



AI-based stroke prediction system using body motion biosignals during walking

Jaehak Yu¹ · Sejin Park^{1,2} · Chee Meng Benjamin Ho³ · Soon-Hyun Kwon¹ · Kang-Hee cho⁴ · Yang Sun Lee⁵ 

Accepted: 12 November 2021 / Published online: 14 January 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

As the world population continues to age, chronic diseases are on the rise. One of these diseases is stroke, which is a dangerous disease that can lead to many social and economic difficulties. Strokes can cause persistent neurological sequelae and physical disabilities; in some cases, motor function of the upper or lower body may be impaired, resulting in abnormal walking patterns or a loss of walking ability. These differences can be captured through walking patterns and gait information. There is therefore a need for research examining systems that can quickly detect signs of stroke based on human biosignals collected during daily life, such as motor function and the walking speed of the upper and lower limbs. By accurately predicting the early symptoms of stroke diseases, neural damage can be reduced with timely visits to medical institutions for treatment. In this paper, we designed and implemented a new AI-based system using real-time motion data for predicting stroke in the elderly. Initial data were collected from elderly Koreans while walking with attached wearable sensors. The sensors were placed on both shoulders and quadriceps. The data were processed, and we obtained a total of 12 motion attributes (angles and acceleration information). Predictive models using machine learning and deep learning algorithms were then constructed. The performance of the proposed system was verified with high prediction accuracies of 98.25% for the C4.5 decision tree model, 98.72% for RandomForest, 96.60% for XGBoost, and 98.99% for long short-term memory (LSTM). Hence, this paper provides a method for quickly and accurately predicting the early onset of stroke based on motion attribute values obtained while walking.

Keywords Real-time stroke prediction · Body motion analysis · Machine learning · Deep learning · Stroke diseases · Walking analysis

✉ Yang Sun Lee
yslee48@hs.ac.kr

Extended author information available on the last page of the article

1 Introduction

The 4th Industrial Revolution has placed substantial focus on information and communications technologies (ICT), such as artificial intelligence (AI), the Internet of Things (IoT), big data, and blockchain. In the healthcare industry, there has been a renewed interest in telemedicine and other smart healthcare services [1–3]. AI is a technology through which machines have the ability to understand, perceive, and judge like humans [1–3]. There have already been several studies suggesting that AI can perform as well as or better than humans at key healthcare tasks, such as diagnosing disease. IoT technology is a hyper-network in which people and objects are connected to the Internet, thus allowing them to create, collect, share, and use information. Even without human intervention, many intelligent devices are capable of making decisions, working as groups, and sending information to the cloud automatically [1, 2]. Big Data refers to the vast amounts of data generated and collected in everyday activities. In the medical field, big data is used for the development of new drug and treatment recommendation services [4]. Blockchain is an innovative technology that does not store transaction books called blocks on a central server, but instead distributes them on personal computers which are then connected like chains for public storage. This is useful for personal health information because it is secure and because it enables individual transactions without the need for intermediaries [5, 6]. These ICT technologies collect, integrate, and store medical data such as Electronic Medical Records (EMR) and genomic data to make it big data, thus making it very easy to be used for various analyses and other purposes. The convergence of big data and artificial intelligence technology is expected to be highly useful in the development of digital health care and smart services based on vast amounts of medical and health data.

With the advancements that have been made in medicine, life expectancy has been prolonged, leading to an increase in the elderly population. The World Health Organization (WHO) reported that, with a rapidly ageing population, the number of chronic diseases can be expected to increase as well. Stroke is one of the highest prevalence diseases, and according to the WHO, 5.7 million people died from a stroke in 2016, ranking third after malignant neoplasm (cancer) and heart disease [7]. Stroke is a disease in which brain cells die because of the clogging or bursting of cerebral arteries, resulting in necrosis of brain cells, and in the worst case, death [8, 9]. Depending on the location and type of stroke, brain dysfunction such as hemiplegia and verbal or consciousness disorder may be observed. Stroke is known to cause disability in both the elderly and in adults [9, 10]. Patients can recover from a stroke if they are identified quickly and transferred to a medical institution for treatment. Therefore, a quick and accurate diagnosis is essential.

Due to its many symptoms, it is difficult to detect stroke early. While there has been substantial research into stroke risk factors, differences in the definition or methodology affect the prediction accuracy. The main risk factors that are consistently used for stroke prediction are smoking, diabetes, high blood pressure,

and obesity [10]. The Framingham Heart Study developed a stroke risk prediction model based on cardiovascular disease and cerebrovascular disease [11]. However, it is difficult to apply such models to elderly Koreans who have very different social and behavioral characteristics than the participants in those studies. In response to this, Jee [13] and Yu [14] developed prediction models of stroke disease for Koreans. Jee et al. [13] developed a prediction model (10-year average stroke risk prediction model) on the probability of stroke within 10 years using EMR items such as age, diabetes, alcohol consumption, smoking, total cholesterol, and body mass index. Meanwhile, Yu et al. [14] attempted a semantic interpretation of stroke patients using the NIHSS (National Institutes of Health Stroke Scale) value for the elderly at Chungnam National University Hospital. However, these initial studies used the same methodology as the Framingham Heart Study, which does not consider the possibility of death or competitive risk due to causes other than stroke. While past medical records and health check-up information are important, there is still a need to grasp the degree of risk in daily life. In other words, comprehensive analysis and prediction systems using patterns and biosignals are urgently required.

Walking is the most basic exercise for the human body [15, 16]. Gait is an alternating motion in which many skeletal muscles move the body in a constant direction through cooperation with various joints of the upper and lower extremities [15, 16]. Research has shown that stroke patients may suffer from a loss of symmetric posture, decreased walking ability, and impaired balance response and motor ability [17, 18]. When one's walking abilities are limited, it often leads to social isolation [17] and reduces the daily activities of that patient [19]. For stroke patients, recovering one's walking ability is an important factor that determines an individual's quality of life. Therefore, patterns in walking have emerged as a key component of early prediction for the onset of stroke and post-stroke patient rehabilitation [20]. Therefore, one should be able to detect stroke as soon as there are changes to a person's walking pattern. Therefore, there is a need for a precise and quick diagnosis method that detects pedestrian asymmetry.

In this paper, we designed and implemented a new AI system that automatically extracts important features from the collected motion data and predicts stroke disease in real time. The motion sensors were attached to all of the study participants, and data were collected from both stroke and elderly patients during walking. The collected data were sent to a server which predicts and analyzes stroke using machine learning or deep learning algorithms. Depending on the algorithms, our system was able to predict stroke using the shoulders and quadriceps angles, angular velocity, and angular acceleration. The performance of the proposed system was verified by confirming the prediction accuracies of 98.25% for the C4.5 decision tree model, 98.72% for RandomForest, 96.60% for XGBoost, and 98.99% for LSTM. Altogether, this study presents a machine learning-based model that can accurately detect and predict the prognostic symptoms of stroke disease in real time. This system overcomes the limitations of previous models that require history and EMR data.

The rest of this paper is structured as follows. Chapter 2 describes stroke and previous research methods. Chapter 3 describes the stroke disease prediction system

using motion biosignals obtained while walking, and Chapter 4 describes the experimental and analysis results in detail. Finally, Chapter 5 discusses the conclusion and future research topics.

