

ORIGINAL RESEARCH ARTICLE

An explainable deep learning-based computational framework for stroke prediction using multimodal radiomics features: A retrospective machine learning study

Hayder M. A. Ghanimi^{1,2}, Dharani Kumar Sunkara Venkata³, Venu Karunanithi⁴, Vidya Sagar Ponnamm⁵, Vedaraj Muthuraj⁶, Vivekanandhan Vijayarangan⁷, Aseel Smerat⁸, and Sudhakar Sengan^{9*}

¹Department of Information Technology, College of Science, University of Warith Al-Anbiyaa, Karbala, Iraq

²Department of Computer Science, College of Computer Science and Information Technology, University of Kerbala, Karbala, Iraq

³Department of Biomedical Engineering, GRT Institute of Engineering and Technology, Chennai, Tamil Nadu, India

⁴Department of Computer Science and Engineering, Kongu Engineering College, Perundurai, Tamil Nadu, India

⁵Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India

⁶Department of Computer Science and Engineering, Panimalar Engineering College, Chennai, Tamil Nadu, India

⁷Department of Information Technology, Vel Tech High Tech Dr. Rangarajan Dr. Sakunthala Engineering College, Chennai, Tamil Nadu, India

⁸Department of Educational Sciences, Al-Ahliyya Amman University, Amman, Jordan

⁹Department of Computer Science and Engineering, Erode Sengunthar Engineering College Erode, Tamil Nadu, India

***Corresponding author:**
Sudhakar Sengan
(sudhakars@esec.ac.in)

Citation: Ghanimi HMA, Venkata DKS, Karunanithi V, *et al.* An explainable deep learning-based computational framework for stroke prediction using multimodal radiomics features: A retrospective machine learning study. *Artif Intell Health*.

doi: 10.36922/AIH026270078

Received: June 30, 2026

Revised: August 6, 2026

Accepted: August 11, 2026

Published online: August 27, 2026

Copyright: © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Abstract

Stroke prediction models (SPMs) serve two important roles: accurately recognizing stroke symptoms and informing treatment decisions. The Synthetic Minority Oversampling Technique is typically used for analysis. An improved deep learning model for multimodal radiomic features and data preprocessing on imbalanced datasets is recommended. The chaotic map-guided approach (CMGA) has been applied to generate a feature-importance score, designed to improve the predictability and performance of a traditional model for feature selection (FS). The current study proposes a new butterfly-based spiking neural network (BB-SNN) to improve classification accuracy and reduce computational time. The efficiency of FS increased by 12% with CMGA-based scoring, and the model's interpretability improved. The proposed SPM achieved 94.80% accuracy, 93.50% sensitivity, and 95.30% specificity, outperforming traditional methods. The deployment of the BB-SNN reduced the training time by 18%, demonstrating the method's efficiency. The comparative analysis of the proposed SPM showed its efficacy in handling radiomics data and addressing class imbalance, making it suitable for the SPM. Using CMGA and BB-SNN substantially enhanced medical diagnostics by effectively predicting strokes. These findings show that the proposed model can be integrated

with diagnostic systems to improve the quality of patient care, as it effectively predicts stroke rates.

Keywords: Bistable spiking neural networks; Synthetic Minority Oversampling Technique; Stroke prediction models; Feature selection; Chaotic map-guided approach

1. Introduction

Stroke is one of the leading causes of mortality and long-term disability worldwide, imposing a substantial burden on individuals, healthcare systems, and economies. Given the significant public health and socioeconomic impact of stroke, accurate prediction of stroke outcomes is essential to enable timely clinical intervention and improve patient management.¹ Early identification of patients at risk of adverse outcomes may support appropriate clinical decision-making, mitigate long-term complications, and facilitate more effective preventive and rehabilitative strategies.²

Different types of imaging data, such as magnetic resonance imaging (MRI) and computed tomography (CT), have provided useful data sources for stroke predictions (SPs). Such datasets are typically complex and, when used correctly, can enhance SP.³ Machine learning (ML) and deep learning (DL) are currently effective computational methods for analyzing radiomics data and extracting useful information. However, their application is subject to limitations, primarily due to the high dimensionality of radiomics features, which can lead to overfitting and slow computation.

Stroke has a high degree of clinical and etiological variability, as well as imaging and functional outcomes. Therefore, accurate SP requires patient stratification based on multi-dimensional features (e.g., demographic, clinical, laboratory, and radiomics). Such a comprehensive layer can be helpful in precision medicine by enabling personalized risk assessment, optimal use of healthcare resources, and the implementation of a clinical decision support system (CDSS). Recent studies have shown that a multimodal patient model incorporating patient-specific data achieved higher prediction accuracy than models based solely on imaging.⁴ Epidemiological studies have shown that acute stroke is one of the main causes of death and disability globally.⁵ Earlier population-based studies have shown cumulative incidence rates of nearly 218 per 100,000 men and 127 per 100,000 women, highlighting the scope of the health care problem and the importance of early prediction. Stroke has several pathological subtypes: atherothrombotic, cardioembolic, lacunar, infarction of

unusual etiology, and intracerebral hemorrhage. These subtypes have distinct clinical presentations and imaging features, underscoring the importance of accounting for subtype variability in future prediction models.⁶

Stroke datasets are typically imbalanced, with few positive stroke samples compared to negative stroke samples.⁷ This drawback is a problem for traditional ML, where the majority class is mostly predicted. These problems should be addressed to develop a sound, effective, and accurate SP model (SPM) that improves the quality of health care services.⁸ DL and other computational techniques are dynamic in SP because they can process large and diverse radiological data obtained from MRI and CT scans.⁹ They can learn high-dimensional features that can capture more details than other methods can detect. Effective detection of early signs of stroke is possible with convolutional neural networks (CNNs) and spiking neural networks (SNNs).¹⁰ Furthermore, feature selection (FS) and class balancing techniques like Synthetic Minority Oversampling Technique (SMOTE) have also been integrated, as noisy features and imbalanced datasets can impact model accuracy and on-time CDSS.¹¹ There is significant clinical and etiological variability of stroke, as well as imaging and functional outcomes. Thus, for accurate prediction of stroke, patient stratification by using multi-dimensional features (such as demographic, clinical, laboratory, and radiomics) is essential. This detailed stratification can support precision medicine by providing individualized risk assessment, efficient utilization of healthcare resources, and CDSS. Recent studies indicate that a multimodal patient model including patient-specific information improves prediction accuracy over imaging-based models alone.

The current study aims to improve DL for SPM by enhancing preprocessing, feature extraction (FE), and classification to overcome the limitations outlined in previous studies.^{12,13} SMOTE helps address class imbalance by generating synthetic examples of minority classes, enabling the model to learn from a better-balanced training set. FS is improved by using chaotic maps to generate feature importance scores¹⁴, thereby leveraging randomness to avoid local optima and identify better

feature subsets. Butterfly optimization algorithm (BOA) optimizes the SPM by adjusting the SNN classifier weights. By following butterflies' natural search behavior, the combination of BOA and SNN yields better classification performance and lower time complexity. This makes SNN dynamic, event-driven, energy-efficient, and biologically plausible.

The proposed butterfly-based SNN (BB-SNN) model is designed to achieve high performance in SPM by leveraging the best FS, appropriate preprocessing, and optimized neural network classification. The proposed BB-SNN has been compared with traditional methods and is suitable for clinical practice. The study aims to enhance health diagnosis by providing a computationally efficient and accurate prognosis of stroke, thereby improving patients' health outcomes. Stroke is a significant global problem, and there is a need for effective SPM to prevent adverse effects and deaths. SPM is challenging due to the high dimensionality of radiomics data and dataset imbalance, which negatively impact ML performance. Some challenges include the lack of an efficient, accurate method for handling noisy features, the challenge of selecting the best features, and the computational expense of neural networks.

We proposed a reliable and effective SPM that is unaffected by the aforementioned limitations. Multimodal radiomics (MMR) and feature-importance score generation using the Chaotic map-guided approach (CMGA) were used to enhance FS. Moreover, the application of BOA to weight tuning in SNN is relevant to classification and computational problems. With these improvements, the proposed BB-SNN could improve SP accuracy, enabling timely diagnosis and enhancing the use of CDSS. The research objectives include:

- (i) To design an efficient MMR-based SPM to improve SP by processing radiomics data.
- (ii) To overcome the problem of imbalanced datasets in stroke by using SMOTE.
- (iii) To apply FS using chaotic maps to improve model interpretability and reduce the computational time required by the model.
- (iv) To enhance classification by applying BOA for weight optimization in SNN.
- (v) To compare the system's performance with existing methods in terms of accuracy, sensitivity, specificity, and time complexity.
- (vi) To achieve fast, accurate SP in clinical practice, the system must be lightweight and computationally efficient.

The article is divided into five parts. **Section 1** explains the importance of accurate SPM, and **Section 2** presents

related works that reveal limitations in methods for addressing class imbalance and FS. **Section 3** presents the methodology, in which the dataset is balanced using SMOTE, FS is performed using CMGA, and weight optimization is conducted using BOA within the SNN. **Section 4** supports the data, signifying that the proposed model achieves higher accuracy and efficiency. **Section 5** presents the conclusion and references for future research.

2. Related works

Stroke is one of the most significant health problems in terms of morbidity and mortality, and thus, there is a need for a reliable SPM for proper patient management and prognosis. Using ML and the latest trends in radiomics, extra-dimensional measures can be derived from imaging and clinical data to improve risk assessment for SP. These practices provide a standardized assessment of factors such as perihematomal edema and hematoma expansion, thereby improving predictive ability.

This section provides an overview of recent research on the development of ML-based SPM. Initially, radiomics FE was used to describe image features; in particular, a CNN that employed transfer learning was combined with state-of-the-art segmentation to predict adverse outcomes and CDSS. Moreover, combining clinical features with radiomics outperforms models based solely on clinical values. [Table 1](#) summarizes the studies reviewed, including the methods used, the results, and the ML used.

Several key challenges must be addressed in developing the SPM, including incorporating MMR data, handling imbalanced data, and balancing computational time and SP accuracy. The limitations of traditional ML and DL include FS, high computational cost, and poor portability. Many current methods fail to consistently manage radiomics data and employ inadequate optimization, thus reducing their precision and reliability. Moreover, the class-imbalanced stroke datasets affect the performance of conventional models and, therefore, they cannot be relied upon for clinical application.

