/app11041761.

[17] Islam, M. S., Hussain, I., Rahman, M. M., Park, S. J., & Hossain, M. A. (2022). Explainable Artificial Intelligence Model for Stroke Prediction Using EEG Signal. Sensors, 22(24), 9859. doi:10.3390/s22249859.

[18] Kumar, S., & Sengupta, A. (2022). EEG Classification for Stroke Detection Using Deep Learning Networks. 2nd International Conference on Emerging Frontiers in Electrical and Electronic Technologies, ICEFEET 2022, 9847883. doi:10.1109/ICEFEET51821.2022.9847883.

[19] Anand Kumar, M., Abirami, N., Guru Prasad, M. S., & Mohankumar, M. (2022). Stroke Disease Prediction based on ECG Signals using Deep Learning Techniques. Proceedings of International Conference on Computational Intelligence and Sustainable Engineering Solution, CISES 2022, 453–458. doi:10.1109/CISES54857.2022.9844403.

[20] Yu, J., Park, S., Kwon, S. H., Cho, K. H., & Lee, H. (2022). AI-Based Stroke Disease Prediction System Using ECG and PPG Bio-Signals. IEEE Access, 10, 43623–43638. doi:10.1109/ACCESS.2022.3169284.

[21] Mak, J., Kocanaogullari, D., Huang, X., Kersey, J., Shih, M., Grattan, E. S., Skidmore, E. R., Wittenberg, G. F., Ostadabbas, S., & Akcakaya, M. (2022). Detection of Stroke-Induced Visual Neglect and Target Response Prediction Using Augmented Reality and Electroencephalography. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 30, 1840–1850. doi:10.1109/TNSRE.2022.3188184.

[22] Kunwar, P., & Choudhary, P. (2023). A stacked ensemble model for automatic stroke prediction using only raw electrocardiogram. Intelligent Systems with Applications, 17, 200165. doi:10.1016/j.iswa.2022.200165.

[23] Elbagoury, B. M., Vladareanu, L., Vlădăreanu, V., Salem, A. B., Travediu, A. M., & Roushdy, M. I. (2023). A Hybrid Stacked CNN and Residual Feedback GMDH-LSTM Deep Learning Model for Stroke Prediction Applied on Mobile AI Smart Hospital Platform. Sensors, 23(7), 3500. doi:10.3390/s23073500.

[24] Lee, J., Kim, Y., & Kim, E. (2024). Data-Driven Stroke Classification Utilizing Electromyographic Muscle Features and Machine Learning Techniques. Applied Sciences (Switzerland), 14(18), 8430. doi:10.3390/app14188430.

[25] Rathakrishnan, K., Min, S. N., & Park, S. J. (2021). Evaluation of ECG features for the classification of post-stroke survivors with a diagnostic approach. Applied Sciences (Switzerland), 11(1), 1–16. doi:10.3390/app11010192.

[26] Yu, J., Park, S., Kwon, S. H., Ho, C. M. B., Pyo, C. S., & Lee, H. (2020). AI-based stroke disease prediction system using real-time electromyography signals. Applied Sciences (Switzerland), 10(19), 6791. doi:10.3390/app10196791.

[27] Laghari, A., Memon, Z. A., Ullah, S., & Hussain, I. (2018). Cyber physical system for stroke detection. IEEE Access, 6, 37444–37453. doi:10.1109/ACCESS.2018.2851540.

[28] Wilkinson, C. M., Burrell, J. I., Kuziek, J. W. P., Thirunavukkarasu, S., Buck, B. H., & Mathewson, K. E. (2020). Predicting stroke severity with a 3-min recording from the Muse portable EEG system for rapid diagnosis of stroke. Scientific Reports, 10(1), 18465. doi:10.1038/s41598-020-75379-w.

[29] Choi, J., Kang, H., Chung, S. H., Kim, Y., Lee, U. H., Lee, J. M., Kim, S. J., Chun, M. H., & Kim, H. (2016). Detecting voluntary gait intention of chronic stroke patients towards top-down gait rehabilitation using EEG. Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, 2016-October, 1560–1563. doi:10.1109/EMBC.2016.7591009.