2 Related work

2.1 Stroke disease

Stroke is a brain disorder caused by the sudden onset of neurological deficiency due to the blockage or bursting of blood vessels in the brain [21, 22]. Stroke patients generally have permanent dysfunction or complications such as motor, language, sensory, perceptual cognition, vision, and quadriplegia [22, 23]. There are two types of strokes: Cerebral infarction (ischemic) caused by the blockage of blood vessels and cerebral hemorrhage (cerebral hemorrhage) caused by the rupture of blood vessels [23]. Cerebral infarction is further divided into thrombosis, wherein blood clots form in the blood vessels damaged by arteriosclerosis, and cerebral embolism, wherein blood clots that have formed in the large arteries block the blood vessels to the brain. Meanwhile, cerebral hemorrhage can be further divided into intracerebral hemorrhage and subarachnoid hemorrhage. Cerebral hemorrhage in the brain parenchyma occurs spontaneously without external shock, and hypertension is its main cause. Subarachnoid hemorrhage is a disease in which a part of a blood vessel wall has weakened in elasticity and then swells in an alveolar shape, after which the cerebral aneurysm ruptures and blood leaks under the arachnoid membrane surrounding the brain [23, 24]. In particular, subarachnoid hemorrhage has been reported to be fatal enough to cause death before reaching the hospital in more than 30% of stroke patients [24].

Acute stroke disease causes disorders in the autonomic nervous system and central nervous system which are accompanied by heart problems such as arrhythmia in the electrocardiogram [25]. Acute stroke is a potentially fatal disease that can also cause permanent dysfunction and complications which cause difficulties in social or economic activities [26]. In particular, 85% of elderly patients in the early stages of stroke show upper limb disability, while 55–75% report upper limb disability over six months after onset [27]. The weakening of the upper limb muscle strength is caused by decreased nerve response and changes in the median nerve boundary, as well as muscle weakness due to inactivity [28]. Most social activities are restricted due to physical dysfunction, thus resulting in a poor quality of life. In addition, low motivation leads to mental disability such as depression, anger, and loss of pleasure, all of which adversely affect function recovery in stroke patients [29]. It is therefore necessary to early detect stroke symptoms that occur in the daily lives of the elderly, respond quickly, and make a thorough diagnosis by visiting the appropriate medical institutions. As a result, there is a need for research that can help minimize the social and economic damage caused by the aftereffects of stroke. Multilateral studies on the immediate detection and diagnosis of stroke diseases and various rehabilitation treatments are urgently needed to improve the quality of life of the elderly.

2.2 Previous Stroke Studies on motion

The human autonomic nervous system, sympathetic nervous system, and musculoskeletal system are integral for walking, and they presuppose human body movements [16]. The conditions of the walking load energy consumption decrease as the load approaches the center of the body. In walking, the feet are the only part of the human body in contact with the supporting ground, and they therefore play a crucial role in all weight loads [30]. The foot not only provides the driving force necessary for body-to-body movement when walking, but it also has a mechanism that adapts to irregular ground by absorbing the physical impact that occurs at such times. Each person's walking and upright posture are inherently unstable because they maintain a high-pressure center on a relatively small base plane provided by their feet. To provide additional stability, the muscles of the waist and legs are activated. This posture adjustment is very complex when performing movement in an unstable state, and it shows high muscle activity because it uses many muscles [31]. The measurement of walking ability can be measured by dividing it into qualitative and quantitative aspects; gait measurement (gait cycle, cadence, swing period, stance period, step length, stride length, and walking speed) provides a quantitative respective [20, 32, 33]. Regarding qualitative measurements, walking asymmetry is a typical measurement value, and it occurs when the base and raised angles of both feet do not match each other during walking. Walking speed is a typical quantitative measurement method, and it has the advantage of being able to be measured relatively easily and simply while ensuring reliability.

For stroke patients, walking asymmetry can typically be observed as a significant difference in the degree of asymmetry in the normal state through changes in the movement of the upper and lower extremities [32, 33]. Akay et al. [34] developed a technique to measure and quantify acceleration information comparing between stroke patients and healthy elderly while walking. They introduced a maximum likelihood estimator (MLE)-based fractal analysis for the complexity of body motion, and the results showed that people with stroke had significantly higher values than healthy people. Chen et al. [35] were able to classify stroke patients based on data from daily living activities collected using wearable devices. In that paper, daily activity data were analyzed, and 11 stroke subjects were selected. The collected data were subjected to machine learning algorithms such as Decision Tree, RandomForest, Support Vector Machine (SVM), and eXtreme Gradient Boosting (XGBoost). The results indicated that XGBoost can accurately predict seven daily activities. Carmo et al. [36] reported a statistical analysis of the changes in arm and leg movement as well as motion values during walking in stroke patients and control groups. The experimental results confirmed that stroke patients had a reduced range of motion in the glenohumeral and elbow joints when walking. Motion data measured during walking focuses on statistical and correlation analysis. Therefore, it can determine whether there is asymmetry due to stroke based on motion data collected in real time while walking. It is essential to research methods that can quickly diagnose stroke before receiving treatment from medical staff.

2.3 AI-based stroke research

Several researchers have applied AI technology to the field of predicting and diagnosing stroke [14, 37–43, 45]. For example, Yu et al. [14] published the results of an analytic and prediction stroke model based on the severity of NIHSS using a C4.5 decision tree algorithm, a representative predictive model of data mining. The rules of the operating principle provided by the C4.5 decision tree were analyzed in detail. However, there is still a need for an in-depth analysis of the data, as the rationale of the decision tree algorithm only provides a partial resolution. Shanti et al. [37] were able to detect the risk of stroke using an artificial neural network (ANN) algorithm. The paper showed that ANN's Backpropagation algorithm was used for learning, and that the consistency and diagnostic accuracy of the prediction was improved. A similar method by Hanifa et al. [38] used an ANN-based predictive model that can detect stroke patients with high accuracy of 95.33% from the experimental data of 300 people. Nevertheless, the classification and prediction of stroke diseases based on the ANN algorithm is still difficult to interpret, and its operating principles only emphasize accuracy. Hanifa et al. [39] presented the results of predicting and verifying stroke risk factors by adjusting the parameter values of the support vector machine (SVM) prediction model. A relatively accurate model which detailed the stroke risk situation was presented using the RBF kernel function of the SVM. However, the focus of that model was on determining the severity and prognosis after the onset of stroke, rather than detecting and predicting symptoms before a stroke. Chiun et al. [40] developed a system for detecting ischemic stroke using the extended patch images of CT scans as the input of the CNN model; it obtained greater than 90% prediction accuracy. Liu et al. [41] proposed a Res-CNN model that automatically classifies acute ischemic stroke lesions in multi-modality MRIs. The results confirmed that this model solved the problem of performance degradation using the residual unit, and that the model performance can be improved through data expansion. Chantamit et al. [43] introduced a method of integrating ICD-10 [44] codes into health records and integrating potential risk factors into predictive patterns and models in EMR information. Based on the integrated EMR information, a deep-learning LSTM-RNN model was applied for stroke prediction.

Recent studies have used different human biosignals to determine and predict stroke diseases in the elderly [23, 42–45]. For example, Choi et al. [23] used Electroencephalogram (EEG) biosignals to detect and predict stroke precursor symptoms. First, a signal at each timepoint was decomposed from raw EEG data with Fast Fourier Transform (FFT), thus giving values for alpha (α), beta (β), gamma (γ), delta (δ), and theta (θ) for six EEG channels. In addition, the ratio values between low β , high β , and θ were extracted and used with the RandomForest algorithm to obtain a stroke prediction accuracy of up to 92.51%. Yu et al. [42] proposed an early stroke detection method based on machine learning and deep learning that measures the electromyography (EMG) of thighs and calves during walking in daily life. They presented a methodology that can accurately detect and predict stroke precursor symptoms with EMG data collected in real time, as opposed to past information such as EMR or CT scans. The experimental results for elderly daily activities were verified with accuracies of 90.38% and 98.958% using the RandomForest algorithm.

Choi et al. [45] proposed a deep learning model that can predict stroke diseases using raw data that has not undergone frequency attributes of EEG biosignals and that has only been subjected to basic pretreatment. These studies confirmed that the LSTM model of deep learning is suitable for the time series analysis of biosignals. However, for biosignals measured and generated in daily life such as walking or driving, the convenience of wearing and the reliability of the data are paramount. Therefore, there is a need for a new alternative for predicting stroke diseases in the elderly during daily activities which can overcome the limitations of the traditional methodology as well as the disadvantages of biosignal measurement and collection.

