Convolutional Neural Networks Improve the Prediction of Hand Movement Speed and Force from Single-trial EEG

Ramiro Gatti, Yanina Atum, Luciano Schiaffino, Mads Jochumsen, and José Biurrun Manresa

***Abstract*—*Objective.* Building accurate movement decoding models from brain signals is crucial for many biomedical applications. Decoding specific movement features, such as speed and force, may provide additional useful information at the expense of increasing the complexity of the decoding problem. Recent attempts to predict movement speed and force from the electroencephalogram (EEG) achieved classification accuracy levels not better than chance, stressing the demand for more accurate prediction strategies. Thus, the aim of this study was to improve the prediction accuracy of hand movement speed and force from single-trial EEG signals recorded from healthy volunteers. *Approach.* A strategy based on convolutional neural networks (ConvNets) was tested, since it has previously shown good performance in the classification of EEG signals. *Main results.* ConvNets achieved an overall accuracy of 84% in the classification of two different levels of speed and force (4-class classification) from single-trial EEG. *Significance.* These results represent a substantial improvement over previously reported results, suggesting that hand movement speed and force can be accurately predicted from single-trial EEG.**

***Index Terms*—Convolutional neural networks, hand movement, speed and force, movement prediction, multi-class classification, single-trial EEG.**

I. Introduction

DECODING brain signals to predict movements is useful in many research areas, such as neuromechanics, neuroscience and robotics [1]. Furthermore, it is also relevant in neurological rehabilitation, since it has potential to facilitate the assessment of the central nervous system in patients, promote neural plasticity, improve motor dysfunction and allow the control of assistive devices through brain-computer interfaces (BCI) [2]. In this regard, motor commands generated prior to or during movement execution can be extracted from specific oscillatory patterns in the electroencephalogram (EEG) [3], [4]. Specifically, the component waves of movement-related cortical potentials (MRCPs) immersed in the EEG, such as the readiness potential and contingent negative variation, carry information about anticipatory behaviour, which can be used to predict movements, i.e., to detect and classify a particular movement before it is actually executed during self-paced or cue-based paradigms [5]–[7].

The movement decoding process is generally focused on detecting a predetermined final state and lacks attention regarding the quality of the action, resulting in simple, rough commands [8]. Research on fine movements of body structures such as fingers [9], or complex movement control [10] is comparatively scarce. It is straightforward to hypothesize that better commands can be achieved if movement kinematics and kinetics are taken into account in the decoding process [11]. Indeed, the decoding of hand movement velocities [12], [13] and 3D trajectories [14] as well as the prediction of force and speed from a specific movement [10], [15] showed promising results. However, recent attempts to predict speed and force from a hand grasping tasks resulted in a classification accuracy not better than chance level [16]–[18], stressing the need for more accurate prediction schemes.

The control strategies generated by the nervous system for goal-directed motor behaviour are extremely complex. Thus, pattern recognition systems used to decode and predict movements require careful engineering and domain expertise to transform raw EEG signals (usually by means of a feature extraction subsystem) into a suitable representation for the classification stage [19]. In this regard, several techniques have been proposed for feature extraction, e.g., common spatial patterns, independent component analysis, and joint time-frequency analysis, and also for classification, e.g., nearest neighbour classifier, linear discriminant analysis, support vector machines (SVMs), and ensemble strategies, among others [20]. An alternative is to use representation learning methods that automatically perform a feature extraction and classification through optimisation algorithms. Deep learning is a paradigmatic example, with multiple levels of representation obtained by combining simple but non-linear modules that transform the input into increasingly more abstract levels [19]. In line with this, a decoding model based on deep learning implemented through convolutional neural networks (ConvNets) recently showed promising results in classification performance using different EEG paradigms [21].

The aim of the present study was to improve the prediction accuracy of hand movement speed and force from single-trial EEG signals recorded from healthy volunteers. Subjects executed an isometric right hand palmar grasp task using two predefined levels of force (20% and 60% of the maximum

R. Gatti, Y. Atum, L. Schiaffino and J. Biurrun Manresa are with the Laboratory for Rehabilitation Engineering and Neuromuscular and Sensory Research (LIRINS), Faculty of Engineering, National University of Entre Ríos, Oro Verde, Argentina (e-mail: rgatti@ingenieria.uner.edu.ar, yatum@ingenieria.uner.edu.ar, lschiaffino@ingenieria.uner.edu.ar, jbiurrun@ingenieria.uner.edu.ar).

R. Gatti and J. Biurrun Manresa are also with the Institute for Research and Development in Bioengineering and Bioinformatics (IBB), CONICET-UNER, Oro Verde, Argentina.

J. Biurrun Manresa is also with the Center for Neuroplasticity and Pain (CNAP), SMI®, Aalborg University, Aalborg, Denmark

M. Jochumsen is with the Center for Sensory-Motor Interaction, SMI®, Aalborg University, Aalborg, Denmark (e-mail: mj@hst.aau.dk).

