An LSTM-CNN Enabled Fine-Grained Human-Machine Interface Using sEMG Trained with a Natural Typing Task

Michael Crouch, Mingde Zheng, Michael S. Eggleston
Augmented Human Devices Group, Nokia Bell Laboratories, New Providence, NJ, 07974 United States

Introduction: The advent of deep learning presents new opportunities for human-machine interfaces (HMI), especially in creating prosthetic devices and virtual interfaces that offer fine-grained control. Surface Electromyography (sEMG) has previously been shown as a promising non-invasive bio-interface **[**1**]**, with successful demonstration of coarse hand gesture recognition when paired with state-of-the-art machine learning (ML) models. These demonstrations, however, are constrained in their ability to detect fine-grained motor function by the limited availability of large-scale annotated datasets which, by comparison, in fields such as image classification often include millions of annotated data points. To overcome the inherent difficulty of generating annotated sEMG data, we devise a keyboard typing task that allows for the rapid acquisition of tens of thousands of labeled datapoints. We demonstrate in the subsequently trained ML model that this large dataset enables highly accurate reconstruction of the fine-grained finger movements involved in typing, with accuracy >81% achieved from sEMG signal alone. Such a system demonstrates the potential for sEMG as a non-invasive HMI for high-resolution control of prosthetic or virtual objects.

Materials and Methods: To study muscle activation under small, controlled finger movements, we designed a typing task, training a neural network to recognize English text solely from forearm muscle activations. A large, novel dataset of sEMG signals was collected, and automatically annotated with the approximately 89,000 keystrokes typed over the course of recording. Recognition was performed in a real-time "streaming" manner, and with approximately 30ms latency. Performance was measured by training the model 10 times and evaluated on a 10% holdout set. Typing averaged 60-70 wpm throughout the task. Data was collected using a state-of-the-art surface biopotential acquisition system made by iWorx, measuring 16 channels per arm at 2 kHz sampling rate on a 25mV scale. The neural network combined convolutional layers, used to extract features from the timeseries data, with LSTM layers, used to analyze the evolution of those features over time. Model training experiments were also performed using variously sized subsets of the full training data, to test the adequacy of the training data.

Results and Discussion: Prior to data training, all 32 channels of neuromuscular signals were verified using iWorx's prescreening visualizer, thereby allowing the capture of subtle myoelectric activations as shown in Figure 1a. The neural network as illustrated in Figure 1b achieved an 81.3% recognition of typed text on a 32-letter keyboard, before any spelling correction or language model was applied. Experiments showed that model performance increased steadily as the amount of training data increased (Figure 1c).

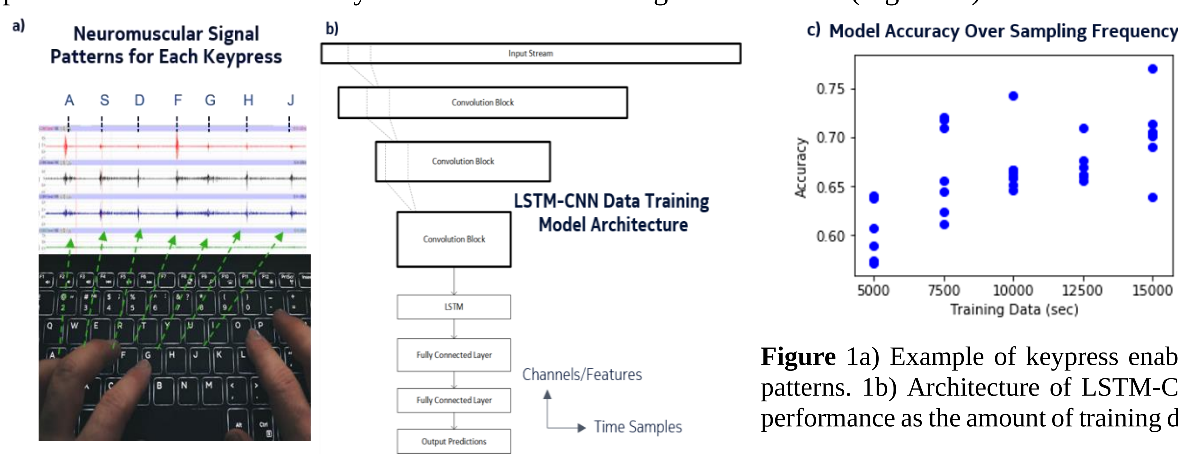

Figure 1a) Example of keypress enabled myoelectric patterns. 1b) Architecture of LSTM-CNN. 1c) Model performance as the amount of training data is increased.

Conclusions: This is, to our knowledge, the first system demonstrating real-time recognition of massive fine-grained finger gestures, and the first showing the sufficiency of surface EMG for a complex natural dexterity task. Ongoing efforts will focus on achieving high recognition accuracy across multiple subjects by implementing a novel transfer learning-based ML model. Our novel system for using keyboard typing as a method of collecting labeled gesture training data was integral to our success in capturing fine-grained gestures.

Reference: [1] Farina, D. et al. The extraction of neural information from the surface EMG for the control of upper-limb prostheses: emerging avenues and challenges. IEEE Trans Neural Syst Rehabil Eng 22, 797–809 (2014).