10.3217/978-3-99161-014-4-011

HIGH-PERFORMANCE NEURAL DECODING OF 14 DUTCH KEYWORDS

Elena C. Offenberg[1], Julia Berezutskaya[1], Zachary V. Freudenburg[1] and Nick Ramsey[1]

[1] Brain Center, University Medical Center Utrecht, Utrecht, Netherlands

E-mail: e.c.offenberg@umcutrecht.nl

ABSTRACT: Brain-computer interfaces (BCIs) can help people with locked-in syndrome to communicate. While continuous speech decoding can be used in everyday communication, navigating a computer menu or interacting with external devices may be easier and more reliable using a small set of distinct command keywords.
In this preliminary study, two able-bodied epilepsy patients, temporarily implanted with high-density electrocorticography (ECoG) electrodes, spoke 14 potential keywords out loud in Dutch. With optimized Support Vector Machines (SVMs), the maximal decoding accuracy reached was a median of 93.3% for 50 repetitions per word (practical chance level 9.6%). We also identified that a minimum of 30 repetitions was needed to achieve this result, and determined that the most relevant electrodes for decoding were on the ventral sensorimotor cortex, close to the central sulcus.

INTRODUCTION

People with a neurodegenerative disease, such as amyotrophic lateral sclerosis (ALS), can over time lose their ability to communicate verbally. Such a loss of communication has a direct negative impact on their perceived quality of life [1]. Fortunately, this effect can be assuaged: using communication devices has been shown to have a positive impact on quality of life and mood in dysarthric people with ALS [2].
Recently, there have been advancements in large-vocabulary decoding using brain-computer interfaces (BCIs) [3,4]. However, these large language models have not yet been shown to work in a home-use scenario. For daily communication or for controlling the home environment, a more robust and stable solution may be preferred. A very robust solution, using high gamma signals for one-dimensional cursor control, has been utilized successfully over several years [5]. However, such one-dimensional control limits the agency of the user. The next step allowing for more complex computer control could be the use of a limited set of keywords. For example, using six-keyword navigation with commands “up”, “down”, “left”, “right”, “enter” and “back” decoded from ECoG signals has recently demonstrated reliable control of computer menus and environment at high accuracy (median accuracy of 90.59%) over several months [6]. However, scaling up the individual word decoding from ECoG has proven rather challenging due to inter-subject variability, limited amounts of data and inherent limitations of decoding from brain signals. Further recent studies of keyword decoding from brain signals achieved 47.1% accuracy decoding 50 words in a paralyzed person with anarthria [7], 74.1% accuracy in a person with vocal paralysis [8], and 92%–100% decoding accuracy of 12 words in able-bodied participants [9].
One very relevant question both for researchers working with limited time and for BCI patients wanting to utilize their assistive devices with as little delay as possible is how much data is necessary to reach an acceptable decoding accuracy. Current decoding attempts are often limited by the small amounts of data researchers are able to collect with transient ECoG recordings. Therefore, determining the minimum amount of data necessary for satisfying decoding accuracies is one of our main goals.
In this study, we acquired high-density ECoG data from two subjects who spoke 14 Dutch words out loud, with S1 repeating each word 50 times. Using this data, we investigated three specific questions: 1. What is the highest accuracy of decoding 14 individual words? 2. How much data is needed to reach this accuracy, and 3. Which cortical areas are relevant for decoding?
We found that an accuracy of 93.3% (chance level 9.6%) could be reached with optimized SVMs with only 30 repetitions per word and that the electrodes contributing to the decoding performance the most were located on the postcentral gyrus of the ventral sensorimotor cortex, close to the central sulcus.

MATERIALS AND METHODS

Two human subjects S1 and S2 (1 male, 1 female, 26 and 46 years old, respectively) with medication-resistant epilepsy were implanted with 32-electrode high density (HD-)ECoG grids with platinum-iridium electrodes, a 4 mm inter-electrode distance and 1 mm exposed diameter. The grids were located on the left hemisphere covering the ventral sensorimotor cortex. Participant S1 had a previous tissue resection in the left temporal cortex. The study was approved by the Medical Ethical Committee of the University Medical Center Utrecht in accordance with the Declaration of Helsinki (2013). The subjects gave written informed consent to participate in research tasks.

Published by
Verlag der Technischen Universität Graz

 10.3217/978-3-99161-014-4-011

The 14 words were candidates for navigational words in Dutch, namely (in alphabetical order): "beneden" (down), "boven" (up), "kiezen" (choose), "links" (left), "noord" (north), "omhoog" (upwards), "omlaag" (downwards), "oost" (east), "rechts" (right), "selecteer" (select), "terug" (back), "verwijder" (“remove”), "west" (west) and "zuid" (south). During each repetition of the task, the 14 words and “-” (for rest trials, referred to as “rest” hereafter) were shown 5 times each in random order and read aloud by the participant. For S1, the task was repeated 10 times over the course of four days, resulting in a total number of 10*5 = 50 repetitions per word. For S2, one run was recorded, resulting in 5 repetitions of each word. HD-ECoG data was recorded using a Micromed system at 2048 Hz. Simultaneously, microphone data was recorded to determine the voice onset times. Voice onsets were defined manually as the first moment during which speech could be audibly perceived. In S1, one trial of the word “terug” was excluded from further analysis, since the subject did not say it during this trial.