3 Stroke disease prediction system using motion information while walking

We proposed an AI-based stroke disease prediction system that uses motion information collected during the daily activities of the elderly. Figure 1 shows the system, which consists of five modules in total: (1) Measures and collects motion data in real time; (2) Collects motion data in real time to extract data preprocess and important attributes; (3) Integrates and distributes collected motion data with ECG, EMG, EEG, and individual EMR data; (4) Predicts motion disease, and (5) Visualizes the results. The proposed system extracts and stores optimized preprocessing and attributes algorithms for various biosignals, such as ECG, EMG, and EEG, as well as motion data collected from the elderly during daily activities. Data for each of the biosignals are used as input values for both machine learning and deep learning models to provide optimal real-time prediction and analysis results for stroke diseases.

3.1 A module for measuring and collecting motion data while walking

In addition to the motion data, other biosignals such as ECG, EEG, and EMG were measured to predict stroke diseases in the elderly. All biosignals were

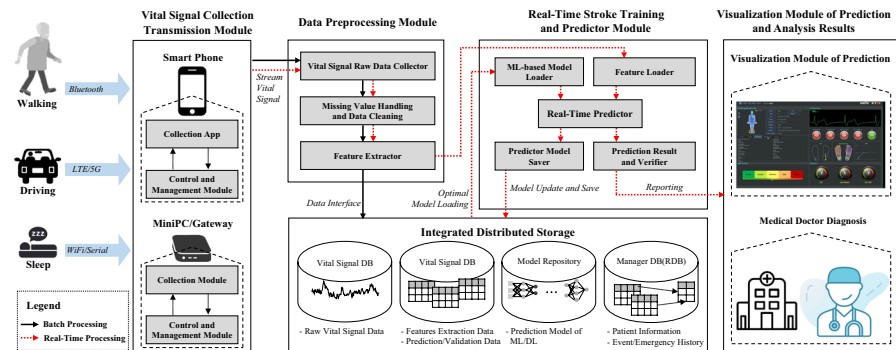


Fig. 1 The overall structure of the real-time stroke disease prediction system using motion biosignals

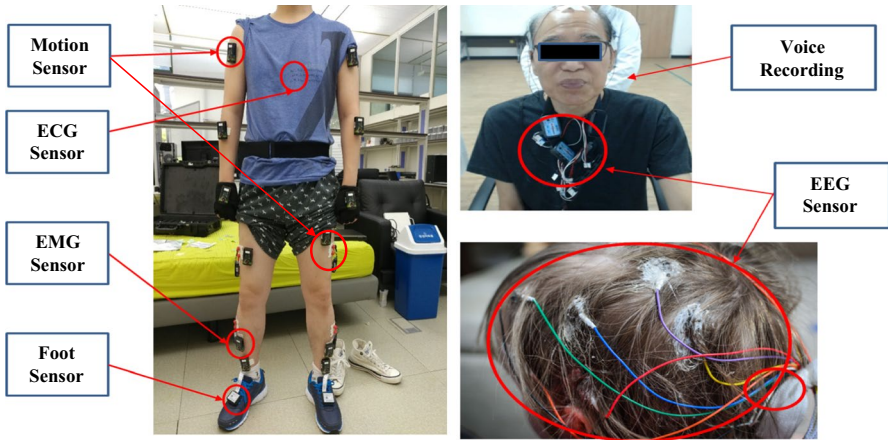


Fig. 2 The location of each sensor for collection of biosignals in real time

measured in real time using the Captiv motion analyzer [46, 47]. The collected data were used to verify the prediction and analysis for stroke in the elderly. Figure 2 shows the attachment locations of all the wearable sensors on each subject.

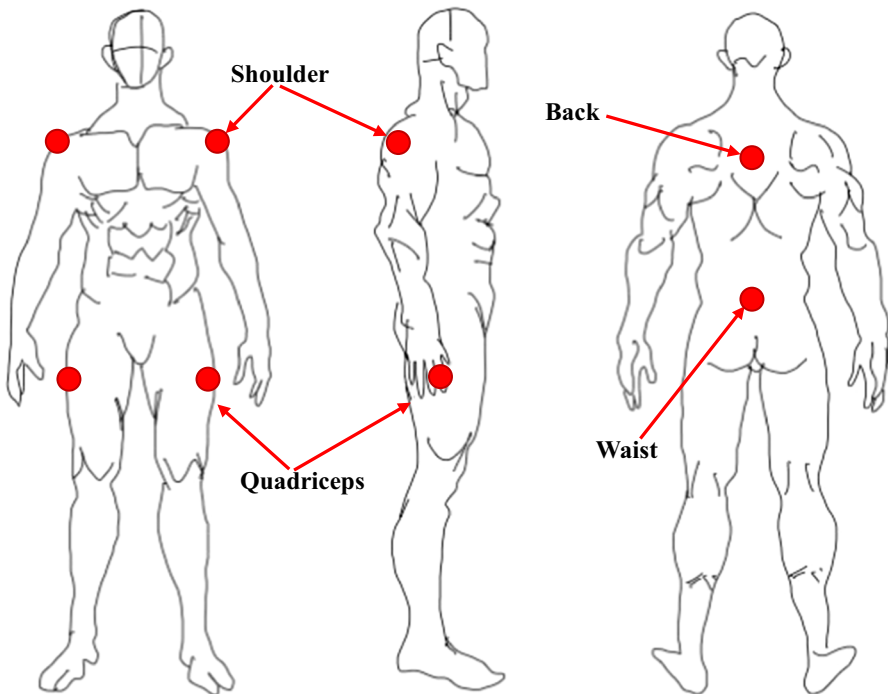


Fig. 3 Real-time measurement and collection of motion biosignals

Figure 3 shows the specific locations of sensors for motion biosignals. The Captiv motion analyzer was chosen for its high accuracy (gyroscope ± 2000 o/s, accelerometer ± 16 g, and magnetometer ± 2.5 gauss) and because it has been used in many medical and rehabilitation fields. To maximize the measurement accuracy, the sampling rate was set to 128 Hz. Raw data of four quaternions (qx, qy, qz, and qw) were collected from six locations on the body: (1) Left and right Shoulders, (2) Left and right quadriceps, (3) Back, and (4) Waist.

3.2 Pre-processing of motion data and extracting important attributes module

The motion data measured at the six locations (Fig. 3) are transmitted to a smartphone or gateway through BLE communication. Other methods, such as Wi-Fi, LTE, and 5G, can also be used depending on the system settings. The messaging standard was JSON (JavaScript Object Notation). Incomplete data or missing values raw data received from the module were removed. Since the categories of the minimum and maximum values are different for each data, a normalization process was also performed. The system was designed to extract the attribute values of angle, angular velocity, and angular acceleration based on the quaternion values of qx, qy, qz, and qw for each position. Table 1 summarizes the raw data of motion extracted and collected in this paper with some additional explanations.

By using the quaternion value for each motion data listed in Table 1, the angle, angular velocity, and angular acceleration properties were extracted for each measurement position and used in the experiment (Table 2). Since the waist and back are reference positions, separate attribute values were not extracted for these.

3.3 Module for integrating and distributed storage of data by multiple biosignals

The motion data-based stroke disease prediction system proposed in this paper can collect and manage various types of biosignals, such as motion, ECG, EEG, EMG, and PPG. The various biosignals including these motions are sequential values according to time. From the time-series data, the suitable attribute values for predicting stroke in the elderly are extracted according to Sect. 3.2. Therefore, for the real-time storage of raw data for each biosignal, the system was designed and developed to integrate and distribute storage as representative MongoDB of Not Only SQL (NoSQL). Depending on each of the biosignals, raw or processed data were extracted for statistical information and important attributes. These will be distributed and stored in a Relational Database (RDB) and NoSQL according to the characteristics. These raw data and important attribute values stored in the RDB and NoSQL can be used for disease prediction and multifaceted analysis based on machine learning and deep learning.

