

Finger gesture recognition with smart skin technology and deep learning

Liron Ben-Ari^{1,†}, Adi Ben-Ari^{1,*}, Cheni Hermon² and Yael Hanein^{1,2,3,4}

¹ School of Electrical Engineering, Tel Aviv University, Tel Aviv, Israel

² X-trodes, Herzlia, Israel

³ Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel

⁴ Tel Aviv University Center for Nanoscience and Nanotechnology, Tel Aviv University, Tel Aviv, Israel

E-mail: yaelha@tauex.tau.ac.il

Abstract

Finger gesture recognition was extensively studied in recent years for a wide range of human-machine interfaces. Surface electromyography (sEMG), in particular, is an attractive enabling technique in the realm of finger gesture recognition and both low and high-density sEMG were previously studied. Despite the clear potential, cumbersome electrode wiring and electronic instrumentation render contemporary sEMG-based finger gestures recognition to be performed under unnatural conditions. Recent developments in smart skin technology provide an opportunity to collect sEMG data in more natural conditions. Here we report on a novel approach based on soft 16 electrode array, miniature and wireless data acquisition unit and neural network analysis to achieve gesture recognition under natural conditions. Finger gesture recognition accuracy values as high as 94.5% were achieved for 8 gestures when the training and test data were from the same session. For the first time, high accuracy values are also reported for training and test data from different sessions, for three different hand positions. These results demonstrate an important step towards sEMG based gesture recognition in non-laboratory settings such as in gaming or Metaverse.

† L. Ben-Ari and A. Ben-Ari contributed equally to this work.

Introduction

Finger gesture recognition (FGR) is a widely studied domain in human-machine interfaces (HMI). Applications include virtual games, where finger gestures can be used instead of a joystick to achieve improved user experience, medical uses, where FGR can be used to help distinguish between normal and abnormal movements [1] or sign-language translation [2, 3], to name just few examples [4]. Several different approaches were explored in recent decades for FGR, including: Video analysis [5, 6], smart gloves [7, 8], smart bands [9] and surface electromyography (sEMG) [10, 11, 12, 13, 14, 15]. sEMG in particular is an attractive approach as it records the electrical activity of arm muscles located away from the fingers so finger movements are not restricted. Moreover, it does not necessitate visual pathway, allowing operation in dark environment or during movement. sEMG is also sensitive to applied force even without any apparent movement (isometric muscle activation).

Despite the great potential of sEMG, such measurements have a number of technical and computational challenges. First, under dynamic activity, motion artifacts are very common. Second, electrode position may vary from session to session or from subject to subject complicating the analysis. Owing to these challenges, sEMG based studies concerned with the detection of hand movements are currently performed in a controlled environment, with the hand held at a fixed position [12, 11, 13]. Also, most studies only report on intra-session classification [12, 13, 15], or inter-session classification with degraded performance [11, 16]. Furthermore, to mediate good electrical contact between the electrodes and the skin wet electrodes are commonly used [12, 11]. These electrodes severely limit the usability of the technology as they limit session duration and electrode number (and therefore separation capacity). Wires, cumbersome amplification and recording instrumentation and relatively large electrode arrays [11, 13] further limit the technology restricting it to clinical or laboratory use and mandate skilled personnel for electrode placement and system operation. Previous studies concerning the identification of finger gestures from sEMG signals were performed in a controlled environment, using relatively bulky electrode arrays that require the use of a conductive gel. Such measurements do not allow continuous real-time tracking, and are far from practical use in HMI applications.

In this investigation we demonstrate FGR under natural conditions for real life applications by addressing the following requirements: First, the system should be compatible with hours of use. Second, recognition of finger gestures should be achieved regardless of the general position of the hand. Third, system performance must be invariant to precise electrode placement in repeated sessions (removal of the wearable device and re-application on a later day).

To achieve these requirements, we used novel printed electrode arrays. The electrodes are printed on thin and soft substrates (Fig. 3(a)) and were studied previously

for various applications [17, 18]. The electrode arrays used in this study were designed specifically to capture arm muscle activity. Moreover, the arrays were designed with an internal ground for simple and quick placement. The thinness and elasticity of the arrays allows excellent mechanical coupling to the skin. In this study we also used, for the first time, a new miniature wearable sensing system that allows continuous sEMG measurement even during dynamic movement. Using such a small, convenient and non-invasive system is an important step towards hand gesture recognition in freely behaving humans. Finally, using deep learning we demonstrate FGR under different hand positions and invariance to precise electrode placement (in repeated sessions).