The following gaps are addressed in the proposed model: MMR is combined with preprocessing methods, including SMOTE, to mitigate class imbalance. Scoring features based on CMGA enhances FS efficiency and increases the model's interpretability. A novel BB-SNN reduces computational cost while enhancing classification performance. This optimization mechanism enables more efficient learning, making SNNs suitable for medical diagnostics. The comparison also highlights the strength and efficiency of the proposed model in handling radiomics data and addressing class imbalance. These boosts position the proposed solution as a computationally efficient and

Table 1. Comprehensive analysis of SPM

Reference	Techniques used	Inference	ML
15	Radiomics FE, ML classifiers	Adding PHE radiomics improved the risk test but did not significantly improve predictive performance	Multiple classifiers, including traditional ML
16	Transfer learning (TL) with CNN, ResNet50, Inception_v3	CNN integrated with radiomics improved hematoma expansion prediction	CNN, Radiomics-Inception_v3, Radiomics-ResNet50
17	Multi-PLD ASL, radiomics, ML-based FS	The combined model presented superior predictive performance for AIS prognosis	XGBoost
18	Radiomics FE, ITK-Snap, AI Kit	Developed a nomogram for SP recurrence in TIA patients	LR
19	Clinical-radiomics nomogram	Achieved high accuracy in predicting recurrent ischemic events in TIA patients	Decision curve analysis
20	Stroke registry analysis, SHAP feature analysis	XGBoost identified key clinical variables for early diagnosis	XGBoost, RF, SVM, CART
21	ML-based FS with SHAP analysis	High AUC and accuracy for mortality prediction in stroke patients	XGBoost
22	LR, Nomogram	Developed models predicting 3-month unfavorable outcomes	Extra trees, CatBoost, RF
23	DT, predictive modeling	Improved prediction of adverse outcomes for patients experiencing reperfusion therapy	DT, CatBoost
24	Radiomics and AI-based segmentation methods	Identified predictors for ischemic and hemorrhagic results in stroke patients	Ensemble ML

Abbreviations: AI: Artificial intelligence; AIS: Acute ischemic stroke; ASL: Arterial spin labeling; AUC: Area under the curve; CART: Classification and regression tree; CNN: Convolutional neural network; DT: Decision tree; FE: Feature extraction; FS: Feature selection; ITK: Interactive toolkit; LR: Logistic regression; ML: Machine learning; PHE: Perihematomal Eedema; PLD: Post-labeling delay; RF: Random forest; SHAP: SHapley Additive exPlanations; SP: Stroke prediction; SPM: Stroke prediction model; SVM: Support vector machine; TIA: Transient ischemic attack; XGBoost: Extreme gradient boosting.

accurate SPM that effectively addresses problems inherent in current models and therefore has the potential for clinical implementation.

Table 2 shows a few studies related to artificial intelligence (AI), ML, DL, Internet of Things (IoT), computer vision, medical imaging, biosensors, robotics, and healthcare analytics. These studies cover a wide range of application areas. The studies demonstrated that the use of modern AI algorithms such as CNNs, bidirectional long short-term memory, transformer networks, graph embeddings, ensemble learning, and quantum-inspired DL substantially improved prediction accuracy, disease diagnosis, medical image analysis, intelligent healthcare systems, rehabilitation, and human-computer interaction. Emerging technologies, such as blockchain, IoT, and multimodal sensing, have been integrated in several studies to enhance healthcare security, automation, and CDSS. The body of research indicates a clear shift towards hybrid, explainable, and intelligent data-driven models. These models incorporate several computational methods to solve complex real-world problems while improving prediction accuracy, automation, and CDSS functions.

3. Proposed methodology

The proposed methodology used state-of-the-art ML to improve data preprocessing, FS, and classification efficiency for SPM. To address class imbalance in the dataset, SMOTE was applied to improve the model's learning from minority classes, rather than being dominated by the majority class. The proposed model performed binary classification of SP rate using multimodal radiomics and demographic features. A CMGA was used to generate the FS score, improving FS efficiency and, consequently, the model's interpretability by identifying the most remarkable features for SP. The new BB-SNN proposed for the classification task reduced computational cost and improved model convergence, making SNN more suitable for medical diagnosis. Integrating SMOTE, CMGA-based FS, and butterfly-based optimization in SNN formed the proposed BB-SNN, a reliable and efficient solution for accurate SPM.

Figure 1 shows the entire computational process of the proposed explainable DL model for precision SP. The workflow started by finding a publicly available electronic health record (EHR) dataset; thereafter, data preprocessing steps were performed, including data cleaning, missing-

Table 2. Comparative summary of representative studies on AI-based healthcare and intelligent systems

Ref.	Research area	Method/model	Key contribution
25	Cognitive modeling	DL + mathematical cognitive modeling	Developed a cognitive framework to model complex sequential human decision-making.
26,27	SP	Ensemble ML	SP risk using unstructured healthcare data with an ensemble cognitive model.
28	Sign language recognition	MP + CNN + BiLSTM	Developed an open-source smartphone-based sign language recognition model.
29,30	Healthcare security	Blockchain + Merkle–Damgård hashing	Improved security and integrity of EHR using blockchain.
31	Cardiovascular disease	DL + ML	Automated chronic heart failure detection with high diagnostic accuracy.
32	Healthcare IoT security	Secure scheduling algorithm	Proposed an adversarially robust scheduling algorithm for secure healthcare big data analytics.
33	Lung cancer	VGG-19 DL	Enhanced lung cancer detection using TL.
34	Biometrics	Quantum dilated CNN	Introduced an attention-based quantum CNN for long-distance iris recognition.
35	Underwater imaging	CNN–transformer + optimization	Improved underwater low-light image quality using multi-scale feature fusion.
36	Prostate cancer	ResNet50 hybrid vision model	Automated histopathological assessment for prostate cancer diagnosis.
37	Liver transplantation	Graph embedding + counterfactual learning	Improved transplantation outcome prediction using graph-based representations.
38	Colorectal cancer	DL + biosensor analysis	Performed benchmarking of AI and biosensor methods for diagnosis and prognosis.
39,40	Rehabilitation medicine	Neuromuscular training	Evaluated rehabilitation strategies for ankle sprain recovery in athletes.
41	HIV healthcare	Psychometric validation	Validated the Indonesian version of the Coping Self-Efficacy Scale.
42	Nursing care	Clinical case study	Provided methods for addressing spiritual care requirements in intensive care.
43	Maternal and child health	Survey analysis	Identified determinants influencing caregivers’ knowledge of infant nutrition.
44	Autism detection	CNN–GRU + wavelet optimization	Developed an EEG-based hybrid model for automated autism detection.
45	Brain–computer interface	EEG signal processing	Processing EEG brain signals to detect and interpret neural activity for controlling or communicating with external devices.

Abbreviations: AI: Artificial intelligence; BiLSTM: Bidirectional long short-term memory; CNN: Convolutional neural network; DL: Deep learning; EEG: Electroencephalography; EHR: Electronic health record; HIV: Human immunodeficiency virus; IoT: Internet of Things; ML: Machine learning; MP: Mean prediction; SP: Stroke prediction; TL: Transfer learning; VGG: Visual geometry group.

value handling, normalization, and train-test splitting. The training dataset was balanced by using SMOTE. The CMGA was then applied to extract multimodal radiomics and clinical features and to rank them to identify the most informative features. The proposed BB-SNN was then used for feature classification to improve prediction performance using the selected features. The model was tested using several performance metrics, including accuracy, sensitivity, specificity, precision, *F1*, Matthews correlation coefficient (MCC), and the Area under the

receiver operating characteristic (ROC) curve (AUC-ROC). Lastly, the explainability module provided feature importance scores and clinically interpretable insights, enabling clear SP and supporting the CDSS.

3.1. Synthetic Minority Oversampling Technique for preprocessing of imbalanced data

The SMOTE balanced the dataset by generating synthetic data points for the minority class, which, in the case of SPM, corresponded to patients who had experienced a stroke. To

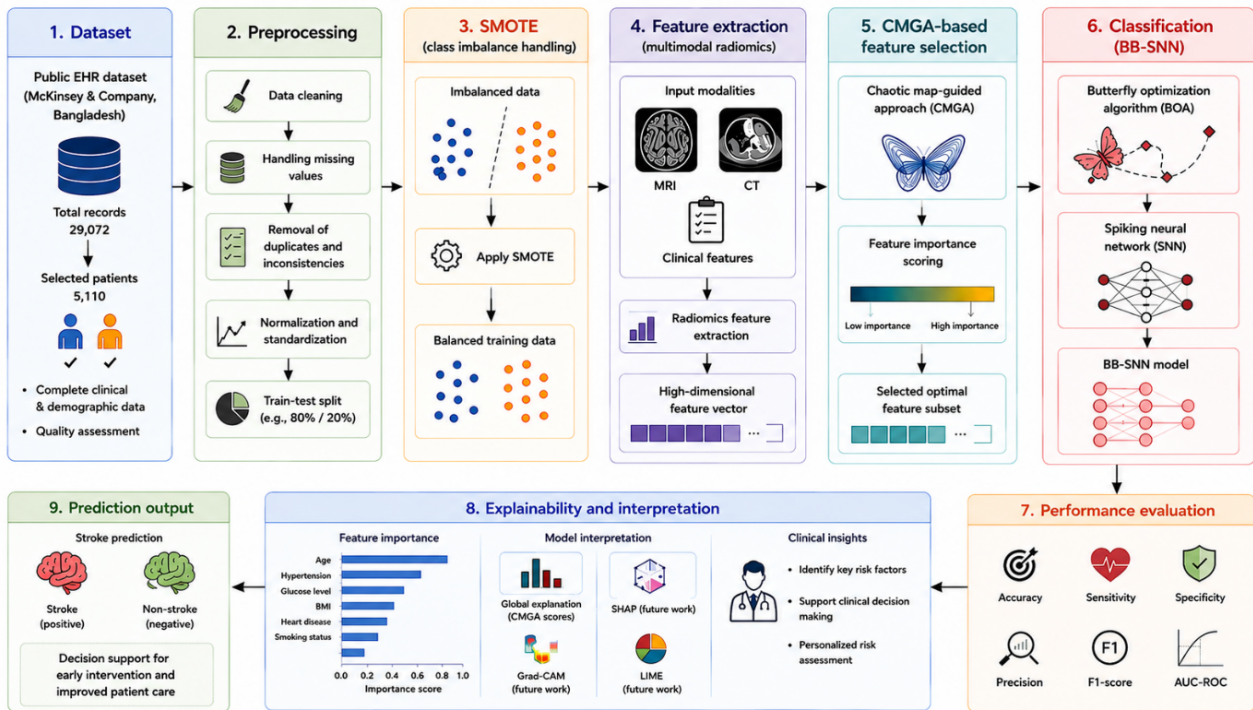


Figure 1. Workflow of the proposed explainable DL framework for precision SP

Abbreviations: AUC: Area under the curve; BB-SNN: Butterfly-based spiking neural network; BMI: Body mass index; CAM: Class activation mapping; CT: Computed tomography; EHR: Electronic health record; LIME: Local interpretable model-agnostic explanations; MRI: Magnetic resonance imaging; ROC: Receiver operating characteristic; SHAP: SHapley Additive exPlanations; SMOTE: Synthetic Minority Oversampling Technique.

prevent data leakage, the dataset was split into training and test sets. SMOTE was applied exclusively to the training data, while the testing dataset remained unchanged. Although alternative imbalance learning methods, such as adaptive synthetic sampling, focal loss, and cost-sensitive learning, are available, SMOTE was selected for its stability and wide adoption in medical classification studies. This helped in mitigating the bias caused by the large number of healthy individuals compared to those who had a stroke. For an assumed minority class instance x_i (e.g., a stroke patient), its k -nearest neighbor was calculated based on the Euclidean distance in the feature space specified in **Equation 1**:

$$Distance(x_i, x_j) = \sqrt{\sum_{p=1}^m (f_{ip} - f_{jp})^2} \quad (1)$$

where,

- x_i, x_j are two instances of the minority class.
- $f_{ip} - f_{jp}$ are the feature values of the p^{th} attribute of the respective instances.