3.4 Machine learning and deep learning-based learning and real-time prediction modules

Real-time motion data collected from the elderly are time-series data that have sequential values over time. It is necessary to consider the time-series properties

Table 1 Description of 24 attributes and meanings extracted from motion raw data

Order	Attribute	Explanation
1–4	qx, qy, qz, and qw values of the right quadriceps	The quaternion attribute values for the qx, qy, qz, and qw positions of the right quadriceps
5–8	qx, qy, qz, and qw values of the left quadriceps	The quaternion attribute values for the qx, qy, qz, and qw positions of the left quadriceps
9–12	qx, qy, qz, and qw values of the right shoulder	The quaternion attribute values for the qx, qy, qz, and qw positions of the right shoulder
13–16	qx, qy, qz, and qw values of the left shoulder	The quaternion attribute values for the qx, qy, qz, and qw positions of the left shoulder
17–20	qx, qy, qz, and qw values at the waist	The quaternion attribute values for the qx, qy, qz, and qw positions of the waist
21–24	qx, qy, qz, and qw values at the back	The quaternion attribute values for the qx, qy, qz, and qw positions of the back

Table 2 Important attributes and descriptions extracted from motion data

Order	Features	Explanation
1	Shoulder_Left_Angle	Angle value of the left shoulder
2	Shoulder_Left_Angle_Velocity	Angular velocity value of the left shoulder
3	Shoulder_Left_Angle_Acceleration	Angular acceleration value of the left shoulder
4	Shoulder_Right_Angle	Angle value of the right shoulder
5	Shoulder_Right_Angle_Velocity	Angular velocity value of the right shoulder
6	Shoulder_Right_Angle_Acceleration	Angular acceleration value of the right shoulder
7	Quadriceps_Left_Angle	Angle value of the left thigh
8	Quadriceps_Left_Angle_Velocity	Angular velocity value of the left thigh
9	Quadriceps_Left_Angle_Acceleration	Angular acceleration value of the left thigh
10	Quadriceps_Right_Angle	Angle value of the right thigh
11	Quadriceps_Right_Angle_Velocity	Angular velocity value of the right thigh
12	Quadriceps_Right_Angle_Acceleration	Angular acceleration value of the right thigh
13	Class	Discriminant information of elderly/stroke patients (labeling with normal/stroke)

of motion data. Therefore, the LSTM model of deep learning, which can express the properties of time series data, was used to predict stroke diseases. This LSTM method has the advantage of solving the long-term dependence problem of RNN in deep learning. Further, a predictive model using machine learning methodology for the values of angle, angular velocity, and angular acceleration attributes from each measurement location was extracted. This machine learning method has the additional advantage of being able to receive predictive performance and hermetic information for stroke disease.

The machine learning- and deep learning-based learning and prediction module proposed in this paper constitutes the following two subblocks (see Fig. 4): first, in the batch processing block, machine learning and deep learning are performed by

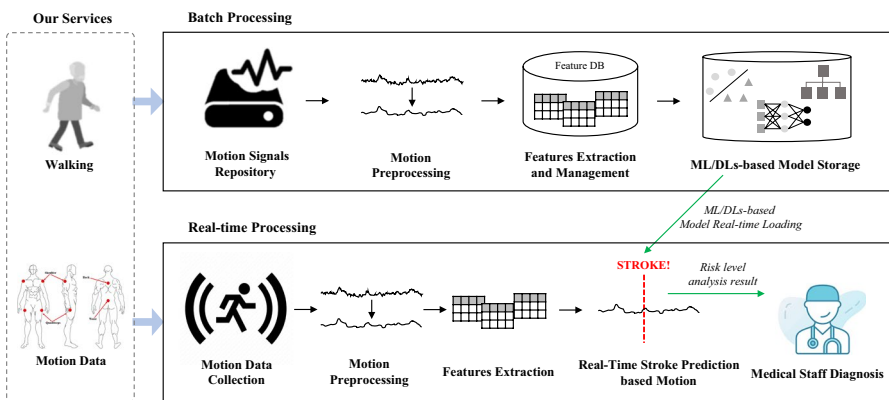


Fig. 4 Machine learning and deep learning-based learning and real-time prediction modules

storing and preprocessing motion data collected in real time to extract important attributes. In the real-time processing block, stroke diseases are predicted as soon as motion data are measured and collected during walking of the elderly, and the risk values are provided to the medical staff, the patient themselves, their next-of-kin, and the visualization module described in Sect. 3.5. In the real-time processing block, raw motion data of more than 3 s were collected, features were extracted and predictions were executed. Through the experiments, it was confirmed that the extracted features for machine learning-based algorithms and raw data for deep learning-based algorithms, provided prediction results within 0.05 s.

3.5 A module that visualizes and provides prediction and analysis results

The visualization and prediction information of stroke created using the motion and the various biosignals obtained during walking are provided (see module on the right). The screen is designed to provide predictive results and semantic analysis information of stroke to the medical staff, the patients themselves, and their next of kin. According to the judgment of stroke risk analysis, medical staff or hospitals may support the patient by contacting the patient and quickly transporting them to the hospital to conduct precise examination and diagnosis services. However, various clinical difficulties and risks exist in predicting and determining stroke diseases using raw data and the important attributes of motion. The judgment of an experienced medical staff with professional medical knowledge is essential in distinguishing whether an elderly person is having a stroke based on various biosignals, including motion. Consequently, this study conducted experiments and verification to provide scientific and meaningful information that enables the faster and more accurate judgment of medical staff and hospitals, rather than a system that directly predicts the precursors and risk of stroke in elderly patients.

4 Experiment and analysis

4.1 Data set and experimental environment

This section describes the verification and performance of the system for the prediction of stroke using motion. Sensors were placed at six locations on the patient. The waist and back sensors were used as reference points while the real-time data were collected from the left and right shoulders and quadriceps at a sampling rate of 128 Hz per second. This motion data positioning is because the joint movement and the movement speed of the upper and lower bodies can be measured in a patient with stroke symptoms during walking. To elaborate, it is possible to check the movement of the upper body and the movement range both in the standing state and during walking, and the movement speed can be determined according to the change of the main joints and muscles of the pelvis in the lower body. These raw data were sent to the server for processing. Next, the angle, angular velocity, and angular acceleration were automatically calculated

and obtained. These motion biosignals are widely used in the study of gait changes and rehabilitation exercises in patients with hemiplegia. In the case of stroke, patients' walking patterns and related factors are analyzed based on fine symptoms, as the decreases in walking speed and the increases in the duration of the foot on the ground compared to the settler are evaluated. The motion information angle, angular velocity, and angular acceleration values of the shoulders and quadriceps are important precursor symptoms of stroke. Therefore, in this paper, we intend to analyze walking disorders caused by minute symptoms or balance abnormalities during walking using motion data that can serve as medical and kinematic parameters.

The motion data of stroke elderly and normal elderly were obtained at Chungnam National University Emergency Center and Department of Rehabilitation Medicine. From 2017 to 2018, various biosignals such as ECG, EEG, and EMG, including motion, were measured and collected from elderly participants over 65 years old [42, 45]. Patients who were confirmed to have had a stroke within one month and who were undergoing rehabilitation treatment were selected as stroke subjects. In 2017, data were obtained from 48 elderly stroke patients and 75 controls; in 2018, data were obtained from 13 stroke patients and 137 controls. To balance the numbers of stroke patients and controls, all 61 stroke patients and 61 randomly extracted controls were selected. All subjects repeated each scenario a total of five times: walking, standing, sitting in a chair, raising and lowering their arms, speaking, and sleeping. The first and last measurements were excluded from the experimental data, as the subjects may have experienced discomfort in wearing the sensors, tension, and fatigue from repetition. The model development in this paper was conducted in an environment using Ubuntu 18.04.4 LTS, an Intel Core i9-10900X CPU, an NVIDIA Quadro RTX 8000 GPU, and 256 MB RAM.