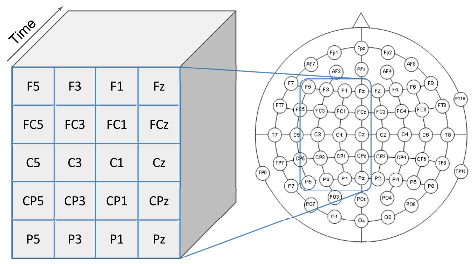

Fig. 1. Input data arrangement for the ConvNet based on the spatial distribution of the recorded channels in healthy volunteers

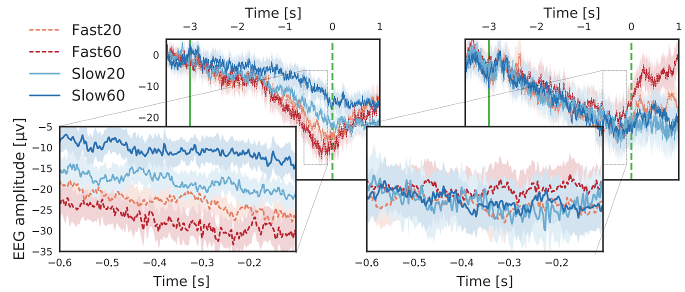

Fig. 2. Representative examples of 3 s EEG trials (back) and the corresponding 500 ms epochs (front) for healthy volunteers relative to movement onset. The solid trace and shading represent mean and 95% confidence intervals for each class, respectively, derived using 5000 bootstrap iterations. Vertical lines represent cue (solid) and movement onset (dashed) times.

voluntary contraction, MVC) and speed (a 3-s slow grasp and a 0.5-s fast grasp). EEG data were minimally pre-processed, in order to minimize user bias. A prediction strategy using ConvNets was implemented and contrasted with results obtained using SVM on the same datasets. Overall classification accuracy, precision and recall were quantified in order to evaluate the performance of the proposed prediction strategies.

II. Materials and Methods

A. Dataset

A dataset consisting of EEG recordings from sixteen healthy volunteers was employed [16]. Written informed consent was obtained from all subjects prior to participation, and the Declaration of Helsinki was respected. The study was approved by the local ethical committee of Region NordJylland (approval no. N-20100067). EEG was recorded during four isometric right palmar grasp tasks with different execution speeds and force levels (expressed as percentage of MVC), categorized as follows: *Slow20*, 3 s to reach 20% MVC; *Slow60*, 3 s to reach 60% MVC; *Fast20*, 0.5 s to reach 20% MVC and *Fast60*, 0.5 s to reach 60% MVC. Forty externally cued repetitions (trials) were performed for each task. A Neuroscan NuAmp Express amplifier was used to record the EEG (Compumedics Ltd., Victoria, Australia) from the electrode locations shown in Fig. 1, in accordance to the 10/10 system. The corresponding EEG channels were referenced to the right earlobe and grounded at nasion. During the experiment, the impedance of all electrodes was kept below 5 kΩ, continuously sampled at 500 Hz and stored for offline analysis. For additional details of the experimental procedure, please refer to [16].

1) Pre-processing: EEG was notch-filtered (50 Hz) uzing a zero-phase filter in order to reduce power line interference and the baseline (1-s interval before the cue) was subtracted from all trials. No further pre-processing or filtering was applied to the EEG signals, and noisy epochs were not removed, in order to minimize user bias. Forty trials per task were executed, resulting in 160 trials per subject. Trials were subsequently segmented into 500-ms epochs, from 600 ms to 100 ms before movement onset (Fig. 2). EEG epochs were finally arranged in a 5 × 4 × 250 matrix with a two-dimensional spatial distribution and the time samples in the third dimension (Fig. 1). Data were divided in 128 trials (80%) for training and validation and 32 trials (20%) for testing. The training and validation set was further split into 102 trials (80%) for training and 18 trials (20%) for validation.

B. Prediction strategies

1) Convolutional Neural Network: The model was based on the EEGnet described in Lawhern et al. [21]. The ConvNet was built in TensorFlow 1.11 [22] using the Keras API [23] and trained on an Dell Precision 7910 workstation with an NVIDIA Titan Xp GPU, using CUDA 9 and cuDNN 7.3.

The model consisted of two blocks (Table I). The input of the first layer was a pre-processed three-dimensional (3D) matrix for each trial, which was reshaped to apply four temporal filters (F_1) to each channel. Following the original net architecture, convolutional kernels of size (1, 64) were applied in the temporal dimension. Kernel weights were initialized with a Glorot uniform technique, without applying a bias vector. The spatial dimension size was kept constant through zero padding without stride. Then, a batch normalization was applied. Afterwards, the matrix was reshaped and its dimensions were permuted in order to apply a depthwise convolution to every temporal slice by means of the wrapper time distributed layer [24]. Two spatial filters of size (Cx, Cy) for each feature map (deep multiplier parameter D) were applied and then the matrix was reshaped and dimensions were permuted again. Afterwards, a batch normalization followed by a Exponential Linear Unit (ELU) activation with alpha = 1, an average pooling of size (1, 4), and drop-out with a rate of 0.25 were applied.

In the second block, a 2D separable convolution of size (1, 16) with eight filters (F_2) was applied. Henceforth, ELU activation, batch normalization, average pooling of size (1, 8), and drop-out were applied using the same hyperparameters as in the first block. Finally, the data was flattened to a single dimension and the four resulting scores of the dense layer were transformed to probabilities by means of a softmax activation.