For each selected neighbor x_j , a new synthetic sample

was formed by selecting a random value between 0 and 1, denoted as λ , and generating a synthetic sample x_{new} as assumed in **Equation 2**:

$$x_{new} = x_i + \lambda(x_j - x_i) \quad (2)$$

This interpolation generated a synthetic point by considering the differences in feature values between and its neighbor. The process was repeated for each instance of the minority class, generating multiple synthetic samples until the class distribution is balanced. The generation of the new dataset is given in **Equation 3**:

$$New\ dataset = \{x_1, x_2, \dots, x_n, x_{new1}, x_{new2}, \dots, x_{newm}\} \quad (3)$$

Once all synthetic data points were generated, the new balanced dataset was used to train the model, providing a more equitable representation of stroke and non-stroke cases. The proposed model performed binary classification of SP rate using multimodal radiomics and demographic features. The objective is not to classify stroke subtypes or predict long-term functional outcomes.

3.2. Feature importance score generation using chaotic map-guided approach

It is important to have FS to improve the model's performance and interpretability. To simplify the identification of the most important features, the CMGA was modeled using chaotic dynamics. Important characteristics, such as age, body mass index (BMI), and glucose levels, could yield different values in children with SP. The logistic map provided in **Equation 4** is used to map each feature F_i (such as age, glucose level, and hypertension status) to a chaotic sequence:

$$C_i(t+1) = r \times C_i(t)(1 - C_i(t)) \quad (4)$$

where,

- r is the chaotic parameter.
 - $C_i(t)$ is the chaotic sequence for feature F_i at time t .
- This modeled how a feature's values evolved, capturing its inherent unpredictability. As features evolved, their importance was determined by how substantially they varied during model training. The dynamic behavior of each feature at different time steps is given by **Equation 5**:

$$C_i(t+k) = r \times C_i(t) \times (1 - C_i(t)) \quad \text{for } k \in [1, T] \quad (5)$$

where,

- T is the number of training iterations, indicating the temporal development of feature dynamics.

The variation in chaotic behavior for every feature as a total variation or entropy over the learning period was quantified, as specified in **Equation 6**:

$$I_i = \sum_{t=0}^T |C_i(t+1) - C_i(t)| \quad (6)$$

where,

- I_i is the importance score of features as F_i .
- Features with higher I_i values were deemed more critical for SPM. The proposed explainability model computed feature importance based on chaotic feature dynamics throughout training. Features with higher CMGA scores contribute more remarkably to the final prediction, enabling clinicians to detect key risk factors such as age, hypertension, glucose levels, and BMI. Although SHapley Additive exPlanations (SHAP), Grad-class activation mapping (CAM), and local interpretable model-agnostic explanations (LIME) were not incorporated in the present work, they should be investigated in future studies to provide patient-level explanations and improve clinical interpretability.

The features were ranked based on their importance score I_i . The most relevant features (hypertension/glucose level) were selected for training, as shown in **Equation 7**:

$$\text{Feature ranking} = \{F_1, F_2, \dots, F_n \text{ sorted by } I_i\} \quad (7)$$

Only the standard and most relevant features would contribute to modeling with this FS, thereby increasing interpretability and prediction accuracy.

3.3. Classification using butterfly-based spiking neural network

The BB-SNN is developed and implemented in SNN to train the SPM's weight parameters. By improving classification performance and reducing computational complexity in SNN-based SP tasks, this mechanism enhances accuracy in classifying subjects as either positive or negative for stroke. BB-SNN was selected because SNNs provide event-driven computation with lower computational complexity than conventional deep neural networks. Butterfly optimization efficiently optimizes synaptic weights while balancing exploration and exploitation, leading to faster convergence, lower computational cost, and improved prediction performance compared to conventional CNN-based optimization methods. The optimization ensures practical training and improved performance on involved and imbalanced data, such as in SP. Membrane potential dynamics govern the behavior of neurons in an SNN. The membrane potential $V_i(t)$ and the i^{th} neuron evolved according to the following differential **Equation 8**:

$$\tau \frac{dV_i}{dt} = -V_i + \sum_{j=1}^n W_{ij} S_j(t) \quad (8)$$

where,

- V_i is the membrane probability of the i^{th} neuron at time t .
- τ is the membrane time constant.
- W_{ij} is the synaptic weight from neuron j to neuron i .
- $S_j(t)$ is the spike output of neuron j at time t .

When the membrane reached a specific level, it triggered a spike and then reset, as indicated in **Equation 9**:

$$V_i(t) = V_{\text{reset}}, \quad \text{if } V_i(t) \geq V_{\text{threshold}} \quad (9)$$

where,

- V_{reset} is the reset probability after a spike is fired.
- $V_{\text{threshold}}$ is the threshold value to generate a spike.

The BOA mimicked butterflies' search patterns to adjust SNN weights. The position of a butterfly X_k in the optimization space was updated according to **Equation 10**:

$$X_{k+1} = X_k + \alpha \cdot \sin(\beta \times D) \times \exp(-\gamma \times D) \quad (10)$$

where,

- X_k is the current weight vector of the k^{th} butterfly.
- α is the scaling factor controlling exploration size.
- β is the search direction parameter.
- D is the Euclidean distance between the current position and the optimal position.
- γ is the convergence factor controlling how the exploration decays over time.

The butterfly's movement combined global search and local refinement, ensuring exploration and exploitation of the solution space. The distance between the current butterfly position " X_k " and the best-found position " X_{best} " is computed in **Equation 11**:

$$D = \sqrt{\sum_{i=1}^n (X_{k,i} - X_{\text{best},i})^2} \quad (11)$$

where,

- D is the Euclidean distance.
- $X_{k,i}$, $X_{\text{best},i}$ are the positions of the k^{th} and best-found butterfly in the i^{th} dimension.

The weight update rule in BOA is given by **Equation 12**:

$$W_{ij}(t+1) = W_{ij}(t) + \delta W_{ij} \quad (12)$$

where,

- W_{ij} is the synaptic weight at time t .
- δW_{ij} is the weight change for the connection from neuron j to neuron i .

The weight change δW_{ij} was calculated based on the butterfly's position update as a function of the distance and search parameters. A local search was performed around the best-found butterfly to refine the solution. The updated position X_k is given in **Equation 13**:

$$X_k = X_{\text{best}} + \rho \times (X_{\text{best}} - X_{\text{current}}) \quad (13)$$

where,

- X_{best} is the best solution found.
- X_{current} is the current butterfly's position.
- ρ is a local search refinement parameter.

The local search helped refine the weight adjustments, which in turn improved the network's performance by reducing errors and accelerating convergence. After the weights were optimized, the neurons fired, and the classification output was read. The description of each neuron's spike response is given in **Equation 14**:

$$Y_i = \sum_{j=1}^n W_{ij} \times X_j \quad (14)$$

where,

- Y_i is the output signal of neuron i .
- W_{ij} is the optimized synaptic weight between neurons i and j .
- X_j is the input feature associated with neuron j .

For SP, if Y_i exceeded a certain threshold, the neuron classified the input as indicative of a stroke, as assumed in **Equation 15**:

$$Y_{\text{pred}} = \begin{cases} 1 & \text{if } Y_i \geq \text{threshold} \\ 0 & \text{if } Y_i < \text{threshold} \end{cases} \quad (15)$$

In **Equation 16**, the true label Y_{true} is assumed, and the prediction error is minimized by calculating the error E between the predicted output and the actual output Y_{pred} :

$$E = \frac{1}{2} \sum_{i=1}^n (Y_{\text{pred}} - Y_{\text{true}})^2 \quad (16)$$

where,

- E is the error.
- Y_{pred} is the predicted output of the network.
- Y_{true} is the true label indicating whether a stroke occurred.

The weights were updated based on the error using gradient descent, or a similar BOA is assumed in **Equation 17**:

$$W_{ij}(t+1) = W_{ij}(t) - \eta \times \frac{\partial E}{\partial W_{ij}} \quad (17)$$

where,

- η is the learning rate.
- $\frac{\partial E}{\partial W_{ij}}$ is the partial derivative of the error with respect to the weight W_{ij} .

The iterative optimization method was run until convergence was reached. The error E was used to monitor convergence. For convergence to succeed, the error should be below a predetermined threshold. The stopping condition for the optimization process is specified in **Equation 18**:

$$\text{Convergence condition} : E(t) - E(t-1) < \epsilon \quad (18)$$

where,

- $E(t)$ is the error at iteration t .
- ϵ is a small threshold to indicate convergence.

If the condition holds, the optimization stops, and the

final weight configuration is used for classification. Once the network converges, the final classification decision is based on the optimized weights. The output Y_{final} is the final SP, as specified in **Equation 19**:

$$Y_{final} = \sum_{i=1}^n W_{ij} \times X_j \quad (19)$$

where,

- Y_{final} is the final decision, signifying the probability of stroke.
- W_{ij} is the optimized synaptic weights.
- X_j is the input features.

If Y_{final} exceeded a defined threshold, the individual was classified as having experienced a stroke; otherwise, they were classified as not having experienced a stroke. The entire SP is assumed in [Algorithm 1](#).

Algorithm 1. for SP with SMOTE, CMGA-based FS, and BB-SNN

Input:

- Imbalanced Dataset D
- Chaotic parameter 'r',
- k-nearest neighbor, SNN Hyperparameters
- Learning Rate 'η', epochs

Output: Final SP with optimized model weights

1. Data preprocessing with SMOTE:

- Separate minority class instances ($D_{minority}$) from the dataset.
- For each** instance x_i in $D_{minority}$:
 - Compute Euclidean distances using **Equation 20**:

$$Distance(x_i, x_j) = \sqrt{\sum_{p=1}^m (f_{ip} - f_{jp})^2} \text{ for } p = 1 \text{ to } m \quad (20)$$

- Find k-nearest neighbor and store them in Neighbors (x_i)

- For each neighbor x_j of x_i :

- Generate 'λ' randomly in [0, 1]
- Generate a synthetic sample using **Equation 21**:

$$x_{new} = x_i + \lambda(x_j - x_i) \quad (21)$$

- Append x_{new} to the dataset

- Repeat until the class distribution is balanced
- Store the new balanced dataset, $D_{balanced}$, using **Equation 22**:

$$New\ Dataset = \{x_1, x_2, \dots, x_n, x_{new1}, x_{new2}, \dots, x_{newm}\} \quad (22)$$

2. Feature Importance Score Generation using CMGA:

- Initialize chaotic sequences $C_i(0)$ for each feature F_i
- For** t = 0 to T:
 - For each** feature F_i : Update chaotic sequence using a logistic map (**Equation 23**):

$$C_i(t+1) = r \cdot C_i(t) \cdot (1 - C_i(t)) \quad (23)$$

- Compute dynamic behavior using **Equation 24**:

$$C_i(t+k) = r \cdot C_i(t) \cdot (1 - C_i(t)) \quad \text{for } k \in [1, T] \quad (24)$$

d. Compute variation for each feature using **Equation 25**:

$$I_i = \sum_{t=0}^T |C_i(t+1) - C_i(t)| \quad \text{for } t = 0 \text{ to } T \quad (25)$$

e. Rank features based on I_i in descending order

f. Select top-ranked features using **Equation 26**:

$$\text{Feature Ranking} = \{F_1, F_2, \dots, F_n \text{ sorted by } I_i\} \quad (26)$$