Materials and methods

Wireless sEMG system

The electrode arrays we tested in this study are based on a technology which was previously described in [18]. Briefly, carbon electrodes and silver traces are screen printed on a thin and soft PU film. A second double sided adhesive PU film is used for passivation and skin adhesive material. Data were recorded with a miniature wireless data acquisition unit (DAU, X-trodes Inc.) which was developed to allow electrophysiological measurements under natural conditions. The DAU supports up to 16 unipolar channels ($2 \mu\text{V}$ noise RMS, 0.5-700 Hz) with a sampling rate of 4000 S/s, 16 bit resolution and input range of $\pm 12.5 \text{ mV}$. A 620 mAh battery supports DAU operation for a duration of up to 16 hr. A Bluetooth (BT) module is used for continuous data transfer. The DAU is controlled by an Android application and the data are stored on a built in SD card and on the Cloud for further analysis. The DAU also includes a 3-axis inertial sensor, to measure the acceleration of the hand during the measurements.

Data collection

16 healthy subjects were recruited (aged 18-30, x female and y male). For the research, we used the data of 8 of them, mainly due to weak signal in people with hair-covered arms. Electrode arrays were placed on the region of the extensor digitorum muscle of the dominant hand. Muscle location was identified by applying strong abduction of the fingers. During the recording each subject sat or stood in front of a table (depending on the position of the hand being examined). An instructional video displayed on a computer was used to guide the subjects.

The experiment consisted of two steps: First, the hand was supported on a table followed by a second stage in which the hand was not supported. The protocol was structured as follows: First, subjects were shown a short video showing different hand gestures. Subjects were instructed to perform specific gestures both through voice and visual instructions presented on a computer screen (for 3 s) and then to stop, and

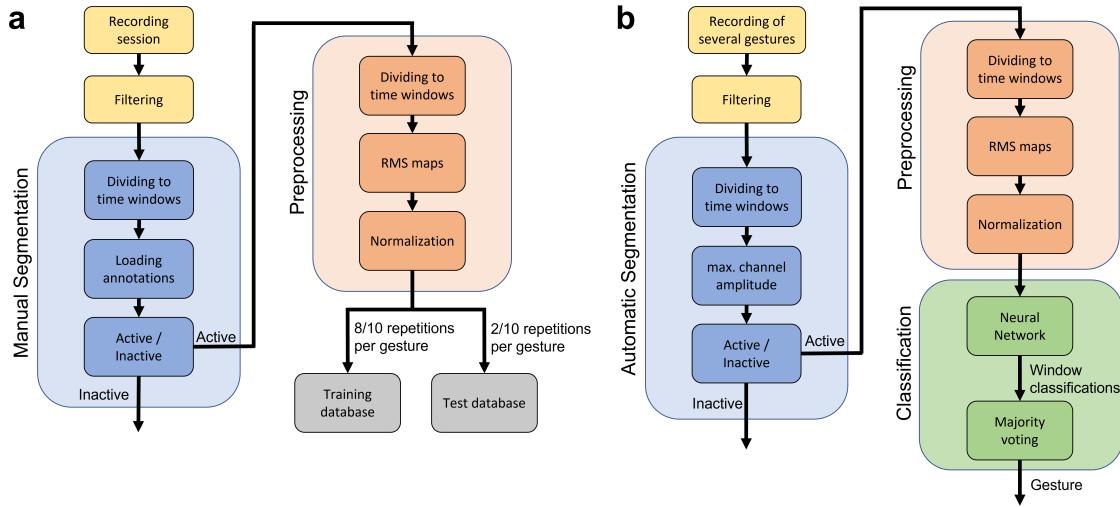


Figure 1. Data analysis flow chart. (a) Constructing training and test databases. (b) Full data analysis scheme, after the network was trained.

rest (another 3 s), after which subjects were instructed to repeat the gesture. The instructions continued until each gesture was performed 10 times. Altogether sEMG was recorded for 10 different finger gestures: Stretching two fingers, stretching three fingers, stretching all the fingers (abduction), making a fist, as well as six movements that represent letters in the Hebrew sign language: "Bet", "Gimel", "Het", "Tet", "Kaf", and "Nun". During the entire process of performing the movements, a representative of the research team monitored that the process is conducted as expected. In addition, a python code was used to send annotation to the android application to mark the timing in which the subject was instructed to start and finish each of the gestures, to assist the analysis stage.

All experiments were conducted on volunteers in accordance with relevant guidelines and regulations under approval from the Institutional Ethics Committee Review Board at Tel Aviv University. Informed consent forms were obtained from all subjects.

Data analysis

Data analysis flow is depicted in Fig. 1.

Filtering: Raw sEMG data were first filtered using a 50 Hz and a 100 Hz comb filters to reduce power-line interference. A 20 - 400 Hz 4th-order Butterworth bandpass filter (BPF) was applied to attenuate non-sEMG components.