The ConvNet architecture was devised taking into account temporal and spectral characteristics of the EEG signals tested in the original experiment, such as sampling rate, window size and frequency resolution of the filters resulting from the convolutional kernels. For this reason, an alternative architecture was also tested, whose parameters were derived from extrapolating the original criteria to match the characteristics of the dataset used in this study. The resulting architecture had the first convolutional kernels of size (1, 250), an average pooling of size (1, 4) in the second block. The rest of the

TABLE I
CONVNET ARCHITECTURE

Block	Layer	Output size	Param. #
1	Input	$(C_y \times C_x \times T)$	0
	Reshape	$(1 \times C_y * C_x \times T)$	0
	Conv2D (1,64) $\times F_1$	$(4 \times C_y * C_x \times T)$	256
	Batch normalization	$(4 \times C_y * C_x \times T)$	16
	Reshape	$(4 \times C_y \times C_x \times T)$	0
	Permute	$(T \times 4 \times C_y \times C_x)$	0
	TD $(C_y, C_x) \times D * F_1$	$(T \times D * F_1 \times 1 \times 1)$	160
	Permute	$(D * F_1 \times 1 \times 1 \times T)$	0
	Reshape	$(D * F_1 \times 1 \times T)$	0
	Batch normalization	$(D * F_1 \times 1 \times T)$	32
	Activation (ELU)	$(D * F_1 \times 1 \times T)$	0
	AveragePooling2D (1, 4)	$(D * F_1 \times 1 \times T/4)$	0
	Dropout (.25)	$(D * F_1 \times 1 \times T/4)$	0
2	SeparableConv2D (1, 16) $\times F_2$	$(F_2 \times 1 \times T/4)$	192
	Batch normalization	$(F_2 \times 1 \times T/4)$	32
	Activation (ELU)	$(F_2 \times 1 \times T/4)$	0
	AveragePooling2D (1, 8)	$(F_2 \times 1 \times T/32)$	0
	Dropout (.25)	$(F_2 \times 1 \times T/32)$	0
	Flatten	$(F_2 * T/32)$	0
	Dense	(N)	228
	Activation (Softmax)	(N)	0
	Total		916

C_x = channels (mediolateral direction), C_y = channels (anteroposterior direction), T = time samples, F_1 = number of temporal filters, TD = TimeDistributed (DepthwiseConv2D), D = depth multiplier (number of spatial filters), F2 = number of pointwise filters, N = number of classes.

parameters remained unchanged.

The learning process consisted of a fixed number of learning steps using mini-batchs of 16 randomly selected trials and the Adam optimization. The initial number of learning steps was set to 500, and validation accuracy and loss curves as a function of the number of learning steps were obtained in order to derive the smallest number of learning steps required to achieve an acceptable classification accuracy. The loss obtained from the validation set was used as metric, and the model was updated if the loss decreased compared to the last saved model. To prevent model overfitting, only the model with the lowest validation loss was kept. In this regard, the relationship between training set size and performance was also analyzed to verify that the training set size was appropriate in relation to the dataset size [25].

2) Support vector machine: SVMs are popular supervised learning models used for classification, that transform a non-linearly separable problem into a linearly separable problem by projecting data into a new feature space through the use of kernel functions, in order to find the optimal decision hyperplane in this feature space. This method was initially proposed to solve two-class problems, although strategies were later developed to extend this technique to multi-class classification problems [26]. For this study, SVMs were implemented for reproducibility purposes, as they would allow a direct comparison with prior studies using the same data [15], [16]. Unlike ConvNets, SVMs require a separate feature extraction stage before classification is performed [16], [27]–[29]. In this regard, ten features were calculated for each 500-ms epoch: 1) Basal amplitude value, using the Hilbert transform to estimate the area envelope, 2) Kurtosis, 3) Curve length, as the sum of consecutive distances between amplitudes, 4) Noise level, as 3 times the standard deviation of the amplitudes, 5) Number of positive peaks, 6) Average nonlinear energy, using the Teager energy operator, 7) Number of zero crossings, 8) Maximum negativity peak, 9) Root-mean-square (RMS) amplitude, and 10) Average power in the interval from 0 to 5 Hz, using Welch power spectral density estimator with a Hamming window and a 50% overlap. With regards to the SVM parameters, a Radial Basis Function (RBF) were used as kernel. Based on a heuristic search, the cost hyperparameter of the SVM was set to 0.001 and the gamma hyperparameter of the kernel was set to 0.0002. Furthermore, a one-against-one strategy was used to implement the multi-class SVM prediction strategy. This method constructed *k(k-1)/2* classifiers, where *k* is the number of classes of the problem. Each classifier used the training data from two classes chosen out of *k* classes. After the training process was over, a voting strategy was used to determine to which class each pattern belongs to. The open source library tool for classification and regression problems (LIBSVM) was used to build the SVMs [30].

C. Data analysis

An individual prediction strategy was trained for each subject (i.e., one ConvNet and 6 SVMs to address the 4-class classification problem), and the same data partitioning for training, validation, and testing was used for the ConvNet and the SVMs. Additionally, ConvNets were trained with the same dataset partitioning, but with randomly scrambled labels, in order to determine the chance classification accuracy level. Furthermore, ConvNets were also trained with increasing number of examples in order to test the effect of training set size on classification accuracy. To ensure the generalizability of the final results, a 5-fold cross-validation procedure was performed, in which a new individual prediction strategy (ConvNet or SVM) was trained for each fold. The overall classification accuracy and Cohen's kappa, together with per-class precision and recall were quantified for each subject, as the mean values of the 5-fold cross-validation procedure computed from the test sets. Finally, to ensure the reproducibility of these results, the source code for the ConvNet is available at [link placeholder] and the dataset is available from from the corresponding author on reasonable request.

D. Statistics

A paired t-test was used to assess differences in overall classification accuracy and kappa values between classifiers (ConvNet vs. SVM) and between ConvNet architectures. A repeated measures analysis of variance was used to assess differences in precision and recall, with Classifier (levels: ConvNet, SVM), Speed (levels: Fast, Slow) and Force (levels: 20% MVC, 60% MVC) as factors. Main effects and two-way interactions were analysed. The Shapiro-Wilk test was performed in order to assess the assumption of normality, which held for all indexes. Performance indexes are reported as $mean \pm standard\ deviation$ unless stated otherwise. P values smaller than 0.05 were regarded as statistically significant.

