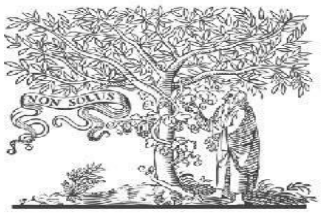
International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

COPY RIGHT



ELSEVIER
SSRN

Link : https://ijiemr.org/downloads/Volume-12/Issue-05

10.48047/IJIEMR/V12/ISSUE05/24

Title **AI-Based Stroke Disease Prediction System Using ECG and PPG Bio-Signals**

Volume 12, Issue 05, Pages: 234-245

Paper Authors

G. Priyanaka, P. Sushanth Reddy, S. Saicharan Reddy

USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per UGC Guidelines We Are Providing A Electronic Bar Code

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

AI-Based Stroke Disease Prediction System Using ECG and PPG Bio-Signals

1. **G. Priyanaka**, Assistant Professor, MTech(CSE), Department of CSE, (Vijay Rural Engineering College(VREC)) piryanakavrec2@gmail.com

2. **P. Sushanth Reddy**, BTech, Department of CSE, (Vijay Rural Engineering College(VREC)) sushanthreddy281@gmail.com

3. **S. Saicharan Reddy**, BTech , Department of CSE, (Vijay Rural Engineering College(VREC)) Saicharanrdy24@gmail.com

ABSTRACT: Because stroke illness often results in death or major disability, aggressive primary prevention and early diagnosis of prognostic signs are critical. Stroke illnesses are classified as ischemic or hemorrhagic, and they should be treated as soon as possible with thrombolytic or coagulant therapy. First, it is critical to notice the precursor symptoms of stroke in real time, which vary by person, and to give professional treatment by a medical institution within the appropriate treatment window. Prior research, however, has concentrated on creating acute therapy or clinical treatment recommendations after the onset of stroke rather than identifying predictive indicators of stroke. Image analysis, such as magnetic resonance imaging (MRI) or computed tomography (CT), has been utilised extensively in recent research to identify and predict prognostic signs in stroke patients. These approaches are not only difficult to identify early in real-time, but they also have drawbacks in terms of extended test times and expensive testing costs. In this research, we present a machine learning-based method for predicting and semantically interpreting stroke prognostic symptoms in the elderly utilising multi-modal bio-signals of electrocardiogram (ECG) and photoplethysmography (PPG) recorded in real-time. We devised and deployed a stroke disease prediction system with an ensemble structure that integrates CNN and LSTM to predict stroke illness in real-time while walking. The suggested system takes into account the ease of wearing bio-signal sensors for the elderly, and bio-signals were captured while walking at a sample rate of 1,000Hz per second from the three electrodes of the ECG and the index finger for PPG. Real-time prediction of elderly stroke patients demonstrated good prediction accuracy and performance.

Keywords *–Deep learning, machine learning, electrocardiogram (ECG), photo plethysmography (PPG), multi-modal bio-*

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

signal, real-time stroke prediction, stroke disease analysis.

1. INTRODUCTION

Stroke is classified as either ischemic (a blood vessel delivering blood to a portion of the brain is blocked) or hemorrhagic (a blood vessel bursts). It is a neurological symptom and condition caused by injury to a specific region of the brain. Stroke is regarded as one of the most dangerous illnesses in contemporary civilization since it may result in mortality in extreme instances, as well as physical and mental impairments such as hemiparesis, speech impairment (aphasia), ataxia, vision impairment, consciousness impairment, and dementia. According to the World Health Organization's (WHO) 2019 Causes of Death Report, issued in December 2020, the top ten causes of death accounted for 55% of all documented deaths in 2019. (about 55.4 million people). Six million of them died as a result of cerebrovascular illness, which was believed to be the second greatest cause of mortality. According to the United Nations (UN), a nation is categorised as an ageing society if the percentage of its people aged 65 and over in the total population is 7% or higher, an aged society if the proportion is 14% or higher, and a super-aged society if the proportion is more than 20%. As a result, the social difficulties of the ageing society are becoming visible enough that the ageing society may be segmented. Furthermore, according to an ageing study report published by Moody's, an international credit rating organisation, as of 2013, Japan, Germany, Italy, and other countries have become super-aged societies, with a percentage of senior people above 20%. According to reports, 34 nations would have become super-aged civilizations by 2030. The prognosis and health condition of stroke patients varies significantly depending on their age and place of onset. According to a prior stroke research, elderly persons 65 and older accounted for more than 66% of overall stroke incidence. Aside from these societal challenges, the incidence and mortality of stroke are likely to become significant social and economic issues.