3. Initialize SNN:

- Define SNN with membrane potentials $V_i(0)$ for all neurons
- Randomly initialize synaptic weights W_{ij}
- Set parameters for butterfly optimization: $\alpha, \beta, \gamma, \rho$

4. Training with BB-SNN:

For Epoch = 1 to epochs **Do**:

For each input instance **X**:

i. Compute membrane potential dynamics using **Equation 27**:

$$\tau \frac{dV_i}{dt} = -V_i + \sum_{j=1}^n W_{ij} S_j(t) \quad (27)$$

ii. If $V_i \geq V_{\text{threshold}}$ generate a spike and reset probability using **Equation 28**:

$$V_i(t) = V_{\text{reset}} \quad (28)$$

b. Compute the output of neurons using **Equation 29**:

$$Y_i = \sum_{j=1}^n W_{ij} \cdot X_j \quad (29)$$

c. Compute classification error using **Equation 30**:

$$Y_{\text{pred}} = \begin{cases} 1 & \text{if } Y_i \geq \text{threshold} \\ 0 & \text{if } Y_i < \text{threshold} \end{cases} \quad (30)$$

5. Butterfly-based weight update:

a. Compute Euclidean distance using **Equation 31**:

$$D = \sqrt{\sum_{i=1}^n (X_{k,i} - X_{\text{best},i})^2} \quad (31)$$

b. Update butterfly position using **Equation 32**:

$$X_{k+1} = X_k + \alpha \cdot \sin(\beta \cdot D) \cdot \exp(-\gamma \cdot D) \quad (32)$$

c. Perform local search refinement using **Equation 33**:

$$X_k = X_{\text{best}} + \rho \cdot (X_{\text{best}} - X_{\text{current}}) \quad (33)$$

d. Update weights using **Equation 34**:

$$W_{ij}(t+1) = W_{ij}(t) + \delta W_{ij} \quad (34)$$

e. Compute weight change using **Equation 35**:

$$W_{ij}(t+1) = W_{ij}(t) - \eta \times \frac{\partial E}{\partial W_{ij}} \quad (35)$$

6. Convergence check:

If $E(t) - E(t-1) < \varepsilon$, stop training

7. Final stroke classification:

a. Compute final neuron output using **Equation 36**:

$$Y_{final} = \sum_{i=1}^n W_{ij} \times X_j \quad (36)$$

b. If $Y_{final} \geq threshold$, classify as stroke

c. Otherwise, classify as non-stroke

8. Return: Optimized BB-SNN weights and SP output.

4. Results and discussion

The dataset used in this study was compiled by McKinsey & Company and included EHRs from people residing in Bangladesh. The data were in CSV format. The initial dataset consisted of 29,072 records, and a sample of 5,110 observations was selected for SP analysis. The 10 predictors included were age, gender, hypertension, marital status, work environment, residential environment, heart disease, average blood glucose level, BMI, and smoking history. Age and average glucose level are numerical variables, whereas gender, marital status, work setting, and residence setting are nominal variables. Some of the binary attributes are hypertension, heart disease, and smoking history. The presence of various health-related features helps develop more effective ML models, given the complex nature of SP, and provides more accurate and valuable data from healthcare datasets. The proposed model should be run on a computer with an Intel Core i7 processor or better, 16 GB of RAM, and an NVIDIA GPU for computational purposes. The software requirements are Python 3.8 or higher, TensorFlow, Keras, SciKit-Learn, and Pandas. Moreover, it is also required to use Jupyter Notebook or another development environment for model implementation and testing.

This study used data from a publicly available EHR dataset compiled by McKinsey & Company, which comprised health data on people living in Bangladesh. The original dataset contained 29,072 records; after

preprocessing and quality control, the number of patient records was reduced to 5,110 for model development. The proposed model is a binary classification model for the SP system that classifies patients as having a stroke or not. The study population comprised records in which all demographic and clinical variables, including age, gender, hypertension, heart disease, average blood glucose level, BMI, marital status, work type, residence type, and smoking history, were complete. To ensure data quality, records with missing, incomplete, inconsistent, or duplicate data were excluded. The publicly available dataset does not include specific details on the number of stroke and non-stroke patients, mean age, gender distribution, stroke type (ischemic or hemorrhagic), or standardized imaging attainment trials. As such, the proposed model was built using the available multimodal radiomics and clinical features, without any subtype-specific analysis. Also, external validation was not conducted because only one public dataset was used in this study. Future work should include validating the proposed model on large-scale, multicenter datasets that include patients from diverse backgrounds and imaging protocols to increase the model's generalizability and clinical applicability.

As shown in [Figure 2](#), the initial data distribution was compared to the distribution obtained from random sampling. The visualization shows that the data are skewed and that random sampling can be used to balance the data for ML training.

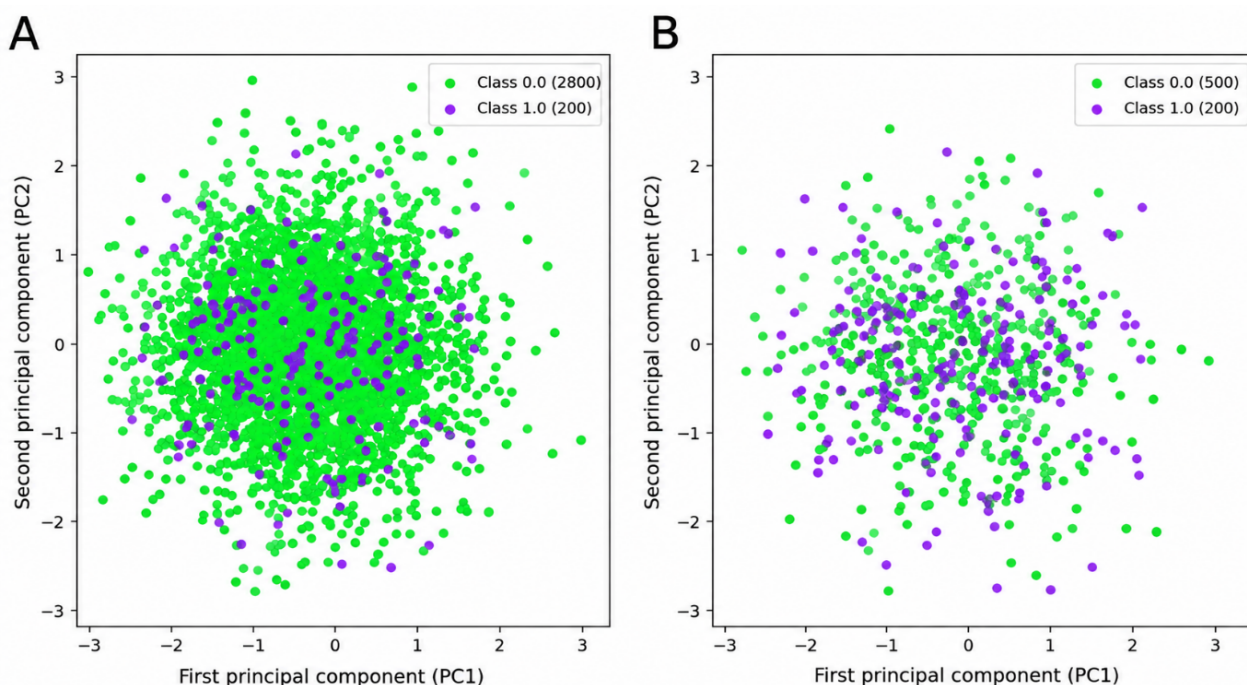


Figure 2. Original (A) and random sampling (B) of data

Figure 3 compares various under-sampling methods. Under-sampling reduces the number of cases in the majority class to equalize class ratios. The diagram shows how these methods help reduce class imbalance in dynamic SP.

Figure 4 shows the heat map of correlations among the dataset's features. It is useful for detecting strong positive or negative relationships between features, thereby helping identify feature requirements and select the most appropriate features for the model.

Figure 5 shows the distribution of the main variables for the samples, specifically regarding the presence or absence of stroke. It represents trends and variations in age, BMI, glucose level, and smoking history. Data from these distributions help interpret the relationship between the features and stroke rate and formulate a better SPM.

Figure 6 displays the FS scores attained based on the chaotic map, displaying the importance of the features in SP. The CMGA improved FS efficiency by detecting the most remarkable features. This improvement makes the model more robust and interpretable, reducing the computational cost while achieving high accuracy.

Figure 7 depicts numerous ML and DL models used, including CNN, ResNet50, Inception v3, and DT, which

were compared to the proposed BB-SNN as shown in Figure 6. The higher AUC value of the BB-SNN also indicates that it has better classification ability in SP compared to other models. The visual comparison shows the improvement in accuracy and sensitivity achieved by the proposed model.

Table 3 and Figure 8 compare the error rates across different epochs for the CNN, ResNet50, Inception v3, DT, and the proposed BB-SNN. As the number of epochs increases, the error rates of conventional models improved only slightly, while CNN, ResNet50, and Inception v3 exhibited higher error rates than BB-SNN at later epochs. DT was initially superior to other models in terms of performance, but it was not scalable with epochs. On the other hand, the proposed BB-SNN exhibited a significantly lower error rate, ranging from 1.9% at 100 epochs to 5.2% at 500 epochs. This trend confirms not only the fast computational speed of the BB-SNN but also its learning convergence, which is due to butterfly-based weight optimization and FS based on a chaotic map. The error rates of all models for each epoch are also compared in Figure 8, which shows that BB-SNN was more stable and exhibited lower error rates compared to the other models. The fact that the model has lower error rates with an increased number of epochs proves the efficiency and learning ability of the model.