Segmentation: Segmentation into time intervals for each gesture (denoted as *active time windows*), may be performed manually (i.e. using annotations made during the recording) or automatically. The automatic segmentation algorithm was implemented as follows: RMS smoothing with a T_0 time window was performed on each channel separately. The maximal value out of the obtained 16 values per window was derived (denoted as A). Each window w was assigned with an initial segmentation of *active* (1) if $A > threshold$, and *inactive* (0) otherwise. The binary time series obtained in the previous stage was filtered with a median filter with kernel size k . We used prior knowledge about the minimal active and inactive time intervals, T_{min}^{active} and $T_{min}^{inactive}$, and set all initially-segmented active windows shorter than T_{min}^{active} to inactive, and all initially-segmented inactive windows shorter than $T_{min}^{inactive}$ to active. We empirically set the parameters to: $T_0 = 50ms$, $threshold = 50\mu V$, $k = 19$, and $T_{min}^{active} = T_{min}^{inactive} = 1s$.

Classification Each active time window, identified in the segmentation stage, was divided into $200ms$ sub-windows. For each sub-window, the RMS value per channel was derived, resulting in 16 values for each sub-window. The obtained 16 values were arranged on a grid according to the spatial locations of the electrodes in the array, resulting in an activation map. The sequence of maps for each active time window was then fed into a classification algorithm. Several algorithmic solutions were explored as detailed below.

Convolutional Neural Network (CNN) In this approach, each map is fed separately into a CNN, which outputs a classification. By conducting majority voting between the classifications obtained for different sub-windows of the same active window, a final classification is obtained. The CNN architecture used in this work is depicted in Fig. 2. This NN consists of two convolutional layers, followed by three fully connected layers. Such architecture was favored as it has a relatively small number of parameters, making it suitable for our relatively small data set as well as the small size of the activation maps. Each layer, apart from the last fully connected one, was followed by a ReLU activation [19] and Batch Normalization [20]. The network was trained with 500-2000 epochs, with a learning rate between 0.0005 and 0.001 (with Adam optimizer [21]), weight decay of 0.0001 and dropout [22] of up to 0.3.

Recurrent Neural Network (RNN) In this approach, the sequence of maps for a certain action is treated as a time-series of dimension 16, which is fed into an LSTM-based RNN. The RNN outputs a classification. The architecture of the RNN consists of one LSTM layer of 12 units, followed by one fully connected layer. The small number of layers was again favored to match the relatively small data set. The network was trained with 1000-2000 epochs, using a learning rate between 0.005 and 0.01 (with Adam optimizer [21]), weight decay of 0.0001 and dropout [22] of up to 0.1

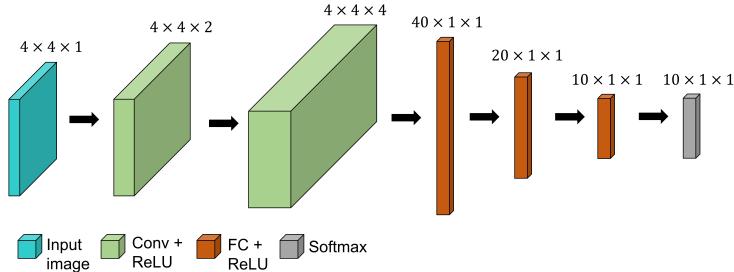


Figure 2. CNN architecture

We note that the different hyper-parameters relate to different tasks. For each task (see Table 1), hyper-parameters are fixed for all subjects.

Classical algorithms In this approach, the classification pipeline follows the same steps as in CNN classification, apart from the CNN which is replaced by either a K-Nearest Neighbor (KNN) classifier or a Multi-class Support Vector Machine (SVM) classifier. For KNN, we used $k = 1$. For SVM, we used a soft margin SVM with RBF kernel and $C = 100$.

Enhancing training quality using HMM To improve classification accuracy the training data set consisted of artificial data, generated as follows: For each of the 10 gestures, a Gaussian Hidden Markov Model (HMM) with c components was defined. Each HMM was trained for I iterations, using the sequences of activation maps belonging to the active windows of the training data set. From each trained HMM, new sequences of activation maps were generated, and these maps were added to the training data set.

We empirically set the parameters to: $c = 4$ and $I = 10$.

Results

To demonstrate reliable finger gesture recognition with the soft electrodes we used sEMG data collected from the arm of healthy volunteers to train and test several different classification models. sEMG data were first collected and prepared following these steps: (1) sEMG recording during hand gesturing (2) sEMG data segmentation (3) constructing RMS maps for each segment and finally (4) classification of each segment with a trained NN. Some of the collected data were used for the training (see Fig. 3 for a schematic presentation of the data flow).

Throughout the text we use the following definitions: For each subject, there are two *sessions*, each recorded on a different date with a new electrode array placed approximately at the same location on the hand. For each one of these sessions, there are

three *hand positions* (see Fig. 3), and for each one, there are 100 *events* corresponding to 10 different *gestures* repeated 10 times each. Events were identified either manually or automatically (see Methods section). In total, we collected 300 events for each session.