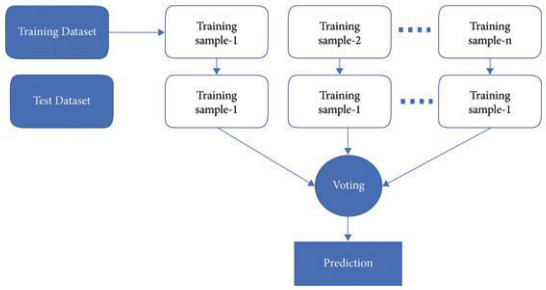


Fig.1: Example figure

The neurological diagnostic and severity information provided by a medical team are used to determine the diagnosis of stroke, which is represented by cerebrovascular illness [6], [10]-[12]. Brain MRI and CT are the most often used

techniques for neurological diagnosis in stroke diagnosis, but other research have shown that bio-signals such as brain waves, muscle, and ECG may also be used to diagnose and prevent stroke disorders [13]-[15]. Furthermore, ultrasonography, echocardiography, cerebral angiography, and single photon emission computed tomography (SPECT) are being utilised to diagnose the most prevalent causes of stroke. Imaging methods like as CT and MRI have recently been popular for stroke detection, but they still have drawbacks in the examination and diagnostic process due to hypersensitive responses to contrast agent medication penetration, radiation exposure, and claustrophobia in a confined environment. Because test findings may include inaccuracies, medical staff judgement based on professional medical knowledge and factual evidence is regarded as critical.

2. LITERATURE REVIEW

Autonomic dysfunction in acute ischemic stroke: An underexplored therapeutic area?:

In patients with acute ischemic stroke, impaired autonomic function, defined by a preponderance of sympathetic activity, is prevalent. Methods for measuring autonomic dysfunction in stroke patients are described in this review. It describes a possible link between ischemic stroke-associated autonomic dysfunction and parameters linked to worse outcomes, such as cardiac problems, blood pressure variability alterations, hyperglycemia, immunological depression, sleep disturbed breathing, thrombotic consequences, and malignant edoema. The insular cortex has been suspected of playing a key role in generating sympathovagal imbalance, but its specific involvement, as well as that of other brain areas, remains unknown. Although sympathetic overactivity seems to be a poor predictive feature in individuals with ischemic stroke, it needs to be explored if treatment techniques that lower sympathetic activity or promote parasympathetic activity may improve prognosis.

Diagnosis and management of acute ischemic stroke: Speed is critical

Stroke is the world's second biggest cause of mortality. 1 The estimated 62 000 strokes that occur in Canada each year impact persons of all ages, from newborns to the elderly, with incidence rates increasing with age. By the age of 80, the lifetime risk of overt stroke is anticipated to be one in four, but the lifetime risk of silent or covert stroke is expected to be closer to 100%. Stroke affects men and women equally and imposes significant social and economic consequences on society, with direct expenses

exceeding $3 billion in Canada each year. 2 Many parallels exist between acute stroke and acute coronary syndrome. We examine the diagnosis and treatment of acute ischemic stroke and compare it to that of acute coronary syndrome to show how fast alleviation of arterial blockage and restoration of normal blood flow may save lives and avoid disability. This narrative review is based on an analysis of relevant clinical studies (Box 1).

Long sleep duration and risk of ischemic stroke and hemorrhagic stroke: The Kailuan prospective study:

The goal of this research was to look at the link between sleep duration and ischemic and hemorrhagic stroke in a community-based sample. The present study comprised 95,023 Chinese adults who were stroke-free at the time of the baseline survey (2006–2007). Cox proportional hazards models were used to compute stroke hazard ratios (HRs) and confidence intervals (CIs) based on sleep duration. After an average of 7.9 years of follow-up, 3,135 patients had a stroke (2,504 ischemic stroke and 631 hemorrhagic stroke). Individuals reporting more than 8 hours had a complete adjusted hazard ratio (95% CI) of total stroke of 1.29 (with 6-8 hours of night sleep included for the reference group) (1.01–1.64). The elderly had a stronger link between lengthy sleep duration and total stroke (HR, 1.47; 95% CI, 1.05-2.07). Only women who reported sleeping more than 8 hours per night were related with hemorrhagic stroke (HR, 3.58; 95% CI, 1.28-10.06) as compared to individuals who reported sleeping 6-8 hours per night. Long sleep duration may be a possible predictor/marker for total stroke, according to this research, particularly in the elderly. Long sleep duration was shown to increase the risk of hemorrhagic stroke solely in women.