[30] Lim, W. M., Kumar, S., & Donthu, N. (2024). How to combine and clean bibliometric data and use bibliometric tools synergistically: Guidelines using metaverse research. Journal of Business Research, 182, 114760. doi:10.1016/j.jbusres.2024.114760.

[31] Waltman, L., van Eck, N. J., & Noyons, E. C. M. (2010). A unified approach to mapping and clustering of bibliometric networks. Journal of Informetrics, 4(4), 629–635. doi:10.1016/j.joi.2010.07.002.

[32] Riley, R. D., Ensor, J., Snell, K. I. E., Archer, L., Whittle, R., Dhiman, P., Alderman, J., Liu, X., Kirton, L., Manson-Whitton, J., van Smeden, M., Moons, K. G., Nirantharakumar, K., Cazier, J. B., Denniston, A. K., Van Calster, B., & Collins, G. S. (2025). Importance of sample size on the quality and utility of AI-based prediction models for healthcare. The Lancet Digital Health, 7(6), 100857. doi:10.1016/j.landig.2025.01.013.

[33] Zantvoort, K., Nacke, B., Görlich, D., Hornstein, S., Jacobi, C., & Funk, B. (2024). Estimation of minimal data sets sizes for machine learning predictions in digital mental health interventions. NPJ Digital Medicine, 7, 361. doi:10.1038/s41746-024-01360-w.

[34] Belhadi, A., Lind, P. G., Djenouri, Y., & Yazidi, A. (2025). Enhanced visibility graph for EEG classification. Frontiers in Neuroscience, 19. doi:10.3389/fnins.2025.1541062.

[35] Becerra, M. A., Duque-Mejia, C., Castro-Ospina, A., Serna-Guarín, L., Mejía, C., & Duque-Grisales, E. (2025). EEG-Based Biometric Identification and Emotion Recognition: An Overview. Computers, 14(8), 299. doi:10.3390/computers14080299.

[36] Stancin, I., Cifrek, M., & Jovic, A. (2021). A review of EEG signal features and their application in driver drowsiness detection systems. Sensors, 21(11), 3786. doi:10.3390/s21113786.

[37] Rajwal, S., & Aggarwal, S. (2023). Convolutional Neural Network-Based EEG Signal Analysis: A Systematic Review. Archives of Computational Methods in Engineering, 30(6), 3585–3615. doi:10.1007/s11831-023-09920-1.

Appendix I

Table A1. Customized checklist for ML and DL prediction models for bio-signal data

Ref. Section/ Topic	Item	Checklist Item	Page
		Title and Abstract	
Title	1	Identify the study as developing and/or validating a prediction model.	
Abstract	2	Provide a summary of objectives, study design, results, and conclusions.	
		Introduction	
Objectives	3	Specify the objectives and aims, including whether the study describes the development or validation of the model or both.	
		Methods and Results	
Data	4a	Specify the source of data.	
	4b	Specify the type of data (e.g., EEG, MEG, ECG, PPG).	
Participants	5	Specify participants' condition (e.g., patients, health).	
Sample Size	6	Explain how the study size arrived at.	
Model	7	Clearly define all techniques and algorithms used in developing and/or validating the pretrained model.	
Outcomes	8	Clearly define the outcome that is predicted by the prediction model.	
	8a	Report on performance measures.	
		Other Information	
Code	9	Provide the code of the study.	

Table A2. Data extraction form

Extracted item	Comments
Author	Name of authors, e.g. Laghari et al. [27]
Objective	Specify the objectives and aims.
Source of data	Specify the source of data, e.g., Hospital name.
Data type	Answer categories: • EEG signal. • EMG signal. • ECG signal. • PPG signal.
Sample Size	Sample size used for building the model.
Techniques	List all machine learning / deep learning algorithms used.
Outcome	List the performance measures used.
Region under study	Specify the region under the study.
Published year	Published year of the study.