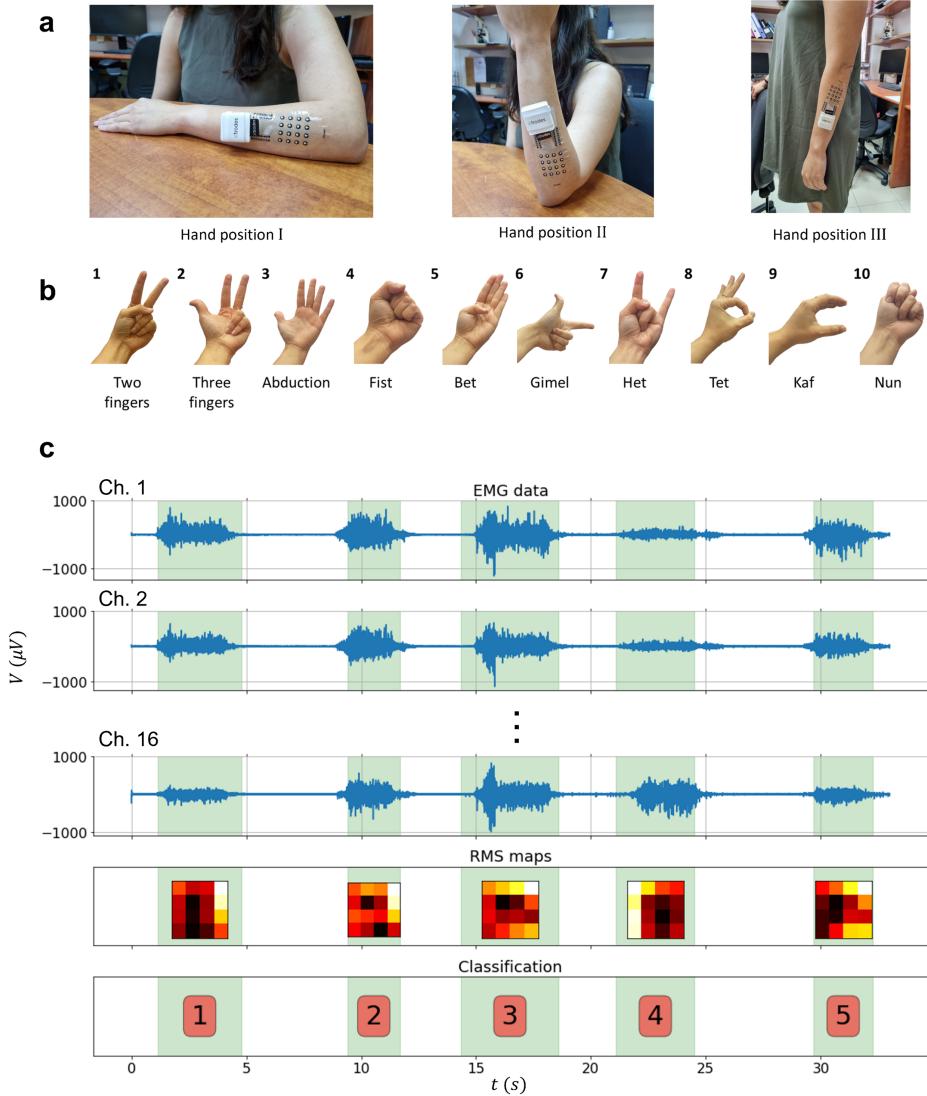


Figure 3. Data collection and analysis scheme. (a) The 3 hand positions examined in the experiment. (b) The 10 hand gestures examined in the experiment. (c) sEMG data from channels 1, 2 to 16. This data is segmented and RMS maps are derived. Corresponding hand gestures are identified algorithmically.

1. sEMG signals

Soft 16 electrode arrays were placed at the region of the extensor digitorum muscle (Fig. 3(a)). Healthy volunteers performed finger gestures (see Methods section) while the sEMG activity of the muscle was recorded. The soft nature of the electrodes along with the small dimension of the wireless DAU allowed subjects to perform natural gestures while recording almost artifact free sEMG data. sEMG data were collected during

10 different hand gestures (Fig. 3(b)) with three arm and body positions (Fig. 3(a)): (1) Arm placed on the table (*position I*), (2) The elbow is placed on the table and the arm is raised by 90 degrees (*position II*), and (3) Subject is standing with the arm next to the body (*position III*). Typical filtered sEMG signals of five gestures in three different electrodes is presented in Fig. 3(c), demonstrating some degree of variability between different gestures. The highlighted regions were automatically segmented. For each such segment, RMS activation maps were generated. These 4 by 4 matrices provide a normalized representation of sEMG activity in the electrode space and is used as an input to the NN.

We selected 10 gestures based on their physiological link with the location of the electrode array. This link is already apparent in the filtered sEMG data (Fig. 3(b)) but it is particularly conspicuous when examining the RMS maps (for clarity, RMS intensity for each electrode was calculated over the entire duration of each action) (Fig. 4). Consecutive activation maps of the same gesture appear consistent within the same hand position, while varying between gestures. Importantly, the same gesture appear to have different maps when the arm position was changed. This result reflects the complexity of the sEMG data. From close examinations of other maps it is evident that small differences in electrode placement (either by repetitive use by the same subject or by different subjects) result with different activation maps. It is therefore important that the classification will be stable against these differences so network training does not have to be fully repeated for each electrode placement, especially when used by the same individual.