III. Results

A. ConvNet validation

The evolution of the validation accuracy and validation loss as a function of the number of learning steps is shown in Fig. 3 (left). It can be observed that the accuracy reaches a stable value after 100 steps while the loss stabilizes after approximately 300 steps, indicating that more training steps would not improve the results and that the ConvNet is not overfitting the data. As a reference, training the ConvNets using 500 steps took approximately 3 min per subject and classifying each new trial took approximately 7 ms. Furthermore, Fig. 3 (right) shows the chance level accuracy obtained after training the model with randomly scrambled labels. In this case, the prediction accuracy is close to the theoretical chance level of 25% for a 4-class classification problem and it does not improve with the number of training steps. Fig. 4 shows the relationship between test accuracy and training set size, from which it can be deduced that the ConvNet strategy can be trained with as few as 80 examples and still achieve an acceptable classification accuracy above 80%.

B. Performance of the classification strategies

The overall classification accuracy for the ConvNet model ($84.0 \pm 7.0\%$) was significantly higher compared to the SVM ($72.4 \pm 7.9\%$; $t_{15} = 5.072$, $p < 0.001$). Likewise, kappa values were significantly higher for the ConvNet (0.80 ± 0.09) that for the SVM (0.63 ± 0.10; $t_{15} = 5.434$, $p < 0.001$). Precision and recall values for all healthy volunteers are shown in Fig. 5. The precision of the ConvNet model ($84.2 \pm 7.0\%$) was higher compared to the SVM ($72.4 \pm 7.9\%$; $F_{1,15} = 27.252$, $p < 0.001$). With regards to recall, it was significantly higher for the ConvNet model ($84.0 \pm 7.0\%$) compared to the SVM ($73.2 \pm 7.5\%$; $F_{1,15} = 23.933$, $p < 0.001$). No further significant main effects or interactions were found for precision or recall. Finally, no significant differences in accuracy were found between the original and the alternative architecture ($83.8 \pm 5.6\%$; $t_{15} = 0.200$, $p = 0.844$).

IV. Discussion

A. Neurophysiological aspects of movement prediction

Building efficient movement decoding models from brain signals is crucial for many biomedical applications, particularly in the BCI field that require precision in online control of assistive devices. Moreover, decoding specific movement features, such as speed, force and/or direction, provides additional degrees of freedom, resulting in more accurate and natural motor commands at the expense of increasing the complexity of the decoding problem [12]–[14], [31]. Early attempts to decode movement from brain signals during movement execution or imagination were focused on classifying between limb movements [32]–[34]. Classification accuracy for these studies was close to 80% for 2 classes [32], [33], and close to 56% for 4 classes [34]. Other studies have tried to decode movement of specific body parts from a single limb, such as wrist [35], or individual finger movements [9], obtaining similar results.

On the other hand, prediction of movement, i.e., decoding movement not during, but before its execution, is a much more difficult task. Considering the brain as a predictive neural system, expectation can be seen as a representation of prediction that serve to sensory or motor areas as preparatory processing prior to an event, particularly in short time scales [36]. Movement intention is the first interesting command to decode from EEG before a movement is executed, as trigger for other more complex motor instructions. It is well known that information about movement intention is encoded in the MRCPs, around 1.5 s prior to movement onset [7]. The timing of the prediction is a relevant feature to study, since it has been shown that a sensory stimulus delivered synchronously with the peak negativity of the MRCP maximizes neural plasticity [37]. Furthermore, kinetic information encoded in the movement intention could be particularly useful; for example, by decoding these movement parameters it would be possible to introduce task variability in the rehabilitation training, which has been shown to maximize the motor learning [38].

It has been already shown that pre-movement EEG contains valuable information about motion. Indeed, detection of voluntary movement from single trial EEG using a matched filter approach demonstrated relatively good performance in a 2-class classification scheme (sensitivity $\approx$ 82.5% for healthy subjects) [39]. However, classification rates for multi-class classification problems are still relatively low in healthy volunteers. As an example, recent studies directed towards the extraction of additional information from movement intention beyond simple detection, such as the prediction of the body part that is about to perform the movement [18], or the classification between different types of movement used in daily life, such as palmar, lateral and pinch grasps [10], resulted in classification accuracies not better than chance levels for the 4-class classification attempts.

In particular, previous work with the same dataset used in this study obtained mean accuracy values of approximately 32-40% for the 4-class classification [16], which is on par with the chance level for that type of problem [40]. These result might be partially explained by the fact that the aim of the study was to obtain a fast prediction scheme using few electrodes and a simple classifier that did not require extensive calibration. As such, only one channel was used as input, and the signals were band filtered using low cut-off frequencies values. However, it was recently suggested that information from the entire EEG spectrum is needed to discriminate between task-related parameters from single-trial movement intention [17].