Table A3. Quality assessment data for each study

Study	Checklist items										
Title and Reference	Title and Abstract		Introduction	Methods and Results							Other info.
	Title	Abstract	Objectives	Data		Participants	Sample size	Models	Outcome		Code
	1	2	3	4a	4b	5	6	7	8a	8b	9
Cyber Physical System for Stroke Detection [27]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Predicting stroke severity with a 3-min recording from the Muse portable EEG system for rapid diagnosis of stroke [28]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Deep Learning-Based Stroke Disease Prediction System Using Real-Time Bio Signals [15]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Machine-Learning-Based Elderly Stroke Monitoring System Using Electroencephalography Vital Signals [16]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Explainable Artificial Intelligence Model for Stroke Prediction Using EEG Signal [17]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No

EEG Classification for Stroke Detection Using Deep Learning Networks [18]	Yes	Yes	Yes	Yes	Yes	No	Unknown	Yes	Yes	Yes	No
Evaluation of ECG Features for the Classification of Post-Stroke Survivors with a Diagnostic Approach [25]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Stroke Disease Prediction based on ECG Signals using Deep Learning Techniques [19]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
A Stacked Ensemble Model for Automatic Stroke Prediction Using only Raw Electrocardiogram [22]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
AI-Based Stroke Disease Prediction System Using Real-Time Electromyography Signals [26]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
A Hybrid Stacked CNN and Residual Feedback GMDH-LSTM Deep Learning Model for Stroke Prediction Applied on Mobile AI Smart Hospital Platform [23]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
AI-Based Stroke Disease Prediction System Using ECG and PPG Bio-Signals [20]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Detecting Voluntary Gait Intention of Chronic Stroke Patients Towards Top-Down Gait Rehabilitation Using EEG [29]	Yes	Yes	Yes	Unknown	Yes	Yes	Yes	Yes	Yes	Yes	No
Detection of Stroke-Induced Visual Neglect and Target Response Prediction Using Augmented Reality and Electroencephalography [21]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Data-Driven Stroke Classification Utilizing Electromyographic Muscle Features and Machine Learning Techniques [24]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No

Table A4. The study author, objective, source of data, type of data, sample size, techniques, outcome, region under study, published year

Author	Objective	Data source	Data type	Sample size	Techniques	Outcome	Region Pub.	Year
Laghari et al. [27]	Reveal the occurrence of stroke in patients who have previously survived a stroke or who have a high risk of experiencing stroke.	Shaheed Mohtarma Benazir Bhutto Medical University.	EEG Signals.	30	ML: SVM.	SVM: Precision: 100% Recall: 99.16%	Pakistan	2018
Wilkinson et al. [28]	Utilizing an inexpensive portable EEG device as a method for prehospital stroke and examining whether EEG data can be used to detect changes in stroke intensity.	University of Alberta Hospital.	EEG Signals.	25	ML: RF - Muse by InteraXon Inc., a device to record EEG signals.	RF: Accuracy: 76% Sensitivity: 63% Specificity: 86%	Canada	2020
Choi et al. [15]	Presented a new way for applying deep learning models to raw EEG data without relying on the frequency features of EEG.	Emergency Medical Center of Chungnam National University Hospital.	EEG Signals.	273	DL: LSTM, Bidirectional LSTM, CNN-LSTM, CNN-Bidirectional LSTM.	Raw values showed the best accuracy. LSTM: Accuracy: 70.1% Bidirectional LSTM: Accuracy: 91.8% CNN-LSTM: Accuracy: 93.7% CNN-Bidirectional LSTM: Accuracy: 94.0%	South Korea	2021
Choi et al. [15]	Developing a health monitoring system that can anticipate the symptoms of stroke diseases in old people in real-time while they walk on a regular basis.	Emergency Medical Center and Rehabilitation Department of Chungnam National University Hospital.	EEG Signals.	273	ML: RF.	RF: Accuracy 92.51%	South Korea	2021
Islam et al. [17]	Classifying stroke and healthy control groups for stroke prediction in active situations.	Korea Research Institute of Standards and Science.	EEG Signals.	123	ML: AdaBoost, XGBoost, and LightGBM.	AdaBoost: Accuracy: 80%	South Korea	2022
Kumar & Sengupta [18]	Develop models to categorize EEG signals as strokes or non-strokes.	Normal and abnormal EEG activity from PhysioNet.	EEG Signals.	Unknown	Deep neural network architecture, RESNET-50, and VGG-16.	RESNET-50: Accuracy: 90% Sensitivity: 100% VGG-16: Accuracy: 90% Specificity: 100% Precision: 100%	Unknown	2022
Rathakrishnan et al. [25]	Develop a classification model using machine learning and ECG signals for diagnosing stroke disease.	Chungnam National Hospital, Daejeon, South Korea.	ECG Signals.	132	ML: SVM, RF, Naïve Bayes, KNN, and LR.	KNN: Accuracy: 96.6% RF: Accuracy: 94.4% SVM: Accuracy: 85.4% Naïve Bayes: Accuracy: 72.7% LR: Accuracy: 66.9%	South Korea	2020
Kumar et al. [19]	Proposing a medical framework to detect abnormalities in the ECG associated with stroke disease.	Indian hospitals.	ECG Signals.	4068	DL: LSTM, CNN, and RNN.	LSTM: Accuracy: 93.78% CNN: Accuracy: 89.25% RNN: Accuracy: 86.19%	India	2022
Kunwar & Choudhary [22]	Developing a classification model based on ECG signals for stroke diagnosis.	The cerebral vasoregulation Dataset.	ECG Signals.	71	Stacking ensemble model of CNN models.	CNN: Accuracy: 99.7% F1: 99.69% Recall: 99.71% Precision: 99.67%	USA	2023