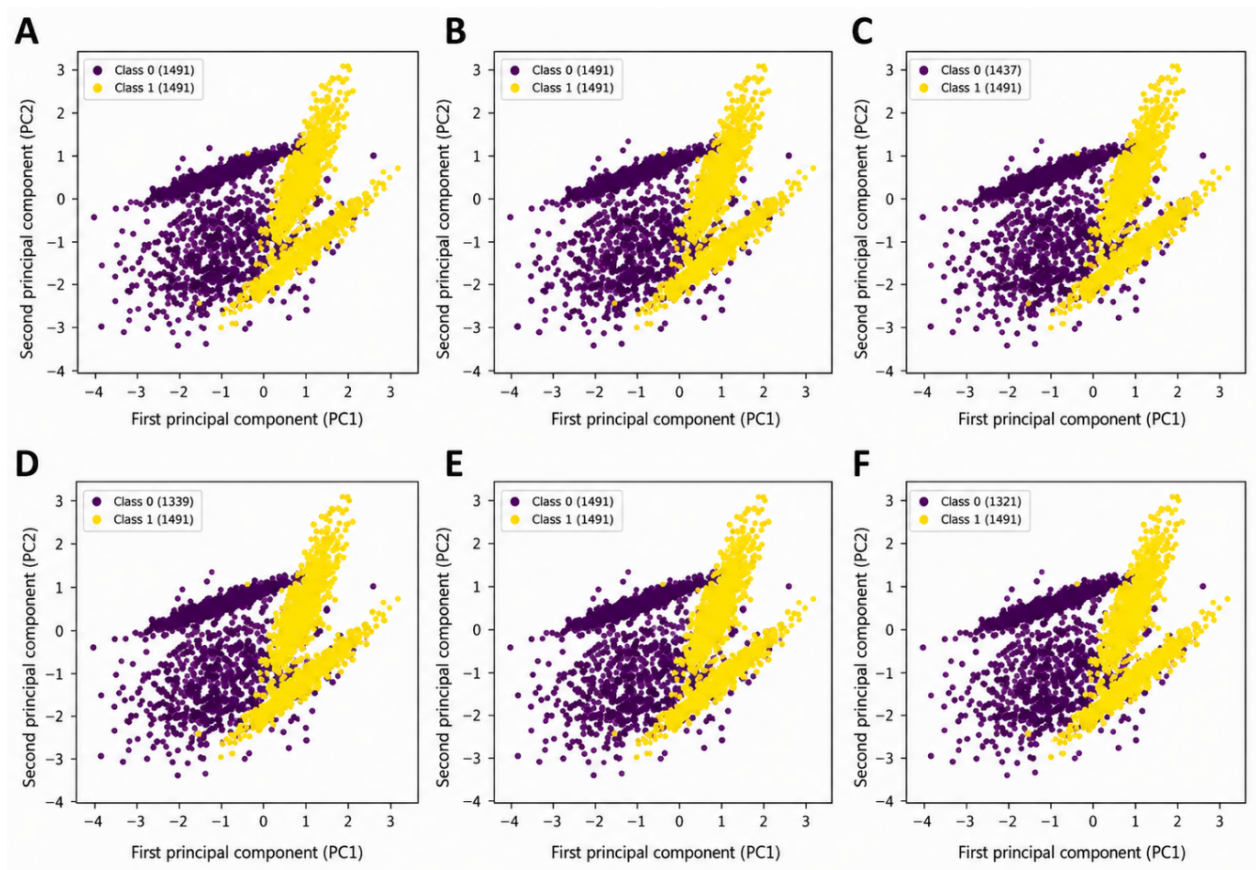


Figure 3. Under-sampled data. (A) Random undersampling. (B) Cluster centroid undersampling. (C) Tomek Links undersampling. (D) Neighborhood cleaning rule undersampling. (E) Near miss undersampling. (F) Edited nearest neighbor undersampling.

Table 3. Comparison of error rate

Epoch	CNN	ResNet50	Inception v3	DT	BB-SNN
100	10.67	11.12	11.45	9.887	1.9
200	11.78	11.56	11.67	10.560	2.3
300	12.87	11.99	12.34	10.830	3.6
400	12.90	12.01	12.77	11.872	4.7
500	13.23	12.30	13.76	11.986	5.2

Abbreviations: BB-SNN: Butterfly-based spiking neural network; CNN: Convolutional neural network; DT: Decision tree.

Figure 9 shows the time taken by several classifiers over 500 epochs. CNN, ResNet50, Inception v3, and DT were more time-consuming, and the values varied, while BB-SNN remained low. The wave-like patterns indicate performance variations and show the efficiency of BB-SNN in terms of computational cost as compared to DL classifiers. In the case of SP, the performance measures, including accuracy, sensitivity, specificity, *F1*-score, and

MCC, were the most important measures of the model's ability to predict and classify stroke cases vs. non-stroke cases accurately. Accuracy is a measure of the model's overall correctness in predicting stroke and non-stroke cases. It is the percentage of correctly classified stroke and non-stroke cases relative to the total number of cases. Sensitivity (or recall) measures the model's ability to predict TP stroke cases accurately. High sensitivity means the model can accurately identify most stroke cases. While sensitivity evaluates the model's capacity to predict stroke cases accurately, specificity determines the model's ability to predict non-stroke cases, thereby reducing the number of false positives (FPs). The *F1*-score combines sensitivity and precision, calculated as a harmonic mean, and is particularly useful for imbalanced datasets, such as stroke data.

The MCC measures the model's accuracy by considering the true positive (TP), FP, true negative (TN), and false negative (FN) rates. It is an accurate measure that provides

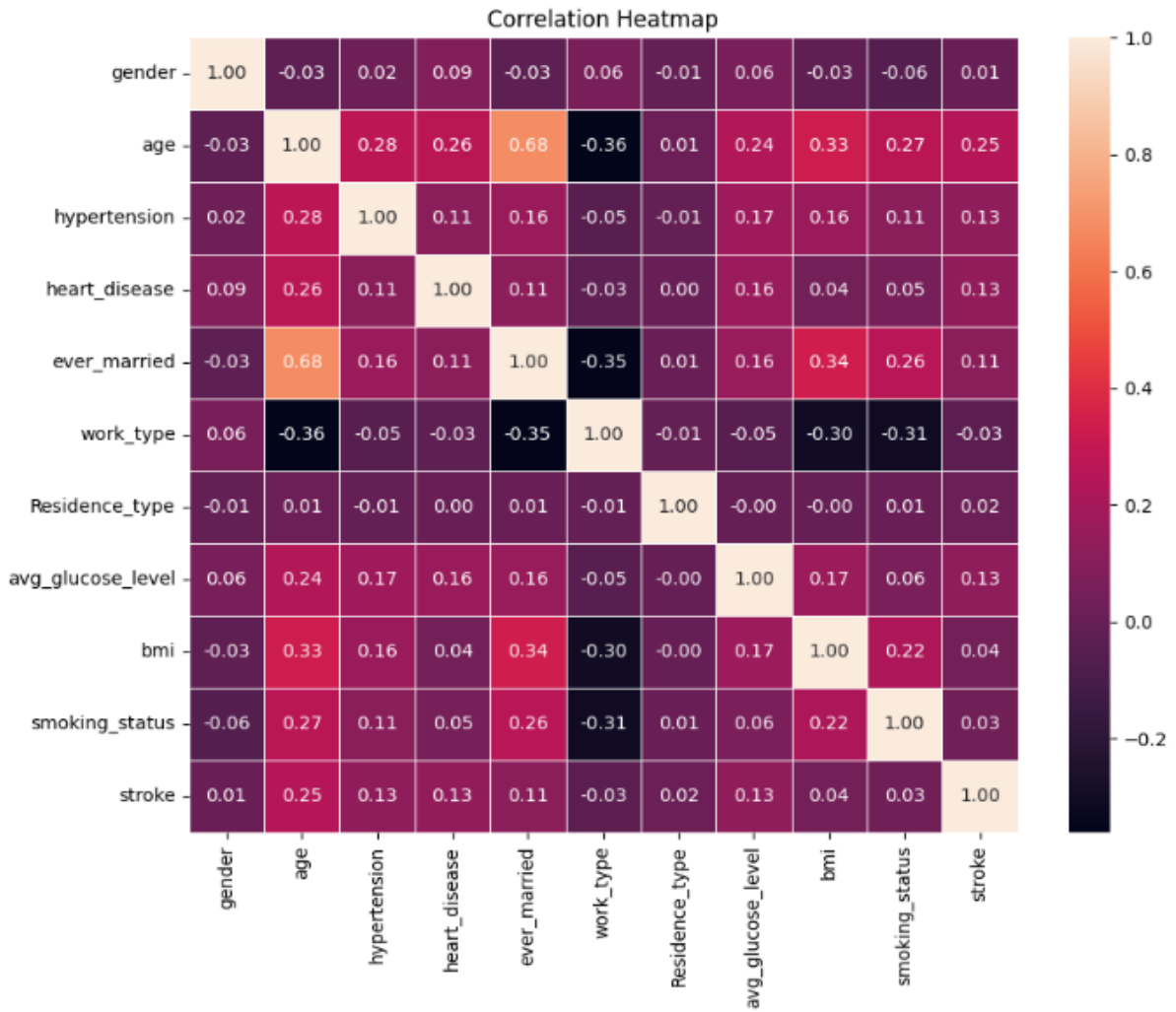


Figure 4. Correlation heat map of features

a more reliable indicator than other metrics when the dataset is skewed. The performance metrics are defined in **Equations 37–41**:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (37)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (38)$$

$$Specificity = \frac{TN}{TN + FP} \quad (39)$$

$$F1 \text{ Measure} = 2 \times \frac{Precision \times Sensitivity}{Precision + Sensitivity} \quad (40)$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (41)$$

These are useful in assessing the use of the SPM and are particularly relevant in medical fields where classification is significant for patient management. The performance is presented in [Table 4](#) and [Figure 10](#).

[Table 3](#) presents the model evaluation metrics: accuracy, sensitivity, specificity, *F1*, and MCC. CNN, ResNet-50, and Inception v3 achieved nearly equal performance, with accuracy higher than 85%. However, BB-SNN achieved the highest accuracy (94.80%), sensitivity (93.50%), and specificity (95.30%) among all models. It also achieved higher *F1*-scores and MCC values, signifying the model's efficiency in SP. The better results are attributed to more

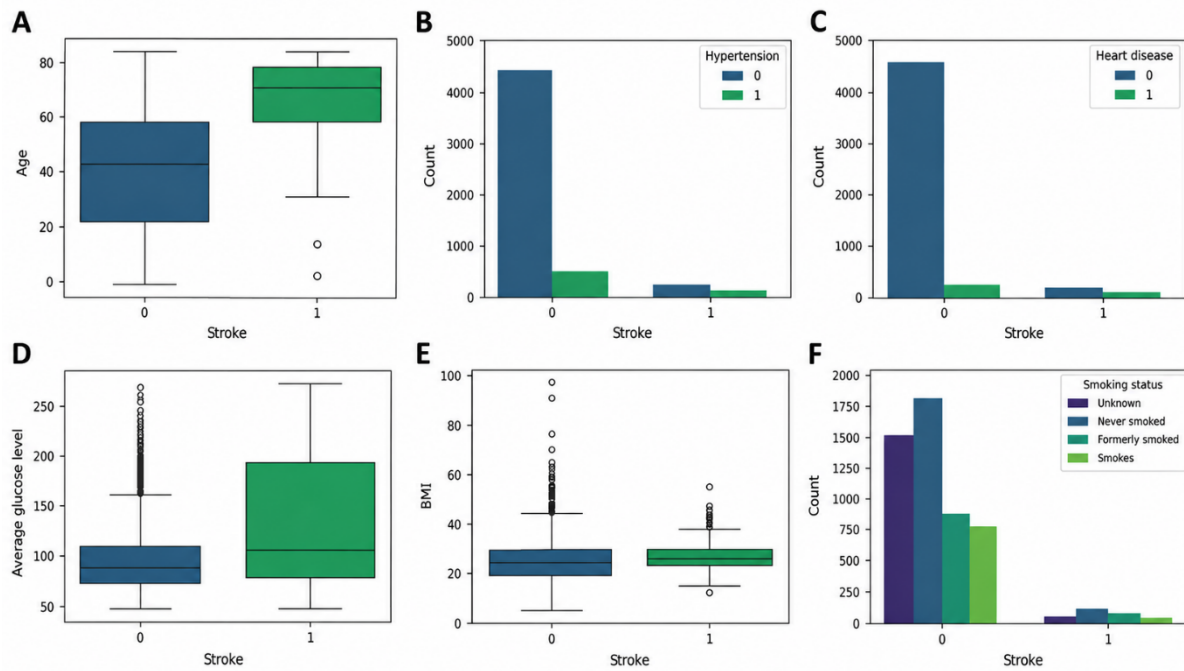


Figure 5. Distribution of diversified features across the presence and absence of stroke. (A) Age comparison for stroke. (B) Hypertension comparison for stroke. (C) Heart disease comparison with stroke. (D) Average glucose level comparison for stroke. (E) Body mass index (BMI) comparison for stroke. (F) Smoking status comparison for stroke.

