Abstract ID: 005

SURFACE ELECTROMYOGRAPHY-BASED RECOGNITION OF SINHALA HANDWRITTEN CHARACTERS USING MACHINE LEARNING

B. M. Seneviratne[a], B. N. S Lankasena[a*], and I. Jayarathne[a]

[a]*Department of Information and Communication Technology, Faculty of Technology, University of Sri Jayewardenepura, Homagama 10206, Sri Lanka.*

[*]*nalaka@sjp.ac.lk*

Abstract

Surface electromyography (sEMG) represents a well-established technique for capturing muscle activation signals across numerous applications. However, sEMG-based handwritten character recognition for underrepresented languages such as Sinhala remains largely unexplored in the literature. This research addresses the existing knowledge gap by proposing a comprehensive framework for recognizing six Sinhala handwritten characters selected based on stroke complexity: ර (r) and ට (ṭa) from ascending letters group, ෘ (a) and ු (u) from descending letters group, and ක (ka) and ග (ga) from middle letters group. To accomplish this objective, a novel dataset was collected using a consumer-grade OpenBCI sEMG acquisition system. Four electrodes were strategically positioned across multiple muscle groups, with particular emphasis on the extensor digitorum muscle due to its role in finger control during handwriting. Eight participants wrote six selected Sinhala characters 50 times each at natural speed while sEMG signals were recorded, generating 300 samples per character. Signal preprocessing was performed using high-pass, low-pass, and notch filtering techniques to remove noise components from the collected dataset. Mean absolute value, root mean square, variance, zero crossing, slope sign change, waveform length, and Willison amplitude features were extracted from the sEMG signals. Several classification algorithms including Random Forest, K-Nearest Neighbors, Naive Bayes, and Support Vector Machine were employed to train predictive models using the derived features, and their effectiveness was assessed using conventional performance metrics. The Random Forest algorithm demonstrated optimal results, attaining a recognition accuracy of 98.96% on the experimental dataset, confirming the practical applicability and robustness of the sEMG- based Sinhala handwriting recognition framework. The study indicates the potential of using machine learning technology with selected sEMG signal characteristics for the classification of Sinhala characters from comparable neuromuscular activation patterns. This work establishes a foundation for developing sEMG-based assistive writing technologies for Sinhala script, benefiting individuals with physical disabilities.

Keywords*: Character Recognition, Machine Learning, Random Forest, Sinhala Handwriting, Surface Electromyography*

Financial assistance from University of Sri Jayewardenepura (*Grant No. ASP/01/RE/TEC/2022/79*) *is acknowledged.*