4.2 Performance indicators of experimental results

This section defines the performance evaluation indicators used to verify the performance of the stroke system with motion data [14, 23, 42, 45, 48].

- Accuracy: The percentage of people who have tested positive for stroke and negative for normal and control groups among all people.
- F1-score: The harmonic mean of precision and recall.
- Recall: The percentage of stroke patients that have tested positive.
- Precision: The proportion of people who are stroke patients among those who have tested positive.

When a stroke patient is misclassified as a normal person, the untreated disease can have an impact on that person's daily life. Therefore, accuracy is the most important performance indicator in the field of disease prediction and healthcare. A predictive model should minimize the rate of misclassification of stroke patients as normal people (see Table 3).

Table 3 Confusion matrix of performance evaluation for stroke prediction

Predicted condition	True condition	
	Stroke group	Control group
Stroke group	TP ^a	FP ^b
Control group	FN ^c	TN ^d

^aTP (True Positive): predicting stroke elderly as stroke elderly

^bFP (False Positive): predicting stroke elderly as general elderly (normal)

^cFN (False Negative): predicting general elderly (normal) as stroke elderly

^dTN (True Negative): predicting general elderly (normal) as general elderly

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{FP} + \text{TN}} \quad (1)$$

$$\text{F1 - Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (2)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (4)$$

4.3 Experiment and analysis based on machine learning

In this section, the attribute values of angle, angular velocity, and angular acceleration for each measurement location were put through a machine learning methodology. This experiment used C4.5 decision tree, C5.0 decision tree, alternating decision tree, RandomForest, logistic regression, naïve Bayes, multi-layer perceptron, SVM, C&RT, XGBoost, and QUEST. Each machine learning-based stroke disease prediction model attempted to achieve the highest accuracy based on different learning and prediction datasets. Two different experiments were conducted. In these experiments, 67% of the randomly extracted data were used for learning and the rest of the data were used for testing. Likewise, data extracted by the same method at a rate of 80% were used for learning, and the remaining 20% of data were used for testing. Finally, the entire dataset was divided into 5-Fold, 10-Fold, and 20-Fold Cross-Validation, and the predictive model verification and experiment were conducted.

In total, 12 motion attribute values (Sect. 3.2) were used in this experiment. Since motion data change at a rate of 128 Hz per second based on the sensors, data were extracted in units of 0.1 s when extracting attribute values. In total, 122 subjects comprising 61 stroke patients and 61 elderly people with a total of

Table 4 Predictive accuracy and F1-score (%) for each algorithm using 12 attribute values of motion

Methods	Data sets									
	Train(67)/ Test(33)		Train(80)/ Test(20)		5-Fold CV		10-Fold CV		20-Fold CV	
	Acc.	F1	Acc	F1	Acc	F1	Acc	F1	Acc	F1
C4.5 DT	96.81	96.80	96.90	90.90	96.83	96.82	96.93	96.92	97.12	97.11
C5.0 DT	94.18	94.18	94.34	94.35	94.24	94.23	94.59	94.59	94.66	94.66
Alternating DT	86.55	86.55	87.06	87.03	87.02	87.03	87.15	87.15	87.16	87.16
RandomForest	97.61	97.61	97.78	97.77	97.76	97.77	97.93	97.93	98.01	98.01
LR	62.86	62.86	63.32	63.31	63.00	63.00	63.12	63.41	63.14	63.14
Naive Bayes	64.17	65.55	65.20	66.42	65.35	66.55	65.29	66.80	65.33	67.27
MLP	84.23	84.23	84.78	84.78	85.76	85.77	87.34	87.34	86.98	86.97
SVM	81.77	81.77	81.57	81.57	81.79	81.76	82.20	82.19	81.25	81.26
C&RT	83.58	83.58	83.72	83.57	83.51	83.52	83.74	83.74	83.71	83.72
XGBoost	95.52	95.51	95.57	95.61	95.61	95.61	95.67	95.66	95.69	95.68
QUEST	80.82	80.82	80.90	80.86	80.87	80.86	81.05	81.05	81.11	81.11

CV cross-validation, Acc. accuracy, F1 F1-score, DT decision tree, LR logistic regression, MLP multi-layer perceptron

Table 5 Recall and Precision (%) for each algorithm using 12 attribute values of motion

Methods	Data sets									
	Train(67)/ Test(33)		Train(80)/ Test(20)		5-Fold CV		10-Fold CV		20-Fold CV	
	Recall	Prec.	Recall	Prec	Recall	Prec	Recall	Prec	Recall	Prec
C4.5 DT	96.80	96.80	96.90	96.91	96.82	96.82	96.92	96.93	97.11	97.11
C5.0 DT	94.19	94.17	94.33	94.37	94.23	94.24	94.60	94.58	94.67	94.65
Alternating DT	86.56	86.55	87.07	87.07	87.03	87.03	87.17	87.14	87.17	87.16
RandomForest	97.62	97.61	97.77	97.80	97.75	97.79	97.92	97.95	98.00	98.02
LR	62.87	62.85	63.33	63.30	62.97	63.03	63.11	63.72	63.14	63.14
Naive Bayes	64.18	66.99	65.18	67.71	65.27	67.89	65.21	68.46	65.30	69.37
MLP	84.22	84.24	84.77	84.79	85.77	85.77	87.34	87.34	86.97	86.98
SVM	81.79	81.76	81.57	81.57	81.75	81.77	82.16	82.23	81.22	81.30
C&RT	83.57	83.59	83.71	83.73	83.55	83.49	83.75	83.74	83.74	83.71
XGBoost	95.51	95.52	95.58	95.56	95.63	95.60	95.68	95.65	95.69	95.68
QUEST	80.81	80.84	80.89	80.92	80.88	80.85	81.07	81.04	81.12	81.10

Prec. Precision

283,734 attribute values (141,867 cases each) were extracted and tested. To verify the performance in this experiment, the four performance indicators defined in Sect. 4.2 were used for verification. Tables 4 and 5 present the data sets classified into each machine learning algorithm using 12 motion attribute values, and the

Table 6 Predictive accuracy and F1-score (%) for each algorithm applying Z-score to 12 attributes of motion

Methods	Data sets									
	Train(67)/ Test(33)		Train(80)/ Test(20)		5-Fold CV		10-Fold CV		20-Fold CV	
	Acc	F1	Acc	F1	Acc	F1	Acc	F1	Acc	F1
C4.5 DT	97.82	97.81	97.88	97.88	98.15	98.15	98.25	98.26	98.22	98.22
C5.0 DT	95.12	95.12	95.11	95.12	95.23	95.24	95.37	95.37	95.35	95.34
Alternating DT	88.01	88.01	88.12	88.12	88.07	88.06	88.36	88.35	88.39	88.37
RandomForest	98.57	98.58	98.60	98.59	98.67	98.66	98.72	98.72	98.71	98.70
LR	64.99	64.97	65.03	65.04	64.97	64.97	65.22	65.22	65.34	65.35
Naive Bayes	66.02	66.02	66.19	66.19	66.33	66.34	66.52	66.53	66.49	66.50
MLP	86.13	86.14	86.19	88.18	86.57	65.55	87.98	87.98	88.06	88.06
SVM	83.54	83.53	83.72	83.73	83.69	83.68	84.28	84.29	84.45	84.44
C&RT	83.37	83.37	85.47	85.46	85.39	85.38	86.61	86.61	86.66	86.44
XGBoost	96.36	96.36	96.47	96.48	96.38	96.38	96.60	96.61	96.63	96.64
QUEST	82.76	82.76	82.79	82.80	82.82	82.82	83.34	83.34	83.16	83.15

Table 7 Recall and precision (%) by algorithm applying Z-score to 12 attributes of motion