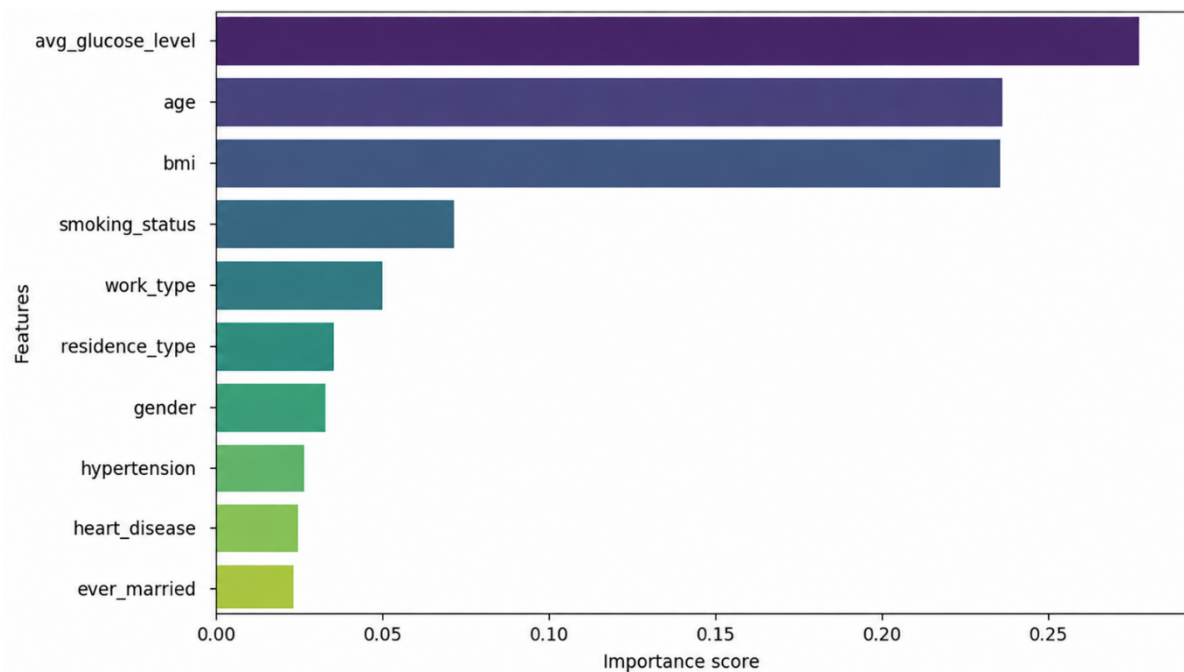


Figure 6. Feature importance for chaotic map-guided approach

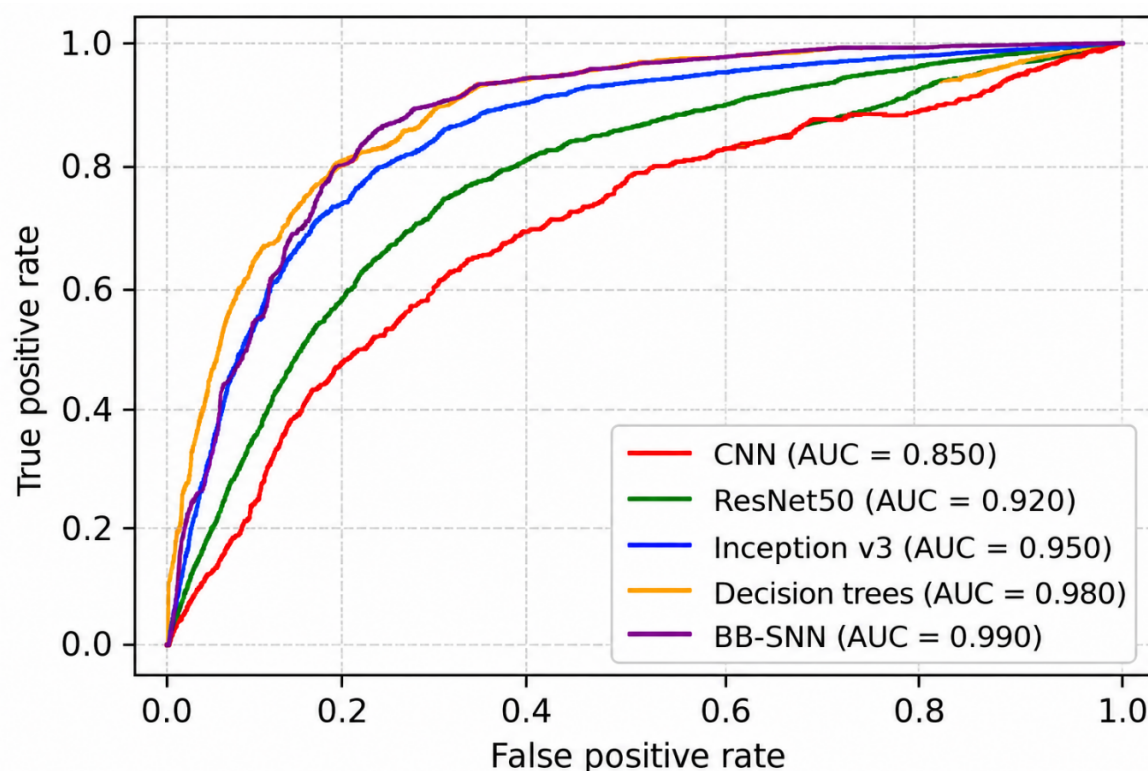


Figure 7. Comparison of receiver operating characteristic curves

Abbreviations: AUC: Area under the curve; BB-SNN: Butterfly-based spiking neural network; CNN: Convolutional neural network.

effective feature and butterfly-based weight optimization. While less computationally intensive in terms of time and resources, DT exhibited relatively lower accuracy, a crucial factor signifying its inefficiency in handling medical data. Figure 10 shows the classification metrics for all the models. BB-SNN outperforms all other methods and is a model for SP, surpassing CNNs and ResNet50.

The proposed model shows promising predictive performance, but a prospective multicenter validation is required before the model can be implemented in routine clinical activities. Practical issues that need to be addressed include integration with hospital information systems, variations in imaging protocols, bias introduced by automation, and clinician acceptance. Therefore, the proposed model should be considered as a CDSS tool rather than an autonomous diagnostic system at the moment. AI is not limited to SP. Predictive models powered by AI can assist in rehabilitation planning by estimating recovery trajectories, supporting individualized therapy, and monitoring functional improvement.

Table 4. Comparison of classification performance

Performance metrics	CNN	ResNet50	Inception v3	DT	BB-SNN
Accuracy	85.20	88.50	87.10	78.30	94.80
Sensitivity	84.00	87.00	86.50	77.00	93.50
Specificity	86.00	89.20	87.80	79.50	95.30
<i>F1</i> -score	84.60	88.00	87.00	78.10	94.10
MCC	82.10	86.50	85.40	76.00	93.20

Abbreviations: BB-SNN: Butterfly-based spiking neural network; CNN: Convolutional neural network; DT: Decision tree; MCC: Matthews correlation coefficient.

Although the proposed framework demonstrates promising predictive performance, its clinical implementation requires prospective multicenter validation before routine deployment. Practical challenges include integration with hospital information systems, differences in imaging protocols, automation bias, and clinician acceptance. Therefore, the proposed model should currently be regarded as a decision-support tool rather than an autonomous diagnostic system.

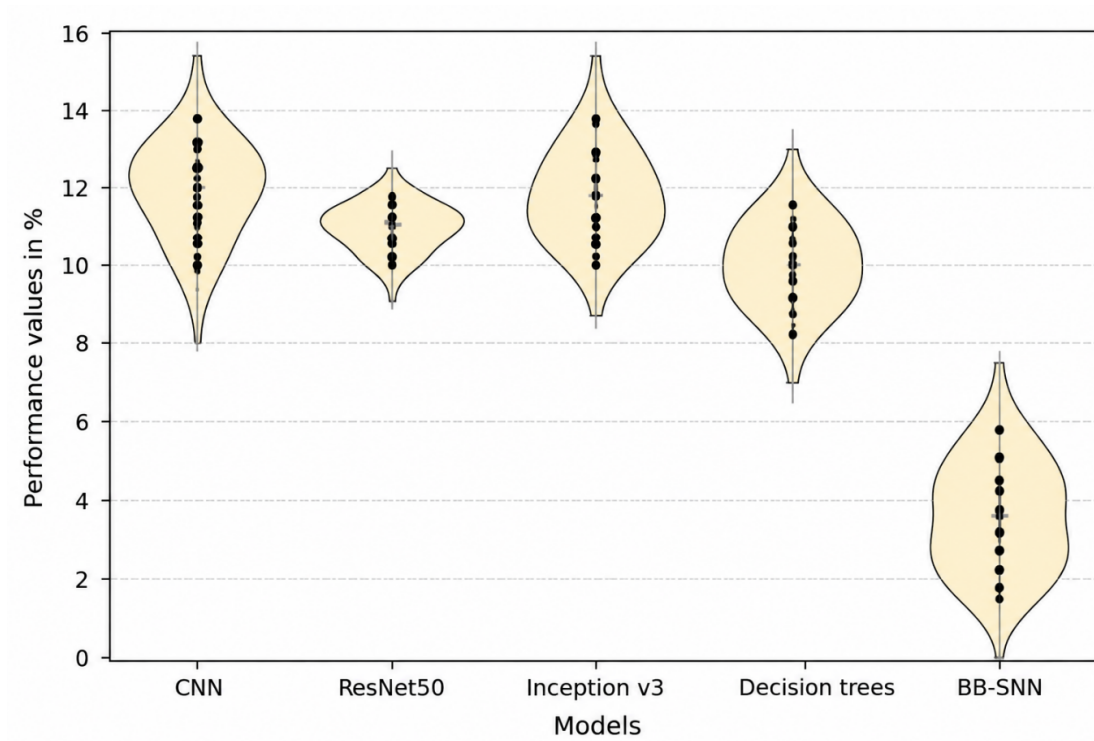


Figure 8. Comparison of error rate
Abbreviations: BB-SNN: Butterfly-based spiking neural network; CNN: Convolutional neural network; DT: Decision tree.