Based on this idea, in this study it was possible to significantly improve the movement prediction accuracy using twenty available channels without additional pre-processing, such as artifact removal or epoch selection. Accuracy levels reached values close to 85% in healthy volunteers, representing an improvement of almost 45% compared to previous results. Therefore, it could be hypothesized that the decoding of complex movement requires more information (in terms of number of channels or features) in order to achieve a classification accuracy comparable with that obtained for simpler movements, such ankle or wrist flexion/extension (binary classification problems) [15], [35], [41], [42]. Finally, previous

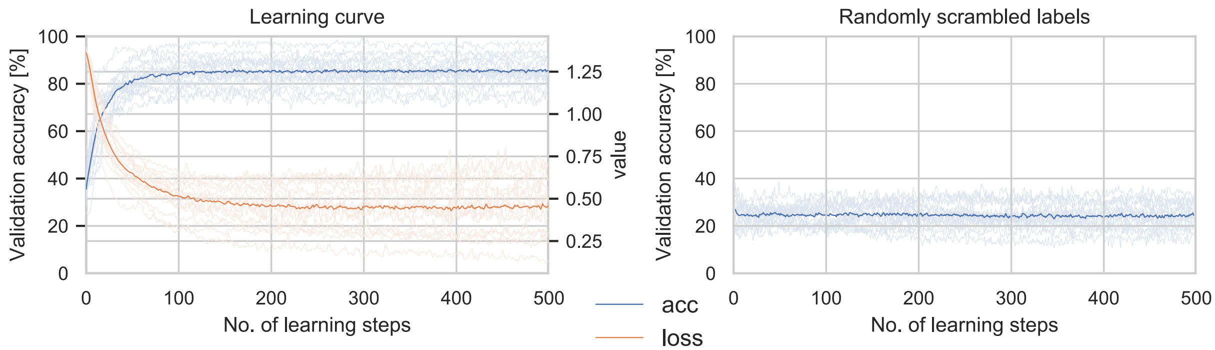

Fig. 3. Left: evolution of the validation accuracy (acc) and validation loss (loss) from all volunteers as a function of the number of learning steps. Right: Validation accuracy from all volunteers as a function of the number of learning steps with randomly scrambled labels. In both the dark line represents the mean validation accuracy for all subjects in each group, and each light line represents the mean validation accuracy for a single subject, derived from the 5-fold cross-validation.

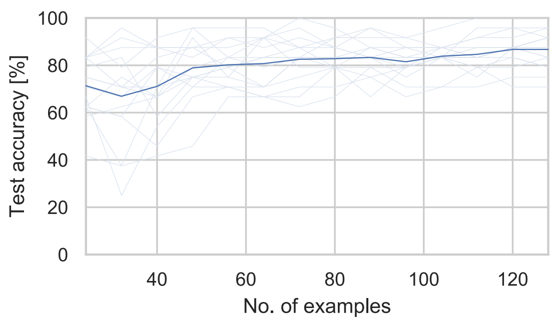

Fig. 4. Test accuracy from all healthy volunteers as a function of the number of examples. The dark line represents the mean test accuracy for all subjects in each group, and each light line represents the mean test accuracy for a single subject, derived from the 5-fold cross-validation.

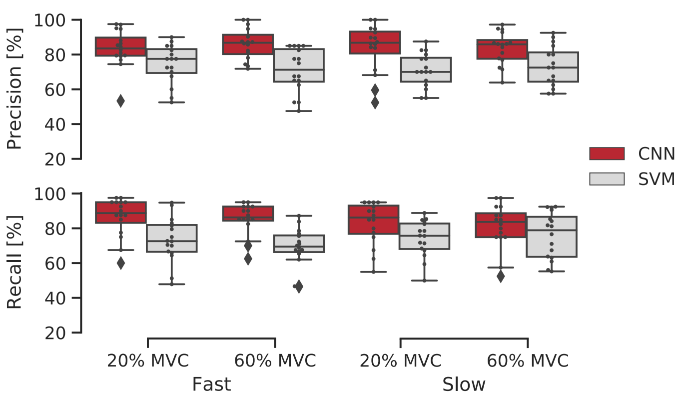

Fig. 5. Classification results over test data. Top: precision and recall for healthy volunteers (n = 16). Boxes represent the median and the 25th and 75th percentiles, whiskers represent 5th and 95th percentiles, diamonds represent values outside of the 5th - 95th percentile range and the individual dots represent the average precision/recall for each individual subject, calculated from the 5-fold cross-validation.

studies have shown that classification of speed tasks achieved higher accuracy in the prediction of ankle dorsiflexion movements [15]. However, it was not the case for hand grasping tasks, since no significant effects of speed or force and no interactions were observed, in line with previous results [16].

B. Methodological aspects of movement prediction

Deep learning methods were originally developed in the computer vision field [43], and recently gained popularity in EEG analysis, in which they are used with the aim of improving classification performance over more traditional approaches, such as linear discriminant analysis, k-nearest neighbours or SVMs [20]. ConvNets are a type of feed-forward deep learning networks that are useful when data have a known topological structure [19], [25]. As a representation learning method, one of the advantages of ConvNets is that feature extraction and classification is intrinsically optimized. Typically, ConvNets consist of a combination of convolutional and pooling layers. The convolutional layer applies mathematical convolution operations through a number of kernels that perform a local weighted sum along the input and return each one in a feature map. Then, the same weights are shared across the input and have only local connections, thereby reducing the amount of network parameters. The pooling layer performs a reduction of the input by applying a function to nearby units, e.g. the maximum value among neighbours, where the units are the pixels of an image or the temporal samples of a biosignal.