An elderly health monitoring system using machine learning and in-depth analysis techniques on the NIH stroke scale:

With the fast transition to an ageing society and the rising interest in healthcare, illness prediction and management through different healthcare equipment and services has received a lot of attention recently. Stroke, as indicated by cerebrovascular illness, is a particularly deadly condition, with high rates of mortality or mental and physical aftereffects in adults and the elderly. The consequences of such stroke disorders are quite harmful since they impede social and economic activity. In this study, we present a novel approach for predicting and analysing stroke severity in older people over the age of 65 using the National Institutes of Health Stroke Scale (NIHSS). Furthermore, we use the C4.5 decision tree algorithm, which is a

framework for prediction and analysis of machine learning approaches. C4.5 decision trees are machine learning methods that give more detailed execution mechanism rules and semantic interpretation analyses. Finally, the C4.5 decision tree approach is validated in this study for classifying and predicting stroke severity, as well as obtaining further NIHSS feature reduction benefits. As a result, during real system operation, the suggested model employs just 13 of the 18 stroke scale parameters, including age, to offer quicker and more accurate service assistance. The method does this by lowering the patient NIH stroke scale measuring time and making the procedure more efficient, with an overall accuracy of 91.11% utilising the C4.5 decision tree algorithm.

Effective anti-aging strategies in an era of super-aging:

The world's cultures in the twenty-first century have encountered issues due to an ageing population, since fertility rates have declined drastically and medical discoveries have increased the average human life duration. The elderly aged 65 and above account for at least 20% of the Korean population, classifying the nation as a super-aging society as defined by the United Nations. The number of old women outnumbers that of senior men, and women outlive males. This research aims to recommend practical techniques for using isoflavones, compounds chemically comparable to the female hormone oestrogen, and to seek for successful anti-aging tactics employing this substance for women to be prepared to enter the old stage in excellent health.

3. METHODOLOGY

Because stroke illness often results in death or major disability, aggressive primary prevention and early diagnosis of prognostic signs are critical. Stroke illnesses are classified as ischemic or hemorrhagic, and they should be treated as soon as possible with thrombolytic or coagulant therapy. First, it is critical to notice the precursor symptoms of stroke in real time, which vary by person, and to give professional treatment by a medical institution within the appropriate treatment window. Prior research, however, has concentrated on creating acute therapy or clinical treatment recommendations after the onset of stroke rather than identifying predictive indicators of stroke. Image analysis, such as magnetic resonance imaging (MRI) or computed tomography (CT), has been utilised extensively in recent research to identify and predict prognostic signs in stroke patients. These approaches are not only difficult to identify early in real-time, but they also have drawbacks in terms of extended test times and expensive

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

testing costs.

Disadvantages:

1. These approaches are not only difficult to identify early in real-time, but they also have drawbacks in terms of extended test times and expensive testing costs.

In this research, we present a machine learning-based method for predicting and semantically interpreting stroke prognostic symptoms in the elderly utilising multi-modal bio-signals of electrocardiogram (ECG) and photoplethysmography (PPG) recorded in real-time. We devised and deployed a stroke disease prediction system with an ensemble structure that integrates CNN and LSTM to predict stroke illness in real-time while walking. The suggested system takes into account the ease of wearing bio-signal sensors for the elderly, and bio-signals were captured while walking at a sample rate of 1,000Hz per second from the three electrodes of the ECG and the index finger for PPG.

Advantages:

1. The real-time prediction of senior stroke patients shown good prediction accuracy and performance.
2. It was experimentally shown that the prognostic symptoms of stroke patients may be predicted with more than 90% accuracy using just ECG and PPG obtained while walking.

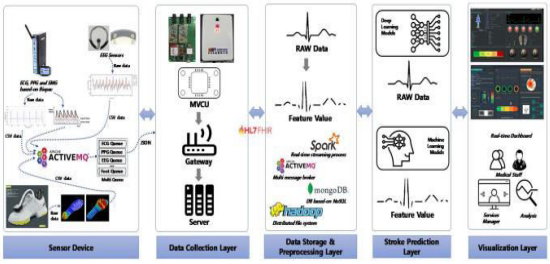


Fig.2: System architecture

MODULES:

To carry out the aforementioned project, we created the modules listed below.