Methods	Data sets									
	Train(67)/ Test(33)		Train(80)/ Test(20)		5-Fold CV		10-Fold CV		20-Fold CV	
	Recall	Prec.	Recall	Prec	Recall	Prec	Recall	Prec	Recall	Prec
C4.5 DT	97.81	97.81	97.90	97.87	98.15	98.15	98.27	98.26	98.21	98.23
C5.0 DT	95.11	95.14	95.14	95.10	95.22	95.26	95.39	95.36	95.33	95.36
Alternating DT	88.00	88.02	88.14	88.11	88.09	88.04	88.39	88.32	88.38	88.36
RandomForest	98.61	98.56	98.61	98.57	98.66	98.67	98.74	98.71	98.71	98.70
LR	64.99	64.96	65.07	65.01	65.00	64.95	65.24	65.21	65.37	65.31
Naive Bayes	66.00	66.05	66.19	66.19	66.41	66.27	66.56	66.50	66.53	66.47
MLP	86.15	86.14	88.22	88.15	86.58	86.52	87.99	87.97	88.09	88.04
SVM	83.51	83.56	83.77	83.69	83.68	83.69	84.33	84.26	84.46	84.43
C&RT	83.33	83.42	85.49	85.44	85.40	85.37	86.63	86.59	86.68	86.61
XGBoost	96.34	96.39	96.49	96.48	96.38	96.38	96.63	96.59	96.62	96.66
QUEST	82.76	82.77	82.83	82.77	82.83	82.81	83.37	83.31	83.14	83.17

performance indicators of prediction accuracy, F1-score, recall, and precision are summarized.

Next, the 12 attributes of motion data were tested after normalization using the Z-score method (Tables 6 and 7). However, the deviation of the maximum and minimum values is very large for the 12 attributes. This creates a problem of bias in the

measurement units, which adversely affects learning. Therefore, stable learning and prediction results can be guaranteed when the same weight is applied by converting the values to the same range from 0.0 to 1.0 for each attribute through the normalization method. In Eq. (5) below, σ and μ respectively refer to the standard deviation and average of the attribute x , α is the weight value, and 1.0 was set in the experiment conducted in this paper.

$$\bar{x}_i = \frac{x_i - \mu}{\sigma} \times \alpha \quad (5)$$

After normalization, the overall performance improvement in terms of prediction accuracy for stroke in the elderly was about 1.0% (Table 6). The prediction model with the highest prediction accuracy was RandomForest, which had 98.72% accuracy. For the parameter setting of the RandomForest algorithm, the number of trees was set to 50, the random seed value was set to 1, the tree depth was set to 30, and the minimum number of node splits was set to 2.

4.4 Experiment and analysis based on deep learning

Experiments on deep learning to predict stroke were performed with LSTM models due to its proven track record in time series analysis [42, 45]. LSTM is a model that overcomes the structural shortcomings of the existing RNN, and it is designed to solve vanishing gradients where the amount of computation increases and the value decreases as the error value is backpropagated [42, 45]. LSTM consists of four cell states along with input, forget, and output gates. This LSTM has a structure that transfers past information to the next state. At this time, the vector value output from each gate is generated through the sigmoid layer and the tanh layer. As a result, the LSTM learns long-term dependent values, and the cell state operates in the order of passing information from the past to the next state. All 24 raw data points from Sect. 3.2 were used. To generate a predictive model, learning data were randomly extracted, and the remaining data not used for learning were used for testing. To verify the experimental and prediction results, the data segmentation ratio was divided into 67% to 33% as well as 80% to 20%.

In this experiment, the important parameters for performance verification using motion data-based deep learning LSTM were set, and they are described as follows. Among the parameters, nUnit means the number of cells in the LSTM network and Iteration means the total number of learning iterations. Next, the learning rate is a scheduling value that adjusts the rate of learning. It is important to set appropriate values as learning rates do not reduce learning errors; overfitting or underfitting may occur when learning progresses too fast or too slow. 1st Decay LR (Learning Rate) can prevent overfitting and induce stable learning error reduction by multiplying the initial learning rate by the 0.1 value from the corresponding number of learning. "2nd Decay LR" was used for stable convergence to determine the optimal predictive model by multiplying the reduced learning rate in "1st Decay LR" by 0.1 again. A hidden node is a parameter value for setting the number of hidden nodes included in a cell in one LSTM. In this

Table 8 Accuracy of predicting stroke diseases based on LSTM using motion raw data (%)

Number	Parameters						Accuracy (%)	
	Iteration	nUnit	Learning rate	1st Decay LR	2st Decay LR	Hidden Neuron	67/33	80/20
1	1000	64	0.01	500	750	192	94.753	94.761
2	1000	256	0.001	500	750	512	95.568	95.564
3	2000	128	0.01	1000	1500	526	96.551	96.562
4	2000	256	0.001	1000	1500	716	97.688	97.731
5	3000	128	0.01	1500	2250	256	97.980	97.982
6	3000	128	0.001	1500	2250	512	98.273	98.281
7	5000	256	0.005	1500	2250	768	98.992	98.994
8	5000	512	0.005	1500	2250	1024	98.885	98.887
9	10,000	256	0.01	2500	3750	764	97.995	97.998
10	10,000	512	0.001	2500	3750	1536	98.436	98.440

experiment, Adam optimizer, which showed a stable performance in prediction accuracy and loss among various optimizers of LSTM, was selected and used to predict stroke in the elderly. As presented in Table 8, when 80% of the data were randomly extracted for learning and the remaining 20% was used for validation, a stroke prediction accuracy of 98.994% was obtained.

Figure 5 shows the prediction accuracy and error rate with 80/20 data from the 7th experiment in Table 8. In Fig. 5b, the error is reduced according to each iteration, and a stable stroke disease prediction is guaranteed. As a result, the LSTM model showed optimal predictive performance when using 5000 learning iterations as well as a learning rate of 0.005 and with the number of hidden neurons (or hidden nodes) in the cell set to three times the nUnit parameter.

Figure 6 shows the ROC (receiver operating characteristic) curve of the LSTM models in Table 8. The ROC curve expresses the threshold and performance of

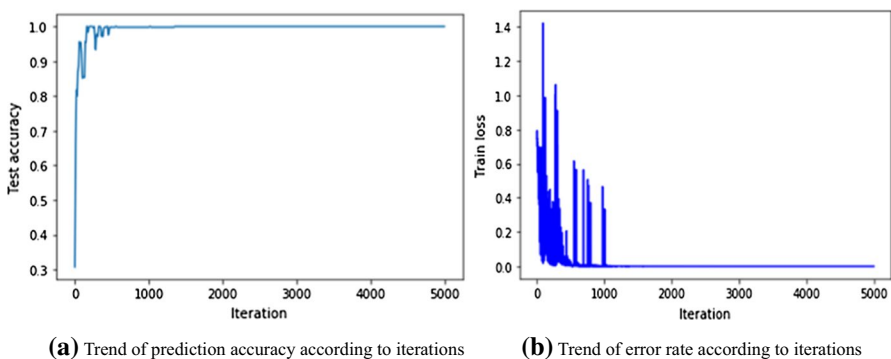


Fig. 5 Changes in prediction accuracy and error rate through LSTM model. **a** Trend of prediction accuracy according to iterations. **b** Trend of error rate according to iterations

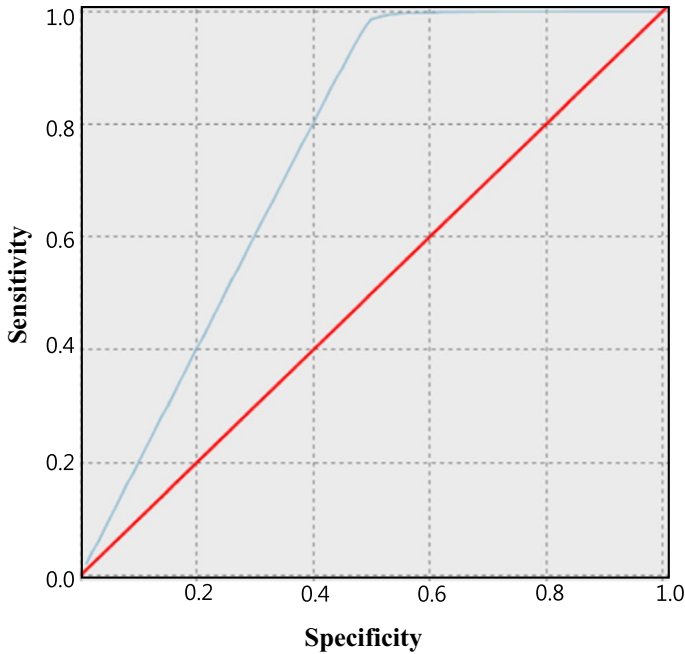


Fig. 6 The ROC curve of LSTM model using raw motion biosignals

binary classification prediction of elderly stroke disease, where the x -axis indicates specificity, and the y -axis indicates sensitivity.

