# ABSTRACT

The electroencephalogram is a conventional tool for diagnosing brain-related illnesses. It is most commonly used for diagnosing epilepsy patients by identifying underlying epileptiform activities traced in the electrograph. Methods for collecting these non-invasive EEG recordings can introduce various types of artifacts, noise, and interference yielding poor quality signals. Some of these undesired signal events are widespread and can completely mask the underlying brain signals. Automated EEG event detection systems, such as seizure detection software, suffer heavily from such poor-quality signals that hinder isolation of the underlying signal components. The focus of this study is to automatically reject a very common type of artifact known as “muscle artifact” using a blind source separation (BSS) method.

Muscle artifacts may appear in the recordings as a result of muscle movements, paroxysmal fast activities, and in most cases agitation and muscle tension. Although clinicians and technologists make their best effort to reduce such noisy events, occasional bursts of such artifacts are inevitable. Any type of artifact in an EEG signal conveys useful diagnostic information but can also confuse the interpreter by masking actual underlying brain activity. This especially hurts the automated seizure detection algorithms and introduces a high false alarm (FA) rate because the muscle artifacts could occur with or without seizure events. Additionally, these systems operate without any knowledge of the clinical settings. Therefore, specificity of such algorithms tends to be quite low. Despite the presence of muscle artifacts which masks the events of interest, such as ictal discharges, the higher dimensional feature space and deep learning system’s posterior probabilities show clear isolation between the two classes. This indicates that both events possess different spectral and/or statistical properties which are possible to isolate.

Traditional digital filters are able to filter high frequency muscle artifacts but they tend to change the shape of the signals which plays an important role during the identification of the epileptiform features. Instead we use a statistical approach known as “Canonical correlation analysis” (CCA) with the aid of linear prediction (LP) model to estimate the underlying frequency components.

We use popular “correlation technique” for the LP coefficient estimation. The correlation technique uses Levinson-Durbin algorithm which ensures the estimated model is minimum phase and hence a stable system. For optimal performance, such LP model prefers preemphasized signals which essentially boosts the high frequency components of the signal. Fortunately the signals obfuscated by the muscle artifacts show similar spectral properties to the preemphasis signal spectrums. In fact these signals, being noisy, help the LP model detect the underlying frequency components more effectively! The lag values associated with the estimated underlying frequency components are used to denoise the signals via CCA. The correlation between linear combination of the lagged version of the signals and the original noisy signal is maximized which gives the contribution of the each lag signals to the original signal. Reconstructed signals via this method preserves the original shape of the signals and filters out the noise which provides a significant value during the visual interpretation and event classification.

Features collected on the denoised signals by our approach shall improve the performance of any seizure detection algorithms and serve as an effective visualization tool. We expect our algorithm to outperform a few traditionally used BSS algorithms and improve the sensitivity of the seizure detection neural network algorithms by at least 5% with ~6 FAs per 24 hours.