Yu et al. [26]	Developing a stroke prediction system with the use of real-time EMG signals.	Emergency Medical Center and the Department of Rehabilitation Medicine at Chungnam National University Hospital	EMG Signals.	558	ML: RF. DL: LSTM.	RF: Accuracy: 90.38% LSTM: Accuracy: 98.96%	South Korea	2020
Elbagoury et al. [23]	Proposing a telemedicine system that predicts heart, and brain stroke.	EMG Lower Limb Dataset mHealth Dataset EMG Physical Action Dataset.	EMG Signals.	38	DL: Stacked CNN + LSTM + GMDH. Explainable AI (XAI).	Stacked CNN + LSTM + GMDH: Accuracy: 99%	Unknown	2023
Yu et al. [20]	Develop multi-models based on ML, ECG, and PPG signals.	Emergency medical center and department of rehabilitation medicine at Chungnam National University Hospital	ECG and PPG Signals.	574	DL: an ensemble structure that combines CNN and LSTM. ML: Decision Tree, RF	Decision Tree: Accuracy: 91.56% RF: Accuracy: 97.51% CNN-LSTM: Accuracy: 99.15%	South Korea	2022
Choi et al. [29]	Decoding stroke patients' gait intentions using EEG signals.	Unknown	EEG Signals.	3	Linear Discriminant Analysis (LDA).	LDA: Accuracy: 73.2% Delay is 0.13 s	Unknown	2016
Mak et al. [21]	Proposing system combines EEG data and Augmented Reality (AR) to identify the presence of Visual-Spatial Neglect (SN) in stroke patients.	University of Pittsburgh Medical Center Inpatient Rehabilitation.	EEG Signals.	226	(RDA+KDE) Classifier	(RDA+KDE): Average train AUC: 0.788 Average test AUC: 0.760	USA	2022
Lee et al. [24]	Examine the impact of the statistical features of muscle activity of the major leg muscles during gait as predictive factors across various models to differentiate between stroke patients and healthy individuals.	Multiple medical institutions across South Korea.	EMG Signals.	240	DT, RF, LR, MLP, SVC, K-NN, NB.	RF:Accuracy: 100% LR: Accuracy: 96% DT: Accuracy: 94% MLP:Accuracy: 99% SVM:Accuracy: 94% NB:Accuracy: 77% KNN:Accuracy 85%	South Korea	2024