5 Conclusion

In this paper, we propose a new system that can predict stroke in the elderly using machine learning and deep learning algorithms based on real-time motion biosignals. The proposed system overcomes the limitations of the current predictive models for stroke occurrence, which are based on the past 10 years of medical history and EMR data, or the stroke risk factors suggested in the Framingham Heart Study. The proposed achieved prediction accuracies of 98.25% for the C4.5 decision tree model, 98.72% for RandomForest, and 96.60% for XGBoost with attribute values obtained during walking. For deep learning experiments, the LSTM model achieved 98.99% accuracy for stroke disease prediction. To conclude, we showed that the impairment of the upper and lower motor functions can be used as a prognostic symptom of stroke disease for early detection. The use of AI can reduce the number of false positives and allow for a preemptive response for stroke to be taken.

In a future research project, we intend to conduct research and develop a service predictive of stroke by organically combining various biosignals and EMR by measuring and collecting real-time biosignals in sleep or driving situations.

Acknowledgements This work was supported by the National Research Council of Science & Technology (NST) grant by the Korea government (MSIP) (No. CRC-15-05-ETRI).

Declarations

Ethical approval The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committee of Chungnam National University (No. 2016-05-022, approved in 5 July 2016).

References

1. Mohammadi M, Al-Fuqaha A, Sorour S, Guizani M (2018) Deep learning for IoT Big data and streaming analytics: a survey. *IEEE Commun Surv Tutor* 20(4):2923–2960. <https://doi.org/10.1109/COMST.2018.2844341>
2. Ajayi OO, Bagula AB, Ma K (2019) Fourth industrial revolution for development: the relevance of Cloud federation in healthcare support. *IEEE Access* 7:185322–185337. <https://doi.org/10.1109/ACCESS.2019.2960615>
3. Garcia AR (2019) AI, IoT, Big data, and technologies in digital economy with blockchain at sustainable work satisfaction to smart mankind: access to 6th dimension of human rights. In: *Smart governance for cities: perspectives and experiences*, pp 83–131. https://doi.org/10.1007/978-3-030-22070-9_6
4. Davenport T, Kalakota R (2019) The potential for artificial intelligence in healthcare. *Future Healthc J* 6(2):94–98. <https://doi.org/10.7861/futurehosp.6-2-94>
5. Jeong S, Kim YW, Youn CH (2014) Personalized healthcare system for chronic disease care in cloud environment. *ETRI J* 36(5):730–740. <https://doi.org/10.4218/etrij.14.2214.0070>
6. Abuidris Y, Kumar R, Yang T, Onginjo J (2021) Secure large-scale E-voting system based on blockchain contract using a hybrid consensus model combined with sharding. *ETRI J* 43(2):357–370. <https://doi.org/10.4218/etrij.2019-0362>
7. World Health Organization (WHO) The top 10 causes of death. Accessed 24 May 2018. <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death>
8. Subudhi A, Dash M, Sabut S (2020) Automated segmentation and classification of brain stroke using expectation-maximization and random forest classifier. *Biocybern Biomed Eng* 40(1):277–289. <https://doi.org/10.1016/j.bbe.2019.04.004>
9. Lee HJ, Lee JS, Choi JC, Cho YJ et al (2017) Simple estimates of symptomatic intracranial hemorrhage risk and outcome after intravenous thrombolysis using age and stroke severity. *J Stroke* 19(2):229–231. <https://doi.org/10.5853/jos.2016.01109>
10. Malik V, Ganesan AN, Selvanayagam JB, Selvanayagam JB, Chew DP, McGavigan AD (2020) Is atrial fibrillation a stroke risk factor or risk marker? An appraisal using the Bradford hill framework for causality. *J Heart Lung Circ* 29(1):86–93. <https://doi.org/10.1016/j.hlc.2019.08.005>
11. Wolf PA, D'Agostino RB, Belanger AJ, Kannel WB (1991) Probability of stroke: a risk profile from the Framingham study. *Am Heart Assoc* 22:312–318. <https://doi.org/10.1161/01.str.22.3.312>
12. D'Agostino PA, Wolf AJ, Belanger, and W. B. Kannel, (1994) Stroke risk profile: adjustment for antihypertensive medication: the Framingham Study. *Am Heart Assoc* 25:40–43. <https://doi.org/10.1161/01.str.25.1.40>
13. Jee SH, Park JW, Lee SY, Nam BH, Ryu HG, Kim SY et al (2008) Stroke risk prediction model: a risk profile from the Korean study. *Atherosclerosis* 197(1):318–325. <https://doi.org/10.1016/j.atherosclerosis.2007.05.014>
14. Yu J, Park S, Lee H, Pyo CS, Lee YS (2020) An elderly health monitoring system using machine learning and In-depth analysis techniques on the NIH stroke scale. *Mathematics* 8(7):1–16. <https://doi.org/10.3390/math8071115>
15. Perry J (1992) Gait analysis: normal and pathological function
16. Abe D, Muraki S, Yasukouchi A (2007) Ergonomic effects of load carriage on energy cost of gradient walking. *Appl Ergonom* 39(2):144–149. <https://doi.org/10.1016/j.apergo.2007.06.001>