2. Constructing Train and Test Databases

Each session, from each hand position, contributed 100 events of 10 repetitions for each of the 10 hand gestures. From each event, we generated normalized RMS maps similar to those described in Fig. 4. For each subject, the maps obtained from eight of the 10 events of each gesture were assigned to the training database, and the remaining maps were added to the test database.

3. Classification models

Based on the results discussed above, we set to realize a NN based classification which is not only accurate but also requiring minimal tuning for newly acquired data (different recording sessions). We implemented both convolutional and recurrent neural networks. For comparison, we also implemented two classical classification algorithms: KNN and SVM. In addition to the original training data set acquired, the training data set of the networks consisted of artificial data (see Methods).

The CNN, KNN and SVM described above receive an RMS maps as input and output a classification. To improve accuracy, for each segment we used a serious of

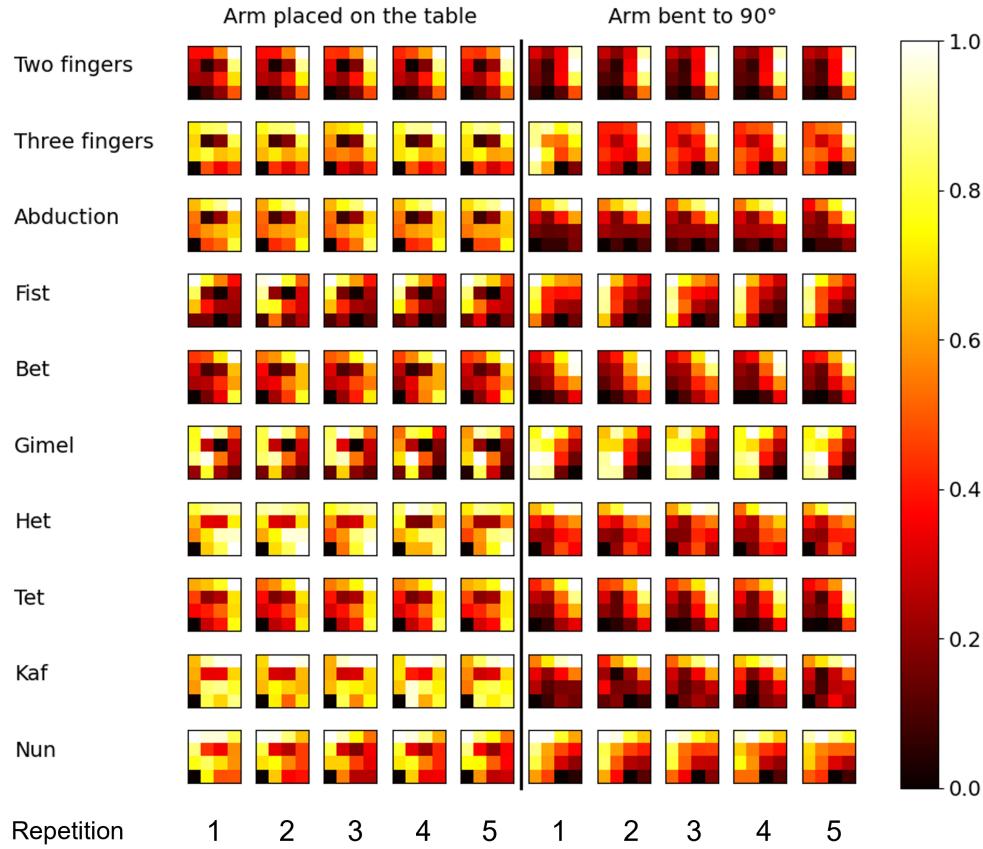


Figure 4. RMS maps. An example of RMS maps of one subject. Each row represents one hand gesture, and the columns represent different repetitions of the same gesture. The first five columns were recorded as the subject's arm was placed on the table, and the next five columns were recorded with the elbow placed on the table, and the arm bent to 90 degrees.

maps and majority voting of single-map classifications, to obtain a final classification for each event.

4. Evaluating the Classification Models

We examined the performance of the models in classifying the sEMG data, focusing on the ability to overcome variability between hand positions and sessions. These capabilities are essential for finger gesture recognition in natural conditions. In order to test these capabilities, we designed three classification tasks (see Tab. 1). Each of these tasks was performed separately with each subject: In *task 1*, we trained the model with events from the training database belonging to hand position II. We then test it on the events from the test database belonging to the same hand position. In *task 2*, we train the model with events from the training database belonging to hand positions I,

II and III. We then test it on the events from the test database belonging to the same hand positions. In *task 3*, we first train the model with all events from the training database belonging to session I, from all subjects (1440 events). After this training stage is finished, we use 20 events from hand position II of session II, i.e. only two repetitions from each gesture, to fine-tune the model. Then, we tested it with 60 events from session II, hand position II.

