ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

A COMPARATIVE WORK OF INCREMENTAL LEARNING AND ENSEMBLE LEARNING FOR BRAINPRINT IDENTIFICATION

Siaw-Hong Liew[1], Yun-Huoy Choo[2], Yin Fen Low[3] and Fadilla 'Atyka Nor Rashid[4]
[1]Faculty of Computer Science and Information Technology Universiti Malaysia Sarawak (UNIMAS), Kota Samarahan, Sarawak, Malaysia
[2]Faculty of Information and Communication Technology, Malaysia
[3]Faculty of Electronics and Computer Engineering, Universiti Teknikal Malaysia Melaka (UTeM), Durian Tunggal, Melaka, Malaysia
[4]Faculty of Information Science and Technology Universiti Kebangsaan Malaysia (UKM), Bangi, Selangor, Malaysia
E-Mail: shliew@unimas.my

ABSTRACT

Electroencephalogram (EEG) signals are nonstationary and vary across time. The static learning model requires large training data to ensure sufficient knowledge acquisition to build a robust model. However, it is very challenging to achieve complete concept learning due to the behavioural changes in model learning. This issue is particularly critical in brainprint identification, where data acquisition in a short time cannot ensure sufficient training data for comprehensive model learning. Thus, dynamic learning, i.e., incremental learning and ensemble learning, presents a better solution for encapsulating EEG signal changes and variations. Both incremental and ensemble learning follow different approaches to manag the concept learning. Incremental learning merges new variations of EEG signals into the existing learning model over time, while ensemble learning uses multiple models for prediction. Nevertheless, limited research works were reported on comparing these two learning methods to prove the efficiency in handling nonstationary data for brainprint identification. Thus, this paper aims to compare incremental learning and ensemble learning for brainprint identification modelling. Incremental Fuzzy-Rough nearest Neighbour (IncFRNN) and Random Forest are selected to represent incremental learning and ensemble learning, respectively. Accuracy, area under the ROC curve (AUC) and F-measure were used to evaluate the classification performance. The experimental results proved that incremental learning outperformed ensemble learning when the training data were limited. The classification results of IncFRNN model were recorded at 0.9160, 0.9827 and 0.9169 while the Random Forest model only yielded 0.8113, 0.9709, and 0.9169 in accuracy, AUC, and F-measure, respectively. The ongoing learning process in incremental learning helps to capture the new changes in EEG signals and improve the classification performance.

Keywords: incremental learning, ensemble learning, electroencephalogram (EEG) signals, brainprint identification.

Manuscript Received 17 February 2023; Revised 21 July 2023; Published 13 August 2023

1. INTRODUCTION

With the advancement of non-invasive Brain-Computer Interface (BCI), Electroencephalogram (EEG) signals have grown into a popular topic in a variety of fields of study due to their high time resolution, low cost and portability [1]. EEG signals are being used as a biometric trait for authentication and identification and have been highlighted recently. Brainprint identification uses EEG signals to identify an individual among a group of persons who are being evaluated (one-to-N matching). In recent years, brainprint identification is catching researchers' attention [2]-[6] corresponding to the rise of security. EEG signal is private and provides uniqueness. Everyone has diverse mental reactions towards different stimuli. EEG signal is outstanding because it is covered in the brain and physically invisible. Other biometric traits, for example, fingerprint, palm print, or face, are effortlessly reachable by physical sensors on the body surface [7]. These are simply violated and inclined to be imitations by third parties. For example, an artificial fingerprint can be made from silicone, gel, or rubber. However, the EEG signal is difficult to replicate at different locations and at different times.

The main challenges in EEG signals classification are the low signal-to-noise ratio (SNR), and nonstationary characteristics within or between persons, where the EEG signals of the same person vary across time. Correspondingly, our brain is easily affected by emotions, moods, feelings, and other surrounding environmental factors [8]. As a result, a classifier might be trained on a limited amount of training data [9]. Generally speaking, a static learning model requires a large or full amount of training data to ensure sufficient knowledge acquisition to build a robust learning model. Static or traditional approaches will become useless for learning new information. The issue will grow difficult if the previously seen data is no longer available when the new data arrives. It is due to the static or traditional approach required to combine the old and new data to retrain the classifier, which is very impractical. In brainprint identification modelling, it is very challenging to achieve complete concept learning due to emotional and behavioural changes in the model learning. This issue is particularly critical in the case of brainprint identification, where data acquisition in a short time cannot ensure sufficient training data for comprehensive model learning [9]. To address the above-mentioned issues, dynamic learning, i.e., incremental learning, and ensemble learning, presents a better solution in encapsulating the changes and variations in EEG signals. However, limited research works were