The ConvNet implemented in this study is based on a recently proposed architecture that demonstrated good performance employing a small number of parameters in the classification of EEG signals recorded using different paradigms [21]. In this model, the first convolutional layer works as a frequential filter, in which the outcome consists of four different band-pass filters that minimize the error at the output. In accordance with the input structures used in image processing, the EEG input to a ConvNet is usually reshaped into a 2D distribution, by arranging channels along the rows and time samples in the columns [44], [45] or by transforming the input into a new space [8], e.g., to a time-frequency domain through Fourier transform and averaging along the channels [46], [47]. The ConvNet implemented in this study considered the localization of the electrodes in order to keep the spatial relationship between them. Furthermore, EEG signals are commonly pre-processed by using temporal and spatial filters, and epochs containing artifacts or with amplitudes above a certain threshold are rejected in order to improve the signal-to-noise ratio [7]. These processes are time-consuming, prone to user bias and may result in the loss of useful information to decode movement. Taking this into consideration, only minimal and automatic pre-processing (baseline correction and notch filtering) was performed in this study prior to the classification stage, and no epochs were removed. Furthermore, it is worth noting that this strategy does not require extremely large datasets for training and the training time is negligible compared to the average setup time for a BCI, which makes it viable for use in rehabilitation.

The present study neither attempted to find *the* optimal strategy for movement prediction nor advance the knowledge on machine learning strategies for single-trial EEG classification. Instead, the main goal of this work was to determine the achievable levels of prediction accuracy from single-trial EEG using state-of-the-art machine learning techniques, and compare the results obtained with recent reports from the literature, using SVMs as classification strategy and feature selection based on temporal and spatial parameters [16]. The prediction results of the ConvNet were better than the SVM for all tasks and all performance indexes in healthy volunteers by an average of 12 percentage points (Fig. 5). This is even more relevant considering that the SVMs implemented in this study (using twenty available channels) already improved the classification accuracy by approximately 32 percentage points compared to the previous study with the same dataset (using only a single channel, C3, plus an eight-channel Laplacian filter) [16]. It is worth mentioning that preliminary tests performed using other classification strategies, such as bagging trees [48] and random forest [49], resulted in similar performances compared to the SVM (accuracies of $72.6 \pm 5.1\%$ and $72.0 \pm 5.7\%$, respectively), significantly lower than ConvNet performance. Furthermore, a systematic investigation regarding movement prediction performed with combinations of spatial filtering (principal component analysis, independent component analysis, common spatial patterns analysis, and surface Laplacian derivation), temporal filtering (power spectral density estimation and discrete wavelet transform), pattern classification (linear and quadratic Mahalanobis distance classifier, Bayesian classifier, multi-layer perceptron neural network, probabilistic neural network, and SVM), and multivariate feature selection strategy using a genetic algorithm, achieved a maximum accuracy of 75% for binary classification [50]. In contrast, the ConvNet shows better results with minimal pre-processing and optimal combination of feature extraction and classification in a multi-class classification scenario.

C. Limitations and future work

Several constraints need to be considered: attempts to use a single ConvNet to predict movements from all subjects resulted in low performance indexes during pilot tests (average accuracy of $27.4 \pm 4.4\%$). This is not an issue in most real-life applications where the decoding is used to control a device for a single subject (and thus an individual ConvNet is trained for each subject), but nevertheless highlights the difficulty in describing a general behavior of the EEG signal in terms of decoding force and speed. The same issue can be observed when attemtping to understand and visualize of the specific features that allow a good classification, since it is not straightforward to extract and interpret physiological information from the network, and these feature vary between subjects. Furthermore, even if high accuracy was achieved offline, it is crucial to perform real-time tests with adequate feedback. Future work will be directed towards testing the strategy with a real application, for which an accurate detection of the movement onset is necessary and an idle state should be considered [51]. Finally, once the definitive scheme has been defined, efficient hardware implementations should be tested in chips or field-programmable gate arrays [19].

V. Conclusion

The results from this study suggest that hand movement speed and force can be accurately predicted from single-trial EEG using convolutional neural networks, although additional considerations are still required to transfer these protocols from laboratory to clinic.

Acknowledgment

The workstation used in this study was provided by the Center for Neuroplasticity and Pain (CNAP), which is supported by the Danish National Research Foundation (DNRF121). The Titan Xp GPU used for this research was donated by the NVIDIA Corporation.

References

[1] A. D. Nordin, W. Z. Rymer, A. A. Biewener, A. B. Schwartz, D. Chen, and F. B. Horak, "Biomechanics and neural control of movement, 20 years later: what have we learned and what has changed?" *Journal of NeuroEngineering and Rehabilitation*, vol. 14, no. 1, p. 91, 2017.

[2] C. Brunner, N. Birbaumer, B. Blankertz, C. Guger, A. Kübler, D. Mattia, J. d. R. Millán, F. Miralles, A. Nijholt, E. Opisso, N. Ramsey, P. Salomon, and G. R. Müller-Putz, "BNCI Horizon 2020: towards a roadmap for the BCI community," *Brain-Computer Interfaces*, vol. 2, no. 1, pp. 1–10, 2015.

[3] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Braincomputer interfaces for communication and control," *Clinical Neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.

[4] S. Machado, F. Araújo, F. Paes, B. Velasques, M. Cunha, H. Budde, L. F. Basile, R. Anghinah, O. Arias-Carrión, M. Cagy, R. Piedade, T. A. de Graaf, A. T. Sack, and P. Ribeiro, "EEG-based brain-computer interfaces: an overview of basic concepts and clinical applications in neurorehabilitation." *Reviews in the Neurosciences*, vol. 21, no. 6, pp. 451–68, 2010.

[5] C. Brunia, "Neural aspects of anticipatory behavior," *Acta Psychologica*, vol. 101, no. 2-3, pp. 213–242, 1999.