Table 1. Classification tasks used to evaluate the proposed algorithms

Task	Training Data	Test Data
1	160 Events from hand position II (all sessions)	40 Events from hand position II (all sessions)
2	1440 Events from hand positions I, II and III (all sessions)	160 Events from hand positions I, II and III (all sessions)
3	240 Events from all hand positions from session 1, and only two repetitions of each gesture from session 2, hand position II (20 events)	60 Events from session 2, hand position II

Overall, we evaluated four classification models (KNN, SVM, CNN and RNN). Obtained accuracy values (averaged over $N = 5$ subjects) are presented in Tab. 2. Best results were obtained for CNN, with average accuracy values of 92.5 for 10 gestures for Task 1. Reducing the number of gestures to 8 increases the accuracy to 94.5

Table 2. Accuracy (%) of different algorithms. N=5

	Task 1	Task 2	Task 3
KNN	80.8 ± 10.3	76.9 ± 7.5	81.7 ± 6.7
SVM	86.7 ± 8.5	81.7 ± 8.5	81.7 ± 7.6
CNN	92.5 ± 2.0	84.7 ± 2.7	84.4 ± 7.9
RNN	69.2 ± 8.2	63.3 ± 8.8	59.4 ± 6.9

In addition, confusion matrices of the three tasks using CNN, the most accurate classification algorithm, are shown in Fig. 5.

5. Full Data analysis Flow

Once having a trained CNN, we can apply the analysis scheme to a newly acquired recording. Namely, we applied automatic segmentation, generated RMS maps, feed each map to the CNN and performed majority voting on the classifications (see Fig.1(b)). To test this scheme, we used one session from each subject (session 2, position II), and apply the automated classification algorithm. For evaluation of the segmentation algorithm, we calculated the intersection over union (IoU) between each manually segmented

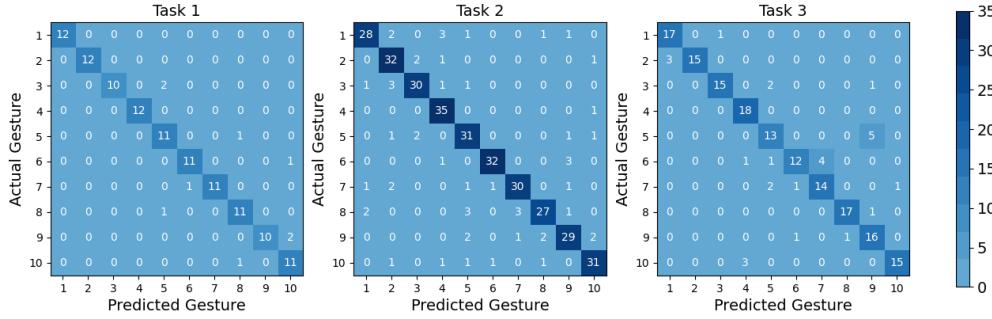


Figure 5. Confusion matrices for CNN. Confusion matrices for tasks 1,2 and 3 for CNN classifier. N=12.

window and its corresponding automatically segmented window. If the IoU value was above 0.2, the segmentation was considered correct (i.e., true positive). Otherwise, we checked whether the gesture was missed by the algorithm (i.e., false negative) or the algorithm incorrectly recognized a non-active segment as active (i.e., false positive). We then calculate the precision and recall values, which were averaged over all subjects. The obtained values are:

$$\text{precision} = \frac{TP}{TP+FP} = 0.984 \pm 0.028$$

$$\text{recall} = \frac{TP}{TP+FN} = 0.997 \pm 0.006.$$

Finally, the CNN was trained using the first eight repetitions of each gesture, and tested on the two last repetitions of each gesture, which are automatically segmented, excluding the false positives. The obtained classification accuracy is $88.33 \pm 5.77\%$. From these results, we see that our proposed segmentation algorithm provides results which are almost identical to the manually obtained ones, and that classification quality is not reduced by the introduction of automatic segmentation.

Discussion and conclusions

In this investigation we demonstrated automated classification of sEMG data recorded using a novel user-friendly wireless system. We presented a NN based algorithm which can classify finger gestures in natural scenarios. Specifically, we demonstrated the ability to perform gesture classification which is insensitive to the position of the hand with accuracy of 84.7%, and to classify hand gestures from a new recording session with accuracy of 84.4%, using only a short calibration step. For the tasks used in the reported investigation, CNN-based model outperforms RNN and the classical models of KNN and SVM.

An important element contributing to the high performances of the system described here are the soft electrode arrays. We have previously found that the sEMG signal to noise ratio (SNR) of these electrode arrays meets the criteria for recording high quality sEMG signals. Implementing an internal ground and using a new miniature wireless system further contributed to our ability to perform sEMG recordings with almost no mechanical artifact even under natural conditions and in different hand positions.