- Data exploration: we will put data into the system using this module.
- Processing: we will read data for processing using this module.
- Data splitting into train and test: Using this module, data will be split into train and test.
- Model construction: Random Forest, Decision Tree, Naive Bayes, AdaBoost Classifier, Logistic Regression, MLP-ANN, Support Vector Machine, Voting

Classifier, BF Tree, BayesNet, CNN, CNN+LSTM, LSTM, BiLSTM, and CNN+BiLSTM. Calculated algorithm accuracy.

- User registration and login: Using this module will result in registration and login.
- Using this module will provide input for prediction.
- Prediction: final predicted shown

4. IMPLEMENTATION

ALGORITHMS:

Random Forest: A Supervised Machine Learning Algorithm that is commonly utilised in Classification and Regression applications. It constructs decision trees from several samples and uses their majority vote for classification and average for regression.

Decision Tree: Decision trees use numerous methods to determine whether or not to divide a node into two or more sub-nodes. The development of sub-nodes promotes the homogeneity of the sub-nodes that arise. In other words, the purity of the node rises in relation to the target variable.

Naive Bayes: A probabilistic classifier, the Naive Bayes classification technique. It is based on probability models with high independence assumptions. The independence assumptions often have little effect on reality. As a result, they are seen as naïve.

AdaBoost Classifier: An AdaBoost classifier is a meta-estimator that starts by fitting a classifier on the original dataset and then fits additional copies of the classifier on the same dataset with the weights of incorrectly classified instances adjusted so that subsequent classifiers focus more on difficult cases.

Logistic Regression: Logistic regression is a statistical analytic approach that uses past observations of a data set to predict a binary result, such as yes or no. A logistic regression model forecasts a dependent variable by examining the connection between one or more existing independent variables.

MLP-ANN: A multilayer perceptron (MLP) is a kind of feedforward artificial neural network that is completely linked (ANN). The word MLP is used ambiguously, sometimes to refer to any feedforward ANN, and sometimes to networks built of many layers of perceptrons (with threshold activation); see Terminology. Multilayer perceptrons are commonly referred to

as "vanilla" neural networks, particularly when just one hidden layer is present.

SVM: Support Vector Machine (SVM) is a supervised machine learning technique that may be used for both classification and regression. Though we call them regression issues, they are best suited for categorization. The SVM algorithm's goal is to identify a hyperplane in an N-dimensional space that clearly classifies the input points.

Voting classifier: A voting classifier is a machine learning estimator that trains numerous base models or estimators and predicts based on the results of each base estimator. Aggregating criteria may be coupled voting decisions for each estimator output.

BF Tree: The breadth-first search (BFS) method searches a tree or graph data structure for nodes that fulfil a set of criteria. It explores all nodes at the current depth level before going on to nodes at the next depth level, starting at the root of the tree or graph.

Bayesian Net:Bayesian networks are a sort of Probabilistic Graphical Model that may be used to create models based on data and/or expert opinion. They may be used for a variety of activities like as prediction, anomaly detection, diagnostics, automated insight, reasoning, time series prediction, and decision making in the face of ambiguity.

CNN: A CNN is a kind of network architecture for deep learning algorithms that is primarily utilised for image recognition and pixel data processing jobs. There are different forms of neural networks in deep learning, but CNNs are the network design of choice for identifying and recognising things.

LSTM: Long short-term memory (LSTM) is a kind of artificial neural network used in artificial intelligence and deep learning. Unlike traditional feedforward neural networks, LSTM has feedback connections. A recurrent neural network (RNN) of this kind may analyse not just single data points (such as photos), but also complete data sequences (such as speech or video).

BiLSTM: BiLSTM stands for Bidirectional Long Short-Term Memory (BiLSTM) In general, LSTM ignores future information in time series processing. BiLSTM processes series data in forward and reverse directions on the basis of LSTM, linking the two hidden layers.