[6] J. Ibáñez, J. I. Serrano, M. D. del Castillo, J. Minguez, and J. L. Pons, "Predictive classification of self-paced upper-limb analytical movements with EEG," *Medical and Biological Engineering and Computing*, vol. 53, no. 11, pp. 1201–1210, 2015.

[7] A. Shakeel, M. S. Navid, M. N. Anwar, S. Mazhar, M. Jochumsen, and I. K. Niazi, "A review of techniques for detection of movement intention using movement-related cortical potentials," *Computational and Mathematical Methods in Medicine*, vol. 2015, no. 346217, 2015.

[8] T. Uktveris and V. Jusas, "Convolutional Neural Networks for Four-Class Motor Imagery Data Classification," in *Intelligent Distributed Computing XI. IDC 2017*, S. M. Ivanović M., Bdic C., Dix J., Jovanović Z., Malgeri M., Ed. Springer, Cham, 2018, pp. 185–197.

[9] K. Liao, R. Xiao, J. Gonzalez, and L. Ding, "Decoding Individual Finger Movements from One Hand Using Human EEG Signals," *PLoS ONE*, vol. 9, no. 1, p. e85192, 2014.

[10] M. Jochumsen, I. K. Niazi, K. Dremstrup, and E. N. Kamavuako, "Detecting and classifying three different hand movement types through electroencephalography recordings for neurorehabilitation," *Medical & Biological Engineering & Computing*, vol. 54, no. 10, pp. 1491–1501, 2015.

[11] K. Jerbi, J. R. Vidal, J. Mattout, E. Maby, F. Lecaignard, T. Ossandon, C. M. Hamamé, S. S. Dalal, R. Bouet, J. P. Lachaux, R. M. Leahy, S. Baillet, L. Garnero, C. Delpuech, and O. Bertrand, "Inferring hand movement kinematics from MEG, EEG and intracranial EEG: From brain-machine interfaces to motor rehabilitation," *Irbm*, vol. 32, no. 1, pp. 8–18, 2011.

[12] T. J. Bradberry, R. J. Gentili, and J. L. Contreras-Vidal, "Reconstructing Three-Dimensional Hand Movements from Noninvasive Electroencephalographic Signals," *Journal of Neuroscience*, vol. 30, no. 9, pp. 3432–3437, 2010.

[13] J. Lv, Y. Li, and Z. Gu, "Decoding hand movement velocity from electroencephalogram signals during a drawing task," *BioMedical Engineering OnLine*, vol. 9, no. 1, p. 64, 2010.

[14] J. H. Kim, F. Bießmann, and S. W. Lee, "Decoding three-dimensional trajectory of executed and imagined arm movements from electroencephalogram signals," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 23, no. 5, pp. 867–876, 2015.

[15] M. Jochumsen, I. K. Niazi, N. Mrachacz-Kersting, D. Farina, and K. Dremstrup, "Detection and classification of movement-related cortical potentials associated with task force and speed," *Journal of Neural Engineering*, vol. 10, no. 5, p. 056015, 2013.

[16] M. Jochumsen, I. Khan Niazi, D. Taylor, D. Farina, and K. Dremstrup, "Detecting and classifying movement-related cortical potentials associated with hand movements in healthy subjects and stroke patients from single-electrode, single-trial EEG," *Journal of Neural Engineering*, vol. 12, no. 5, p. 056013, 2015.

[17] M. Jochumsen, C. Rovsing, H. Rovsing, I. K. Niazi, K. Dremstrup, and E. N. Kamavuako, "Classification of Hand Grasp Kinetics and Types Using Movement-Related Cortical Potentials and EEG Rhythms," *Computational Intelligence and Neuroscience*, vol. 2017, pp. 1–8, 2017.

[18] V. Morash, O. Bai, S. Furlani, P. Lin, and M. Hallett, "Classifying EEG signals preceding right hand, left hand, tongue, and right foot movements and motor imageries," *Clinical Neurophysiology*, vol. 119, no. 11, pp. 2570–2578, 2008.

[19] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[20] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, "A review of classification algorithms for EEG-based braincomputer interfaces," *Journal of Neural Engineering*, vol. 4, no. 2, pp. 1–24, 2007.

[21] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "EEGNet: a compact convolutional neural network for EEG-based braincomputer interfaces," *Journal of Neural Engineering*, vol. 15, no. 5, p. 056013, 2018.

[22] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, M. Kudlur, J. Levenberg, R. Monga, S. Moore, D. G. Murray, B. Steiner, P. Tucker, V. Vasudevan, P. Warden, M. Wicke, Y. Yu, and X. Zheng, "TensorFlow: A system for large-scale machine learning," *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16)*, pp. 265–284, 2016, 1605.08695.

[23] F. Chollet and others, "Keras," Github repository https://github.com/fchollet/keras, 2015.

[24] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," *CoRR*, vol. abs/1610.02357, 2016.

[25] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, available in http://www.deeplearningbook.org.

[26] X.-X. Niu and C. Y. Suen, "A novel hybrid CNNSVM classifier for recognizing handwritten digits," *Pattern Recognition*, vol. 45, no. 4, pp. 1318–1325, 2012.

[27] P. Geethanjali, Y. K. Mohan, and J. Sen, "Time domain Feature extraction and classification of EEG data for Brain Computer Interface," in *2012 9th International Conference on Fuzzy Systems and Knowledge Discovery*, no. Fskd. IEEE, 2012, pp. 1136–1139.