The classification accuracy of the system described here match the state of the art, while doing so with significantly fewer electrodes. Several recent studies reported on sEMG based FGR (Table 3). Atzori *et al.* [12] established the NinaPro database for sEMG based hand movement classification, and used linear discriminant analysis (LDA), k-nearest neighbors (KNN), support vector machine (SVM) and multi-layer preceptron (MLP) to classify gestures. Using carefully placed 10 electrodes, they were able to distinguish between 52 hand gestures with accuracy of 76%. In a later study, CNN was used for classification of the same database, achieving accuracy of $66.59 \pm 6.40\%$ [14]. Other studies focused on high density EMG (HD-EMG) recordings. Rojas *et al.* [10] used activation maps obtained from HD-EMG recordings of the forearm muscles to classify between 12 hand gestures with accuracy of 90%. Amma *et al.* [11] characterized another database for sEMG based FGR (CSL-HDEMG), using 192-electrode array and a naive Bayes classifier to discriminate 27 gestures with accuracy of up to 90%. Geng *et al.* [13] introduced a new database (CapgMyo), consisting of 8 gestures recorded using a 128-electrodes array. Using a convolutional neural network (CNN), they were able to reach accuracy of up to 99.5%. They also achieved recognition accuracy of 96.8% and 77.8% on the CSL-HDEMG and NinaPro databases, respectively. Later studies achieved improved results using these databases. Wei *et al.* [15] used a Multi Stream CNN, reaching accuracy of 99.8%, 95.4% and 85% on the CapgMyo, CSL-HDEMG and NinaPro databases, respectively, while Padhy [16] proposed a multilinear singular value decomposition (MLSVD) approach resulting in accuracies of 98.0% and 98.6% for CapgMyo, and CSL-HDEMG databases, respectively. The first classification task introduced in this work is similar to classification tasks described in previous studies. A comparison is provided in Tab. 3. From this comparison, it is apparent that our system achieves similar accuracy to previous methods, while utilizing a much smaller and more convenient electrode array with no need for careful electrode placement. Moreover, task 2 and 3 examined in this work have no equivalents in the literature.

In the investigation reported here, data were processed off-line. For many FGR applications on-line analysis is desired and can be achieved if data transfer and analysis times are fast enough. In the scheme presented here total analysis time was approximately 1 s, which mainly results from the segmentation algorithm latency.

Although sEMG requires electrode placement at close proximity to the muscle, the use of 16 electrode arrays and CNN analysis negate the need for very precise placement,

Table 3. Comparison between various classification models applied to position I.

Reference	Electrodes	Gestures	Method	Accuracy (%)
Atzori (2015)[12]	10 OttoBock	52	SVM	76.0
Amma (2015)[11]	192 dry	27	Naive Bayes	90.4
Geng (2016)[13]	10 OttoBock	52	CNN	77.8
Geng (2016)[13]	192 dry	27	CNN	96.8
Geng (2016)[13]	128 dry	8	CNN	99.5
Wei (2019)[15]	10 OttoBock	52	MS CNN	85.0
Wei (2019)[15]	192 dry	27	MS CNN	95.4
Wei (2019)[15]	128 dry	8	MS CNN	99.8
Padhy (2021)[16]	192 dry	27	MLSVD	98.0
Padhy (2019)[16]	128 dry	8	MLSVD	96.8
This study	16 dry	10	CNN	92.5
This study	16 dry	8	CNN	94.5

allowing accurate classification despite the variability of multiple sessions.

In this investigation we used 16 electrode arrays. These 4 by 4 arrays clearly provide more information than low resolution sEMG, contributing to effective discrimination between gestures. Higher electrode resolution may contribute to improved resolution in particular if discrimination between more gestures is needed. It is important to note that increasing electrode count may increase data analysis time and DAU dimensions, tempering with the ultimate goal of real time analysis under natural conditions.

sEMG signals contain information on applied force. This information can be important in many applications. In the current study we did not utilize this feature: RMS maps were normalized and as such, amplitude information was discarded. Moreover, spectral analysis was not implemented to gain additional information about the applied force. These topics remain for future investigations.

As we demonstrated here, sEMG has many important benefits compared with video analysis and smart gloves. On the downside, sEMG gesture separation is limited to specific degrees of freedom associated with the targeted muscle. sEMG may gain from the combination of other technologies such as video or smart gloves, to improve network training and resolution. For example, in this study we did not exploit three-dimensional acceleration data which was recorded by a built in 3 axis inertial sensor implemented in the wireless DAU.

To conclude, the results we presented here demonstrate an important step in achieving FGR in natural conditions, using sEMG signals. In particular, the use of a minimally interfering wearable system to measure sEMG from the arm demonstrates

several important advantages, making the system a possible candidate for gaming or Metaverse applications.

Acknowledgments

The authors thank Liron Amihay, David Buzaglo, Kerith Aginsky and Param Gandhi for advice and assistance with sEMG data collection. YH thanks Anat Mirelman for many fruitful discussions.