17. Carr JH, Shepherd RB, Nordholm L, Lynne D (2002) Investigation of a new motor assessment scale for stroke patients. *Phys Ther* 65(2):175–180. <https://doi.org/10.1093/ptj/65.2.175>
18. Østbye T, Levy AR, Mayo NE (1997) Hospitalization and case-fatality rates for subarachnoid hemorrhage in Canada from 1982 through 1991. The Canadian collaborative study group of stroke hospitalizations. *Stroke* 28(4):793–798. <https://doi.org/10.1161/01.str.28.4.793>
19. Guralnik JM, Ferrucci L, Balfour JL, Volpato S, Di Iorio A (2001) Progressive versus catastrophic loss of the ability to walk: implications for the prevention of mobility loss. *J Am Geriatr Soc* 49(11):1463–1470. <https://doi.org/10.1046/j.1532-5415.2001.4911238.x>
20. Mauritz KH (2002) Gait training in hemiplegia. *Eur J Neurol* 9(s1):23–29. <https://doi.org/10.1046/j.1468-1331.2002.0090s1023.x>
21. Seo K, Kang MJ, Kim GS, Lee JH, Suh SH, Lee K (2020) National trends in clinical outcomes of endovascular therapy for ischemic stroke in South Korea between 2008 and 2016. *J Stroke* 22(3):412–415. <https://doi.org/10.5853/jos.2020.01928>
22. Radomski MV, Trombly CA (2013) Occupational therapy for physical dysfunction, 7th, North American Edition, Alphen aan den Rijn. Wolters Kluwer, Netherlands
23. Choi YA, Park S, Jun JA, Ho CMB, Pyo CS, Lee H (2021) Machine-learning-based elderly stroke monitoring system using electroencephalography vital signals. *Appl Sci* 11(4):1–18. <https://doi.org/10.3390/app11041761>
24. Johansson BB (1999) Hypertension mechanisms causing stroke. *Clin Exp Pharmacol Physiol* 26(7):563–565. <https://doi.org/10.1046/j.1440-1681.1999.03081.x>
25. Gottesman RF, Hillis AE (2010) Predictors and assessment of cognitive dysfunction resulting from ischaemic stroke. *Lancet Neurol* 9(9):895–905. [https://doi.org/10.1016/S1474-4422\(10\)70164-2](https://doi.org/10.1016/S1474-4422(10)70164-2)
26. Pikija S, Trkulja V, Ramesmayer C, Mutzenbach JS, Oberpfalzer MK, Hecker C, Bubel N, Füssel MU, Sellner J (2018) Higher blood pressure during endovascular thrombectomy in anterior circulation stroke is associated with better outcomes. *J Stroke* 20(3):373–384. <https://doi.org/10.5853/jos.2018.01305>
27. Feys HM, De Weerd WJ, Selz BE et al (1998) Effect of a therapeutic intervention for the hemiplegic upper limb in the acute phase after stroke a single-blind, randomized, controlled multicenter trial. *Stroke* 29(4):785–792. <https://doi.org/10.1161/01.str.29.4.785>
28. Patten C, Lexell J, Brown HE (2004) Weakness and strength training in persons with poststroke hemiplegia: rationale, method, and efficacy. *J Rehabil Res Dev* 41(3A):293–312. <https://doi.org/10.1682/jrrd.2004.03.0293>
29. Kim YM, Park YH (2014) The grading of cognitive state comparisons with different distances across three conditions in stroke survivors. *Phys Ther Rehabil Sci* 3(2):134–141. <https://doi.org/10.14474/ptrs.2014.3.2.134>
30. Saltzman CL, Nawoczenski DA (1995) Complexities of foot architecture as a base of support. *J Orthop Sports Phys Ther* 21(6):354–360. <https://doi.org/10.2519/jospt.1995.21.6.354>
31. Aruin AS, Latash ML (1995) Directional specificity of postural muscles in feed-forward postural reactions during fast voluntary arm movements. *Exp Brain Res* 103(2):323–332. <https://doi.org/10.1007/BF00231718>
32. Kollen B, Kwakkel G, Lindeman E (2006) Hemiplegic gait after stroke: is measurement of maximum speed required? *Arch Phys Med Rehabil* 87(3):358–363. <https://doi.org/10.1016/j.apmr.2005.11.007>
33. Sungkarat S, Fisher BE, Kovindha A (2011) Efficacy of an insole shoe wedge and augmented pressure sensor for gait training in individuals with stroke: a randomized controlled trial. *Clin Rehabil* 25(4):360–369. <https://doi.org/10.1177/0269215510386125>
34. Akay M, Sekine M, Tamura T, Higashi Y, Fujimoto T (2004) Fractal dynamics of body motion in post-stroke hemiplegic patients during walking. *J Neural Eng* 1(2):111–116. <https://doi.org/10.1088/1741-2560/1/2/006>
35. Chen PW, Baune NA, Zwir I, Wang J, Swamidass V, Wong AWK (2021) Measuring activities of daily living in stroke patients with motion machine learning algorithms: a pilot study. *J Environ Res Public Health* 18(4):1634. <https://doi.org/10.3390/ijerph18041634>
36. Carmo AA, Kleiner AFR, Costa PH, Barros RM (2012) Three-dimensional kinematic analysis of upper and lower limb motion during gait of post-stroke patients. *Braz J Med Biol Res* 45(6):537–545. <https://doi.org/10.1590/s0100-879x2012007500051>
37. Shanthi D, Sahoo G, Saravanan N (2009) Designing an artificial neural network model for the prediction of thrombo-embolic stroke. *Int J Biometr Bioinform IJBB* 3(1):10–18

38. Hanifa SM (2010) Prediction of stroke risk through stacked topology of ANN model. *Int J Adv Res Comput Sci* 1(4):170–179. <https://doi.org/10.26483/ijarcs.v1i4.164>
39. Hanifa SM, Raja-S K (2010) Stroke risk prediction through non-linear support vector classification models. *Int J Adv Res Comput Sci* 1(3):47–53. <https://doi.org/10.26483/ijarcs.v1i3.60>
40. Chin CL, Lin BJ, Wu GR, Weng TC, Yang CS, Su RC, Pan YJ (2017) An automated early ischemic stroke detection system using CNN deep learning algorithm. In: 2017 IEEE 8th International Conference on Awareness Science and Technology (iCAST), pp 368–372. <https://doi.org/10.1109/ICAawST.2017.8256481>
41. Liu L, Chen S, Zhang F, Wu FX, Pan Y, Wang J (2019) Deep convolutional neural network for automatically segmenting acute ischemic stroke lesion in multi-modality MRI. *Neural Comput Appl* 32:6545–6558. <https://doi.org/10.1007/s00521-019-04096-x>
42. Yu J, Park S, Kwon SH, Ho CMB, Pyo CS, Lee H (2020) AI-based stroke disease prediction system using real-time electromyography signals. *Appl Sci* 10(19):6791. <https://doi.org/10.3390/app10196791>
43. Chantamit P, Goyal M (2018) Long short-term memory recurrent neural network for stroke prediction. *Mach Learn Data Min Pattern Recognit LNCS 10934*:312–323. https://doi.org/10.1007/978-3-319-96136-1_25
44. WHO (2004) ICD-10: international statistical classification of disease and related health tenth revision, vol 2, Geneva. https://www.who.int/classifications/icd/ICD-10_2nd_ed_volume2.pdf
45. Choi YA, Park SJ, Jun JA, Pyo CS, Cho KH, Lee HS (2021) Deep learning-based stroke disease prediction system using real-time bio signals. *Sensors* 21(13):4269. <https://doi.org/10.3390/s21134269>
46. Rocha-Ibarra E, Oros-Flores M-I et al (2021) Kinect validation of ergonomics in human pick and place activities through lateral automatic posture detection. *IEEE Access* 9:109067–109079. <https://doi.org/10.1109/ACCESS.2021.3101964>
47. Guillaume JF, Laroche D, Babault N (2019) Kinematics and electromyographic activity of horse riders during various cross-country jumps in equestrian. *Sports Biomech* 20(6):680–692. <https://doi.org/10.1080/14763141.2019.1597154>
48. Yu JH, Lee HS, Im YH, Kim MS (2010) Real-time classification of internet application traffic using a hierarchical multi-class SVM. *KSII Trans Internet Inf Syst* 4(5):859–876. <https://doi.org/10.3837/tiis.2010.10.00>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Jaehak Yu¹ · Sejin Park^{1,2} · Chee Meng Benjamin Ho³ · Soon-Hyun Kwon¹ · Kang-Hee cho⁴ · Yang Sun Lee⁵ 

Jaehak Yu
dbzzang@etri.re.kr

Sejin Park
sjpark@etri.re.kr; sjpark@kriss.re.kr

Chee Meng Benjamin Ho
hocmb@sewon3h.com

Soon-Hyun Kwon
kwonshzzang@etri.re.kr

Kang-Hee cho
khcho@cnu.ac.kr

¹ Department of KSB (Knowledge-Converged Super Brain) Convergence Research, Electronics and Telecommunications Research Institute (ETRI), Daejeon 34129, Republic of Korea

- ² Center for Medical Metrology, Korea Research Institute of Standards and Science (KRISS), Daejeon 34113, Republic of Korea
- ³ AI Research Group, Sewon Intelligence, Ltd., Seoul 04512, Republic of Korea
- ⁴ Department Department of Rehabilitation Medicine, Chungnam National University College of Medicine, Daejeon 35015, Republic of Korea
- ⁵ Division of Computer Engineering, Hanshin University, Gyeonggi-do 18101, Republic of Korea