[28] W. Chaovalitwongse, Y. Jeong, M. K. Jeong, S. Danish, and S. Wong, "Pattern Recognition Approaches for Identifying Subcortical Targets during Deep Brain Stimulation Surgery," *IEEE Intelligent Systems*, vol. 26, no. 5, pp. 54–63, 2011.

[29] V. Rajpurohit, S. F. Danish, E. L. Hargreaves, and S. Wong, "Optimizing computational feature sets for subthalamic nucleus localization in DBS surgery with feature selection," *Clinical Neurophysiology*, vol. 126, no. 5, pp. 975–982, 2015.

[30] C.-C. Chang and C.-J. Lin, "LIBSVM: A Library for Support Vector Machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, pp. 1–27, 2011.

[31] H. A. Agashe and J. L. Contreras-Vidal, "Reconstructing hand kinematics during reach to grasp movements from electroencephalographic signals," in *33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2011, pp. 5444–5447.

[32] G. Pfurtscheller, C. Neuper, A. Schlögl, and K. Lugger, "Separability of EEG signals recorded during right and left motor imagery using adaptive autoregressive parameters," *IEEE Transactions on Rehabilitation Engineering*, vol. 6, no. 3, pp. 316–325, 1998.

[33] E. Yorn-Tov and G. Inbar, "Selection of relevant features for classification of movements from single movement-related potentials using a genetic algorithm," in *2001 Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2001, pp. 1364–1366.

[34] G. Pfurtscheller, C. Brunner, A. Schlögl, and F. Lopes da Silva, "Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks," *NeuroImage*, vol. 31, no. 1, pp. 153–159, 2006.

[35] Y. Gu, K. Dremstrup, and D. Farina, "Single-trial discrimination of type and speed of wrist movements from EEG recordings," *Clinical Neurophysiology*, vol. 120, no. 8, pp. 1596–1600, 2009.

[36] A. Bubic, D. Y. Von Cramon, and R. Schubotz, "Prediction, cognition and the brain," *Frontiers in Human Neuroscience*, vol. 4, p. 25, 2010.

[37] N. Mrachacz-Kersting, S. R. Kristensen, I. K. Niazi, and D. Farina, "Precise temporal association between cortical potentials evoked by motor imagination and afference induces cortical plasticity," *The Journal of Physiology*, vol. 590, no. 7, pp. 1669–1682, 2012.

[38] J. W. Krakauer, "Motor learning: its relevance to stroke recovery and neurorehabilitation." *Current Opinion in Neurology*, vol. 19, no. 1, pp. 84–90, 2006.

[39] I. K. Niazi, N. Jiang, O. Tiberghien, J. F. Nielsen, K. Dremstrup, and D. Farina, "Detection of Movement Intention from Single-Trial Movement-Related Cortical Potentials," *Journal of Neural Engineering*, vol. 8, no. 6, p. 066009, 2011.

[40] G. Mueller-Putz, R. Scherer, C. Brunner, R. Leeb, and G. Pfurtscheller, "Better than random: A closer look on bci results," *International Journal of Bioelectromagnetism*, vol. 10, no. 1, pp. 52–55, 2008.

[41] H. Shibasaki and M. Hallett, "What is the Bereitschaftspotential?" *Clinical Neurophysiology*, vol. 117, no. 11, pp. 2341–2356, 2006.

[42] Y. Gu, O. F. Do Nascimento, M. F. Lucas, and D. Farina, "Identification of task parameters from movement-related cortical potentials," *Medical and Biological Engineering and Computing*, vol. 47, no. 12, pp. 1257–1264, 2009.

[43] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105.

[44] Z. Tang, C. Li, and S. Sun, "Single-trial EEG classification of motor imagery using deep convolutional neural networks," *Optik - International Journal for Light and Electron Optics*, vol. 130, pp. 11–18, 2017.

[45] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, "Deep learning with convolutional neural networks for EEG decoding and visualization," *Human Brain Mapping*, vol. 38, no. 11, pp. 5391–5420, 2017.

[46] C.-v. Soare, "Brain Computer Interface using Machine Learning," in *13th International Conference on Human Computer Interaction, RoCHI*, A. Iftene and J. Vanderdonckt, Eds. Iasi, Romania: Matrix Rom, 2016, pp. 65–68.

[47] Y. Lu, H. Jiang, and W. Liu, "Classification of EEG Signal by STFT-CNN Framework: Identification of Right-/left-hand Motor Imagination in BCI Systems," in *The 7th International Conference on Computer Engineering and Networks*, Shanghai, China, 2017.

[48] R. Prashanth, S. Dutta Roy, P. K. Mandal, and S. Ghosh, "High-Accuracy Detection of Early Parkinson's Disease through Multimodal Features and Machine Learning," *International Journal of Medical Informatics*, vol. 90, pp. 13–21, 2016.

[49] H. B. Li, W. Wang, H. W. Ding, and J. Dong, "Trees Weighting Random Forest Method for Classifying High-Dimensional Noisy Data," in *2010 IEEE 7th International Conference on E-Business Engineering*. IEEE, nov 2010, pp. 160–163.

[50] O. Bai, P. Lin, S. Vorbach, J. Li, S. Furlani, and M. Hallett, "Exploration of computational methods for classification of movement intention during human voluntary movement from single trial EEG," *Clinical Neurophysiology*, vol. 118, no. 12, pp. 2637–2655, 2007.

[51] E. Lew, R. Chavarriaga, S. Silvoni, and J. d. R. Millán, "Detection of self-paced reaching movement intention from EEG signals," *Frontiers in Neuroengineering*, vol. 5, no. 00013, 2012.