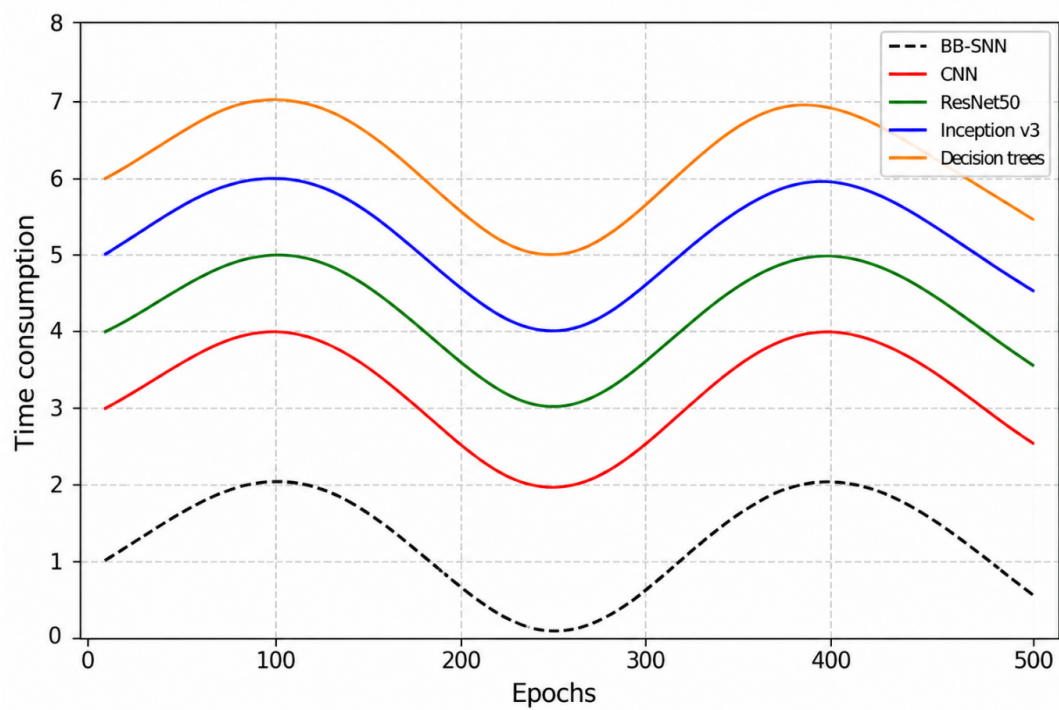


Figure 9. Comparison of the time consumption of classifiers.
Abbreviations: BB-SNN: Butterfly-based spiking neural network; CNN: Convolutional neural network.

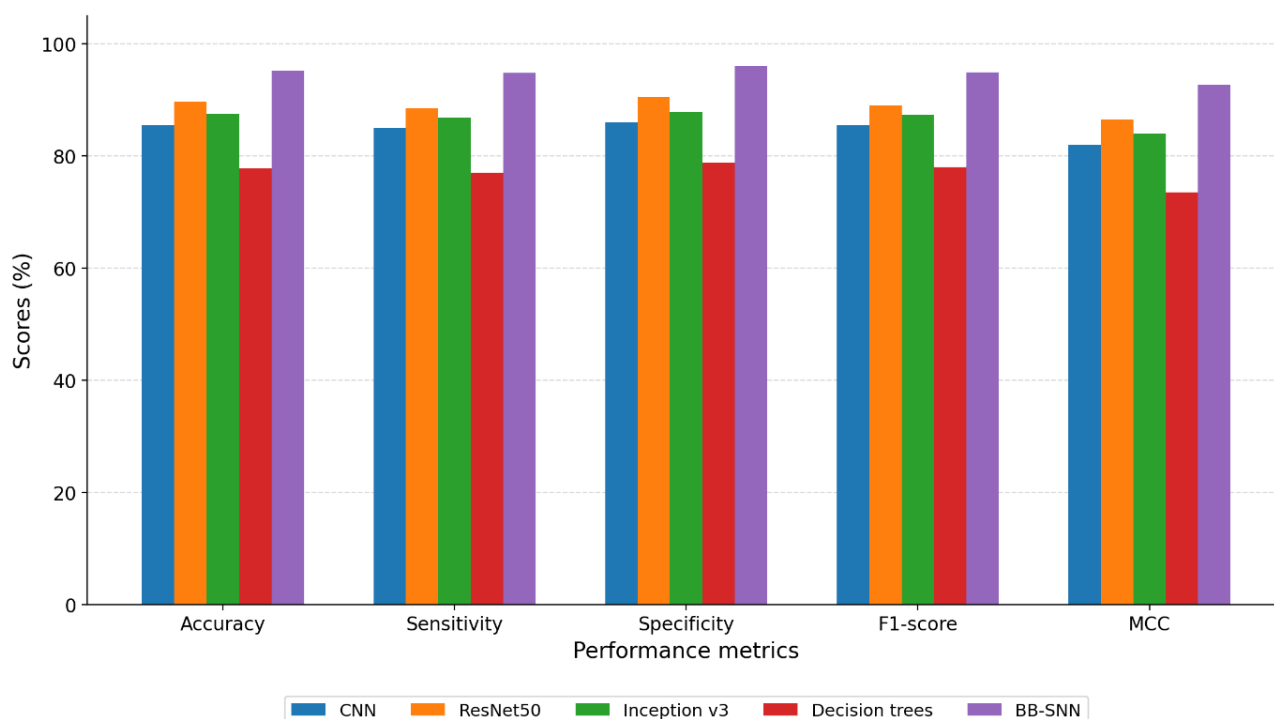


Figure 10. Comparison of classification performance. Abbreviation: MCC: Matthew correlation coefficient.

Furthermore, population-level screening models may enable early identification of individuals at elevated stroke risk before symptom onset, thereby helping preventive healthcare strategies. These developments further reinforce the role of AI within precision medicine. The proposed model was evaluated using a publicly available retrospective dataset without external multicenter validation. The model predicts stroke rate but does not differentiate stroke subtypes or estimate long-term functional outcomes.

These findings are consistent with earlier epidemiological data from Catalonia, which similarly reported divergent incidence and case-fatality rates between ischemic and hemorrhagic stroke subtypes⁴⁶, reinforcing the need for subtype-aware prediction models. Furthermore, Gasull and Arboix⁴⁷ highlighted that molecular and biomarker profiles differ meaningfully across stroke subtypes, supporting our rationale for incorporating laboratory and multi-dimensional features—rather than imaging alone—into the proposed stratification framework. Although CMGA provides feature-level explainability, patient-specific interpretation using SHAP or LIME was not included. Future studies should validate the model using larger, prospective, multicenter cohorts.

5. Conclusion

The proposed SPM addresses some limitations of existing models by incorporating MMR and improved preprocessing for class imbalance via SMOTE. The FS is completed more efficiently with CMGA-based FS scoring, thereby increasing the model's interpretability and reliability. The proposed BB-SNN reduces computational cost and improves classification rates for complex medical diagnostic problems, thereby enhancing the use of SNNs. This training process is faster and more effective, making the system a computationally efficient solution for SP. The comparative analysis confirms that it is more efficient at handling radiomics data and addressing class imbalance than traditional models. This development places the model in a good position for implementation in clinical practice settings. In addition to accuracy, the proposed method includes early SP, which has the potential not only to support early treatment but also to improve the management of stroke risks in the health sector. Extending the proposed model for real-time SP by applying continuous data flow should be the primary emphasis of future work. More state-of-the-art explainability methods should be employed to explain BB-SNN's results. This may improve the generalizability of the results by incorporating patient-specific factors and datasets from multiple centers into the analysis. Moreover, the use of federated learning can improve data privacy and security while enabling

multi-site training across hospitals, thereby increasing the system's scalability and practicality.

Future work should explore prospective multicenter validation, subtype-specific SP, integration of genomic and wearable sensor data, federated learning for privacy-preserving distributed training, explainable AI methods such as SHAP and Grad-CAM, and deployment within CDSS for real-time SP risk assessment.

Acknowledgments

None.

Funding

None.

Conflict of interest

The authors declare they have no competing interests.

Author contributions

Conceptualization: Sudhakar Sengan

Data curation: Vivekanandhan Vijayarangan, Aseel Smerat

Formal analysis: Sudhakar Sengan

Investigation: Sudhakar Sengan

Methodology: Sudhakar Sengan

Resources: Venu Karunanithi, Vidya Sagar Ponnamp

Software: Hayder M. A. Ghanimi, Vedaraj Muthuraj

Validation: Dharani Kumar Sunkara Venkata

Visualization: Aseel Smerat

Writing—original draft: Sudhakar Sengan

Writing—review & editing: Sudhakar Sengan

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

The dataset analyzed in this study is publicly available from Kaggle at <https://www.kaggle.com/datasets/fedesoriano/stroke-prediction-dataset>

Availability of code

The source code developed for this study is publicly available through GitHub and archived on Zenodo to facilitate reproducibility and future research. The complete implementation, including the preprocessing pipeline, feature selection module, model training framework, evaluation scripts, and supporting documentation, can be accessed from the GitHub repository: <https://github.com/sudhakarsengan/Deep-Learning-Framework-for-Precision-Stroke-Prognosis>.

A permanent archived version of the repository is available through Zenodo: <https://zenodo.org/records/21496687>. The archived version provides a citable snapshot of the source code and associated research materials, ensuring long-term accessibility and reproducibility. The repository contains the implementation of the proposed framework, installation instructions, dependency specifications, and documentation required to reproduce the reported experiments.

References

1. Huang S, Diao S, Wan Y. Application of machine learning methods in predicting functional recovery in ischemic stroke patients. In: Innovative Scientific Research: Theory, Methodology, Practice: Proceedings of the 1st International Scientific and Practical Conference; September 3-6, 2024; Boston, MA. International Science Group; 2024:240-252. doi: 10.46299/ISG.2024.2.1
2. Barron K, Blaivas M, Blaivas L, Sadler J, Deal I. Bedside ultrasound to identify and predict severity of dysphagia following ischemic stroke: human versus artificial intelligence. *Ultrasound Med Biol*. 2024;50(1):99-104. doi: 10.1016/j.ultrasmedbio.2023.09.008
3. Chen M, Qian D, Wang Y, *et al*. Systematic review of machine learning applied to the secondary prevention of ischemic stroke. *J Med Syst*. 2024;48(1):8. doi: 10.1007/s10916-023-02020-4
4. GBD 2021 Stroke Risk Factor Collaborators. Global, regional, and national burden of stroke and its risk factors, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet Neurol*. 2024;23(10):973-1003. doi: 10.1016/S1474-4422(24)00369-7
5. Adams HP Jr, Bendixen BH, Kappelle LJ, *et al*. 3rd. Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. *Stroke*. 1993;24(1):35-41. doi: 10.1161/01.STR.24.1.35
6. Thorvaldsen P, Asplund K, Kuulasmaa K, Rajakangas AM, Schroll M. Stroke incidence, case fatality, and mortality in the WHO MONICA Project. *Stroke*. 1995;26(3):361-367. doi: 10.1161/01.STR.26.3.361
7. Kongsut S, Piriyaakuntorn P. Unlocking the potential of HB/RDW ratio as a simple marker for predicting mortality in acute ischemic stroke patients after thrombolysis. *J Stroke Cerebrovasc Dis*. 2024;33(9):107874. doi: 10.1016/j.jstrokecerebrovasdis.2024.107874
8. Yi JS, Ki HJ, Jeon YS, *et al*. The collateral map: prediction of