5. EXPERIMENTAL RESULTS

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

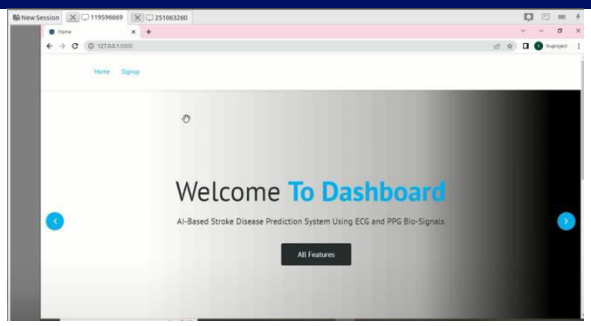


Fig.3: Home screen

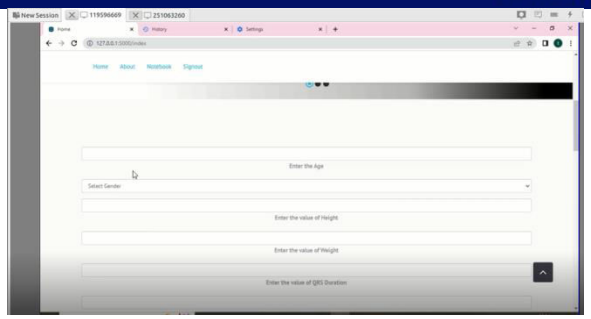


Fig.6: Main screen

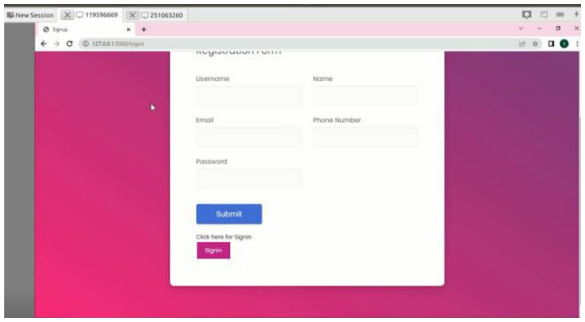


Fig.4: User registration

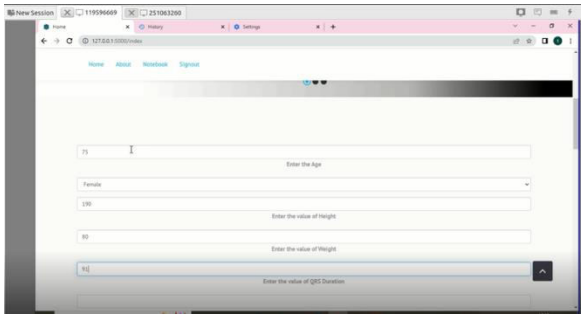


Fig.7: User input

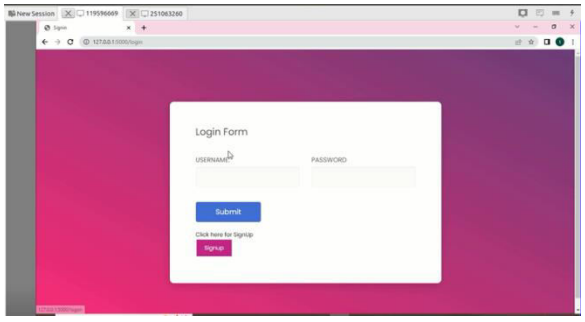


Fig.5: user login

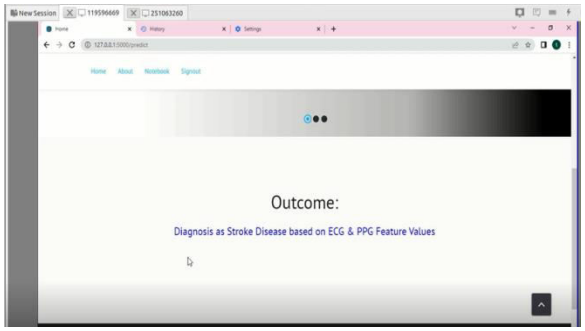


Fig.8: Prediction result

6. CONCLUSION

In this research, we offer a system that enables semantic analysis of illnesses in the elderly by using numerous biological signals of ECG and PPG recorded while walking in the elderly's

everyday lives. The proposed system captures numerous ECG and PPG biosignals in real time and may identify and predict prognostic signs of stroke illness in the elderly. Using numerous biosignal data, a machine learning-based prediction model research was undertaken, which included separating the signal waveform into particular portions, and reasonably accurate prediction results and semantic interpretations were produced using this model. In this work, it was experimentally demonstrated that utilising the suggested features, the prognostic symptoms of stroke patients may be reliably predicted by more than 90% based purely on ECG and PPG obtained while walking. To summarise the experimental and verification findings, we demonstrated that by splitting stroke and general elderly into 10-folder CV datasets, we can properly forecast 91.56% C4.5 Decision Tree, 97.51% RandomForest, and 99.15% CNN-LSTM models for deep learning. The method presented in this work has high academic value since it can reliably predict prognostic symptoms and the development of stroke by recording ECG and PPG at a reasonable cost and with minimal discomfort throughout everyday life. With a high repetition probability, various bio-signal data gathered in everyday life may give objective interpretation information to stroke patients or medical professionals. The trial findings demonstrated that this technology may be utilised for practical healthcare services such as reducing stroke aftereffects and preventing emergency situations via continuous monitoring. We will perform in-depth analyses and predicting experiments of stroke illness in the future by assessing numerous bio-signals such as EEG, EMG, foot pressure, and motion, as well as electronic medical records (EMRs) and MRI image data.