References

- [1] Kostas Nizamis, Noortje H.M. Rijken, Robbert van Middelaar, João Neto, Bart F.J.M. Koopman, and Massimo Sartori. Characterization of Forearm Muscle Activation in Duchenne Muscular Dystrophy via High-Density Electromyography: A Case Study on the Implications for Myoelectric Control. *Frontiers in Neurology*, 2020.
- [2] Jian Wu, Lu Sun, and Roozbeh Jafari. A wearable system for recognizing american sign language in real-time using imu and surface emg sensors. *IEEE journal of biomedical and health informatics*, 20(5):1281–1290, 2016.
- [3] Feng Wen, Zixuan Zhang, Tianyi He, and Chengkuo Lee. Ai enabled sign language recognition and vr space bidirectional communication using triboelectric smart glove. *Nature Communications*, 12(1):5378, 2021.
- [4] Amit Konar and Sriparna Saha. *Introduction*, pages 1–33. Springer International Publishing, Cham, 2018.
- [5] Yuxiao Chen, Long Zhao, Xi Peng, Jianbo Yuan, and Dimitris N. Metaxas. Construct dynamic graphs for hand gesture recognition via spatial-temporal attention. In *BMVC*, 2019.
- [6] Yuecong Min, Yanxiao Zhang, Xiujuan Chai, and Xilin Chen. An efficient pointlstm for point clouds based gesture recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.
- [7] Lidia Santos, Nicola Carbonaro, Alessandro Tognetti, Josa Luis Gonzalez, Eusebio De la Fuente, Juan Carlos Fraile, and Javier Perez-Turiel. Dynamic gesture recognition using a smart glove in hand-assisted laparoscopic surgery. *Technologies*, 6(1), 2018.
- [8] T. Primya, G. Kanagaraj, K. Muthulakshmi, J. Chitra, and A. Gowthami. Gesture recognition smart glove for speech impaired people. *Materials today : proceedings*, 2021.
- [9] Rajarajan Ramalingame, Rim Barioul, Xupeng Li, Giuseppe Sanseverino, Dominik Krumm, Stephan Odenwald, and Olfa Kanoun. Wearable smart band for american sign language recognition with polymer carbon nanocomposite-based pressure sensors. *IEEE Sensors Letters*, 5(6):1–4, 2021.
- [10] Monica Rojas-Martínez, Miguel A Mañanas, and Joan F Alonso. High-density surface emg maps from upper-arm and forearm muscles. *Journal of neuroengineering and rehabilitation*, 9(1):85–85, 2012.
- [11] Christoph Amma, Thomas Krings, Jonas Böer, and Tanja Schultz. Advancing muscle-computer interfaces with high-density electromyography. In *Proceedings of the 33rd Annual ACM Conference on human factors in computing systems*, CHI ’15, pages 929–938. ACM, 2015.
- [12] Manfredo Atzori, Arjan Gijsberts, Ilya Kuzborskij, Simone Elsig, Anne-Gabrielle Mittaz Hager, Olivier Deriaz, Claudio Castellini, Henning Muller, and Barbara Caputo. Characterization of a benchmark database for myoelectric movement classification. *IEEE transactions on neural systems and rehabilitation engineering*, 23(1):73–83, 2015.
- [13] Weidong Geng, Yu Du, Wenguang Jin, Wentao Wei, Yu Hu, and Jiajun Li. Gesture recognition by instantaneous surface emg images. *Scientific reports*, 6(1):36571–36571, 2016.

- [14] Manfredo Atzori, Matteo Cognolato, and Henning Müller. Deep learning with convolutional neural networks applied to electromyography data: A resource for the classification of movements for prosthetic hands. *Frontiers in neurorobotics*, 10:9–9, 2016.
- [15] Wentao Wei, Yongkang Wong, Yu Du, Yu Hu, Mohan Kankanhalli, and Weidong Geng. A multi-stream convolutional neural network for semg-based gesture recognition in muscle-computer interface. *Pattern recognition letters*, 119:131–138, 2019.
- [16] Sibasankar Padhy. A tensor-based approach using multilinear svd for hand gesture recognition from semg signals. *IEEE sensors journal*, 21(5):6634–6642, 2021.
- [17] Lilah Inzelberg, David Rand, Stanislav Steinberg, Moshe David-Pur, and Yael Hanein. A wearable high-resolution facial electromyography for long term recordings in freely behaving humans. *Sci Rep*, 8(1):2058, Feb 2018.
- [18] Lilah Inzelberg, Moshe David-Pur, Eyal Gur, and Yael Hanein. Multi-channel electromyography-based mapping of spontaneous smiles. *J Neural Eng*, 17(2):026025, Apr 2020.
- [19] Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6):84–90, 2017.
- [20] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. 2015.
- [21] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. 2014.
- [22] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.*, 15(1):1929–1958, jan 2014.