- lesion growth and penumbra after acute anterior circulation ischemic stroke. *Eur Radiol.* 2024;34(3):1411-1421.
doi: 10.1007/s00330-023-10084-6
9. Truong B, Zheng J, Hornsby L, Fox B, Chou C, Qian J. Development and validation of machine learning algorithms to predict 1-year ischemic stroke and bleeding events in patients with atrial fibrillation and cancer. *Cardiovasc Toxicol.* 2024;24(4):365-374.
doi: 10.1007/s12012-024-09843-8
10. Mochetti MM, Silva EGP, Correa AAF, *et al.* Neuron-specific enolase at admission as a predictor for stroke volume, severity and outcome in ischemic stroke patients: a prognostic biomarker review. *Sci Rep.* 2024;14(1):2688.
doi: 10.1038/s41598-024-53080-6
11. Shen H, Huasen BB, Killingsworth MC, Bhaskar SM. Introducing the futile recanalization prediction score (FRPS): a novel approach to predict and mitigate ineffective recanalization after endovascular treatment of acute ischemic stroke. *Neurol Int.* 2024;16(3):605-619.
doi: 10.3390/neurolint16030045
12. Wang Q, Yin J, Xu L, Lu J, Chen J, Chen Y, *et al.* Development and validation of outcome prediction model for reperfusion therapy in acute ischemic stroke using nomogram and machine learning. *Neurol Sci.* 2024;45(7):3255-3266.
doi: 10.1007/s10072-024-07329-7
13. Li N, Li YL, Shao JM, Wang CH, Li SB, Jiang Y. Optimizing early neurological deterioration prediction in acute ischemic stroke patients following intravenous thrombolysis: a LASSO regression model approach. *Front Neurosci.* 2024;18:1390117.
doi: 10.3389/fnins.2024.1390117
14. Wang Y, Zhou Y, Hu H, *et al.* Development and validation of a clinical prediction model for ischemic stroke recurrence after successful stent implantation in symptomatic intracranial atherosclerotic stenosis. *J Clin Neurosci.* 2024;123:137-147. doi:10.1016/j.jocn.2024.03.028
15. Dierksen F, Sommer JK, Tran AT, *et al.* Machine learning models for 3-month outcome prediction using radiomics of intracerebral hemorrhage and perihematomal edema from admission head computed tomography (CT). *Diagnostics.* 2024;14(24):2827.
doi: 10.3390/diagnostics14242827
16. Lu M, Wang Y, Tian J, Feng H. Application of deep learning and radiomics in the prediction of hematoma expansion in intracerebral hemorrhage: a fully automated hybrid approach. *Diagn Interv Radiol.* 2024;30(5):299-312.
doi: 10.4274/dir.2024.222088
17. Wang Z, Shen Y, Zhang X, *et al.* Prognostic value of multi-PLD ASL radiomics in acute ischemic stroke. *Front Neurol.* 2025;15:1544578.
doi: 10.3389/fneur.2024.1544578
18. Guan L, Wei B, Guo D, *et al.* A clinical-radiomics nomogram for predicting early ischemic stroke risk in patients with transient ischemic attack. *World Neurosurg.* 2024;190:e199-e211.
doi: 10.1016/j.wneu.2024.07.090
19. Shi MG, Feng XM, Zhi HY, Hou L, Feng DF. Machine learning-based radiomics in neurodegenerative and cerebrovascular disease. *MedComm.* 2024;5(11):e778.
doi:10.1002/mco2.778
20. Abujaber AA, Imam Y, Balkhi I, Yaseen S, Nashwan AJ, Akhtar N. Utilizing machine learning to facilitate the early diagnosis of posterior circulation stroke. *BMC Neurol.* 2024;24(1):156.
doi: 10.1186/s12883-024-03638-8
21. Abujaber AA, Albalkhi I, Imam Y, Nashwan A, Akhtar N, Alkhawaldeh IM. Machine learning-based prognostication of mortality in stroke patients. *Heliyon.* 2024;10(7):e28869.
doi:10.1016/j.heliyon.2024.e28869
22. Abdul Raheem AK, Dhannoon BN. A novel deep learning model for drug-drug interactions. *Curr Comput Aided Drug Des.* 2024;20(5):666-672.
doi: 10.2174/0115734099265663230926064638
23. Mohammed MA, Abdulkareem KH, Dinar AM, Garcia Zapirain B. Rise of deep learning clinical applications and challenges in omics data: a systematic review. *Diagnostics.* 2023;13(4):664.
doi: 10.3390/diagnostics13040664
24. Mohammed MA, Lakhan A, Abdulkareem KH, Garcia-Zapirain B. A hybrid cancer prediction based on multi-omics data and reinforcement learning state action reward state action (SARSA). *Comput Biol Med.* 2023;154:106617.
doi: 10.1016/j.compbimed.2023.106617
25. Ananthakrishnan G, Ghanimi HMA, Babu KS, *et al.* Mathematical cognitive modelling enhanced by deep learning for understanding human decision making in complex sequential tasks. *JIM.* 2026;29(4):1033-1044.
doi: 10.47974/jim-2573
26. Ghanimi HMA, Akilandeswari K, Hanumat PA, Sengan S, Prakash BP, Bommiseti RK, An Ensemble Cognitive Model for Stroke Prediction Using Unstructured Health Information Powered by Machine Learning. *J. Mach. Comput.* 2025;5(1):611-621.
doi: 10.53759/7669/jmc202505048
27. Eskandari M, Alkhafaji AHA, Al-Asady AM, *et al.* Mebendazole as an Adjunct Therapy with Mesalamine to Increase Efficacy and Maintenance Therapy for Ulcerative Colitis Patients: A Pilot Study. *Curr Pharm Des..*

- 2026;32(8):609-616.
doi: 10.2174/0113816128372513250616182826.
28. Bakhshi A, Akbari M, Maleki F, *et al.* Artificial Intelligence in Non-Alcoholic Fatty Liver Disease and Fibrosis: A Narrative Review. *Curr Pharm Des.* 2026;32.
doi: 10.2174/0113816128422676260219073220
29. Demirci S, Uludüz E, Eyüboğlu S, *et al.* Linking photophobia, sleep disturbances, and migraine chronicity: Evidence from a retrospective analysis. *Brain Behav.* 2026;16(3):e71252.
doi: 10.1002/brb3.71252
30. Rahmani MKI, Ghanimi HMA, Jilani SF, *et al.* Early Pathogen Prediction in Crops Using Nano Biosensors and Neural Network-Based Feature Extraction and Classification. *Big Data Res.* 2023;34:100412.
doi: 10.1016/j.bdr.2023.100412
31. Sai PHVST, Bharathi VVP, Nandan U, *et al.* Detection of Chronic Heart Failure Using Deep Learning and Machine Learning. In: *Smart Innovation, Systems and Technologies.* Springer Nature Singapore; 2025:535-543.
doi: 10.1007/978-981-97-8355-7_46
32. Samypen Y, Chakkaravarthy M. Adversarially-Robust Secure Scheduling Algorithm for Big Data Analytics in Healthcare IoT Networks. In: *2025 International Conference on Intelligent Innovations in Engineering and Technology (ICIET).* IEEE; 2025:1-5.
doi: 10.1109/iciiet65921.2025.11378465
33. Rajaram R, Chakkaravarthy M, Selvam J. Cutting-Edge Lung Cancer Detection with VGG-19. In: *Lecture Notes in Networks and Systems.* Springer Nature Singapore; 2025:205-216.
doi: 10.1007/978-981-96-3102-5_15
34. Midhunchakkaravarthy D, Shirke-Deshmukh S, Deshpande V. NAS-QDCNN: Neuron attention stage-by-stage quantum dilated convolutional neural network for iris recognition at a distance. *Biomed Eng Appl Basis Commun.* 2025;38(01).
doi: 10.4015/s1016237225500218
35. Midhunchakkaravarthy D, RangaNarayana K, Narayana VL. Underwater low-light image enhancement using a hybrid CNN-transformer framework with multi-scale feature fusion and cheetah optimization. *Discov Comput.* 2026;29(1).
doi: 10.1007/s10791-026-09925-x
36. Degadwala S, Midhunchakkaravarthy D, Khan S. Automating Histologic Assessment of Prostate Cancer with a ResNet50-Based Hybrid Vision Model. *TCSST.* 2025;7(3):295-311.
doi: 10.36548/jtcsst.2025.3.001
37. Midhunchakkaravarthy D, Kumar CK, Narayana L. Design of an Improved Model for Liver Transplantation Outcome Prediction Using Graph Embedding and Counterfactual Fusions. In: *2025 9th International Conference on Inventive Systems and Control (ICISC).* IEEE; 2025:374-382.
doi: 10.1109/icisc65841.2025.11188077
38. Midhunchakkaravarthy D, Nagamani GM, Narayana VL. Quantitative benchmarking and cross-modal analysis of deep learning, machine learning, and biosensor frameworks for early colorectal cancer diagnosis and prognosis. *J Theor Appl Inf Technol.* 2025;103(18):7544-7564. <https://www.jatit.org/volumes/Vol103No18/21Vol103No18.pdf>
39. Memon AG, Chandran SP, Nawaz R, Talpur RA, Memon F. Comparative Effects of Neuromuscular Training and Mobilization with Movement on Pain, Range of Motion, Balance, and Function in Footballer with Ankle Sprain. *Malays J Med Health Sci.* 2025;21(5).
doi: 10.47836/mjmhs.21.5.7
40. Memon AG, Chandran SP, Annosha, Zakir T, Sulaman M. Effects of Neuromuscular Training with and Without Kinesio-Tape on Pain, Range of Motion, Balance and Function in Footballer with Ankle Sprain. *Journal of Liaquat University of Medical & Health Sciences,* 2025;24(01):38-44.
doi: 10.22442/jlumhs.2024.01174
41. Rosyad YS, Aljaberi MA, Natarajan SB, Pramesona BA. The Coping Self-efficacy Scale Among People Living With HIV: Validity and Reliability Indonesian Version. *MJMHS.* 2025;21(s3):136-140.
doi: 10.47836/mjmhs.21.s3.21
42. Karmakar M, Panduragan SL, Said FM. A Case Report on Meeting the Spiritual Need of Intubated Conscious Patient. *Patient Exp J.* 2025;12(3):196-200.
doi: 10.35680/2372-0247.2042
43. Supriatin, Dioso RI, Che Hassan H, Maretalinia. Determinants of Caregiver's Knowledge of Infant's Early Nutrition in Indonesia. *MJMHS.* 2025;21(s3):105-110.
doi: 10.47836/mjmhs.21.s3.16
44. Sarvalingam P, Vajravelu A, Bin Abdul Jamil MM, Ambar RB, Selvam J, P R. A Novel Hybrid CNN-GRU Framework with Wavelet-Based Feature Optimization for Automated Autism Spectrum Disorder Detection Using EEG Signals. *Adv Eng Intell Syst.* 2026;005(02).
doi: 10.22034/aeis.2026.576366.1514
45. Moorthi M, Vanithalakshmi M, Hariharan N, Shyam KK, Yong LC, Athithan A. EEG based Brain-Computer Interface For Robotic Application. In: *2025 6th International Conference on Smart Electronics and Communication (ICOSEC).* IEEE; 2025:574-579.
doi: 10.1109/icosec67334.2025.11459790
46. Marrugat J, Arboix A, García-Eroles L, *et al.* Estimación de

la incidencia poblacional y la mortalidad de la enfermedad cerebrovascular establecida isquémica y hemorrágica en 2002 [The estimated incidence and case fatality rate of ischemic and hemorrhagic cerebrovascular disease in 2002 in Catalonia]. *Rev. Española Cardiol.* 2007;60(6):573-580. [In Spanish].

doi: 10.1157/13107113

47. Gasull T, Arboix A. Molecular mechanisms and pathophysiology of acute stroke: emphasis on biomarkers in the different stroke subtypes. *Int J Mol Sci.* 2022;23(16):9476. doi: 10.3390/ijms23169476