REFERENCES

[1] S. De Raedt, A. De Vos, and J. De Keyser, ‘‘Autonomic dysfunction in acute ischemic stroke: An underexplored therapeutic area?’’ J. Neurol. Sci., vol. 348, nos. 1–2, pp. 24–34, Jan. 2015.

[2] C. P. Warlow, ‘‘Epidemiology of stroke,’’ Lancet, vol. 352, pp. 1–4, Oct. 1998.

[3] K.-D. Seo, M. J. Kang, G. S. Kim, J. H. Lee, S. H. Suh, and K.-Y. Lee, ‘‘National trends in clinical outcomes of endovascular therapy for ischemic stroke in South Korea between 2008 and 2016,’’ J. Stroke, vol. 22, no. 3, pp. 412–415, Sep. 2020.

[4] T. D. Musuka, S. B. Wilton, M. Traboulsi, and M. D. Hill, ‘‘Diagnosis and management of acute ischemic stroke: Speed is critical,’’ Can. Med. Assoc. J., vol. 187, no. 12, pp. 887–893, Sep. 2015.

[5] Q. Song, X. Liu, W. Zhou, L. Wang, X. Zheng, X. Wang, and S. Wu, ''Long sleep duration and risk of ischemic stroke and hemorrhagic stroke: The Kailuan prospective study,'' Sci. Rep., vol. 6, no. 1, pp. 1–9, Sep. 2016.

[6] J. Yu, S. Park, H. Lee, C. S. Pyo, and Y. S. Lee, ''An elderly health monitoring system using machine learning and in-depth analysis techniques on the NIH stroke scale,'' Mathematics, vol. 8, no. 7, pp. 1–16, Jul. 2020.

[7] World Health Organization. The Top 10 Causes of Death. Accessed: Apr. 22, 2022. [Online]. Available: https://www.who.int/newsroom/fact-sheets/detail/the-top-10-causes-of-death

[8] UNDESA World Social Report 2020. Accessed: Apr. 22, 2022. [Online]. Available: https://www.un.org/development/desa/dspd/world-socialreport/2020-2.html

[9] S. Park, M. J. Yang, S. N. Ha, and J. S. Lee, ''Effective anti-aging strategies in an era of super-aging,'' J. Menopausal Med., vol. 20, no. 3, pp. 85–89, Dec. 2014.

[10] B. C. Meyer and P. D. Lyden, ''The modified national institutes of health stroke scale: Its time has come,'' Int. J. Stroke, vol. 4, no. 4, pp. 267–273, Aug. 2009.

[11] D. Schlegel, S. J. Kolb, J. M. Luciano, J. M. Tovar, B. L. Cucchiara, D. S. Liebeskind, and S. E. Kasner, ''Utility of the NIH stroke scale as a predictor of hospital disposition,'' Stroke, vol. 34, no. 1, pp. 134–137, Jan. 2003.

[12] J. Yu, D. Kim, H. Park, S. Chon, K. H. Cho, S. Kim, S. Yu, S. Park, and S. Hong, ''Semantic analysis of NIH stroke scale using machine learning techniques,'' in Proc. PlatCon, Jeju, South Korea, Jan. 2019, pp. 82–86.

[13] J. Yu, S. Park, S. H. Kwon, C. M. B. Ho, C. S. Pyo, and H. Lee, ''AIbased stroke disease prediction system using real-time electromyography signals,'' Appl. Sci., vol. 10, no. 19, pp. 1–19, Sep. 2020.

[14] Y. A. Choi, S. Park, J. A. Jun, C. M. B. Ho, C. S. Pyo, H. Lee, and J. Yu, ''Machine-learning-based elderly stroke monitoring system using electroencephalography vital signals,'' Appl. Sci., vol. 11, no. 4, pp. 1–18, Feb. 2021.

[15] M. Lee, J. Ryu, and D. Kim, ''Automated epileptic seizure waveform detection method based on the feature of the mean slope of wavelet coefficient counts using a hidden Markov model and EEG signals,'' ETRI J., vol. 42, no. 2, pp. 217–229, Apr. 2020.