Data pre-processing consisted of notch filtering of line noise (50 Hz) and its harmonics, common average re-referencing, and high frequency band (HFB) component extraction (70-170 Hz) using a Morlet wavelet decomposition in 1 Hz frequency bins, implemented via MNE-Python [10]. The high-frequency components were then averaged across frequencies, log-transformed and downsampled to 100 Hz.

The HFB signals of each run were “re-calibrated”, i.e. normalized individually per run using the mean and standard deviation of a 2-second rest period prior to the beginning of the task. The signals were then concatenated across runs and split into trials of 0 to 1 second after each voice onset time. Extending the trial length and including signals from before the voice onset, namely from 0.5 seconds before to 2 seconds after, did not change the accuracy results and was therefore not further pursued.

Due to the limited sample size, theoretical chance levels and practical chance levels differ [11]. Therefore, we used a binomial cumulative distribution to derive statistical significance thresholds for the obtained accuracies [11]. The practical chance levels were set at $p < 10^{-3}$ for the given sample sizes.

One repetition of every word was used as validation, and another as part of the test set, resulting in as many folds as there were repetitions per word. The flattened trial data (vectorized electrodes x time-points) was used to train an optimized Support Vector Machine (SVM) with a linear kernel in a one-vs-one approach with leave-one-group-out nested cross-validation. In the inner loop, the SVM regularization parameter was optimized using an automatic hyperparameter selection library Optuna [12], while the outer loop was necessary for cross-validating the classification results.

Since SVMs do not inherently provide probability estimates, the class membership probability estimates for the SVM were calculated with Scikit-learn [13], which uses Platt scaling and five-fold cross-validation. The electrode weights were determined by the L2 norm of the respective coefficients in the trained SVM. As per calculation of the L2 norm, we summed over the time dimension. The electrodes with the largest absolute classifier weights have the biggest impact on the classification.

RESULTS

Accuracy and Misclassified Trials

In S1, when trained and tested using all 50 repetitions per word or rest trial, the SVM reached a median accuracy of 93.3 ± 6.7% across folds. For S2 with 5 repetitions, a median accuracy of 73.3 ± 6.7% across folds was reached. The practical chance levels were 9.6% and 17.3% for S1 and S2, respectively - thus, both results were well above chance.

Not performing the re-calibration, which normalized the HFB data per run for S1, did not change the decoding accuracy.

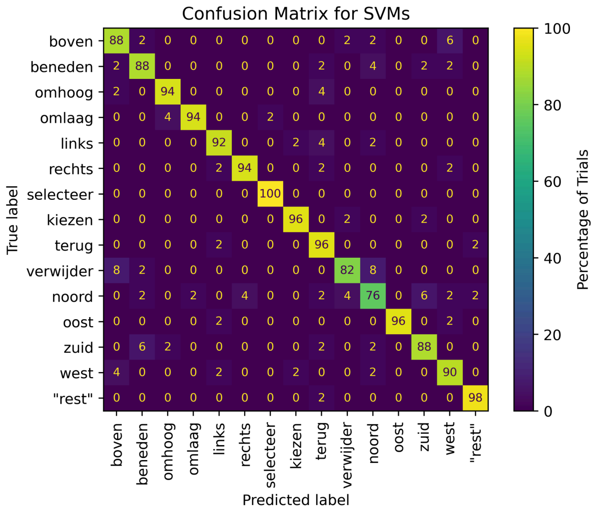

Figure 1: Normalized confusion matrix for S1 for 14 words and rest, with 50 repetitions for every word but “terug” (49 repetitions).

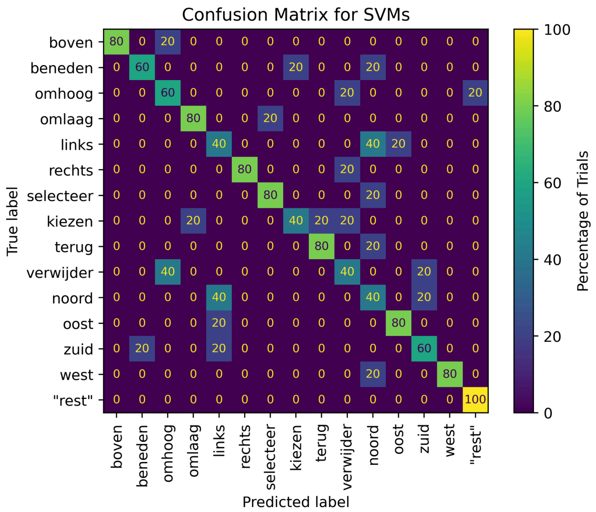

Figure 2: Normalized confusion matrix for S2 for 14 words and rest, with 5 repetitions for every word.

Published by
Verlag der Technischen Universität Graz

10.3217/978-3-99161-014-4-011

As can be seen in Figures 1 and 2, the prediction accuracy varied between different words. Although for S1, the highest accuracy was a 3-syllable word, “selecteer”, there was no significant correlation between decoding accuracy and length of the words (Pearson coefficient of 0.16, *p*-value 0.90).
To see how uncertain the trained SVM for S1 was about its predictions, we visualized the calculated class probabilities for each trial, sorted by words (Figure 3). For trials classified correctly (in gray), the probability often peaked at the near-maximum for the target class and was quite low for all non-target classes. For the misclassified trials (in red), there were often several probability peaks that included the target class.
Quantitatively, across all misclassified trials, the target class was assigned the second or third highest probability in 78% of cases, showing that even in trials with incorrect predictions, correct patterns were still being picked up by the SVM. In total, in 98.3% of trials, the correct class was among the top three predictions.

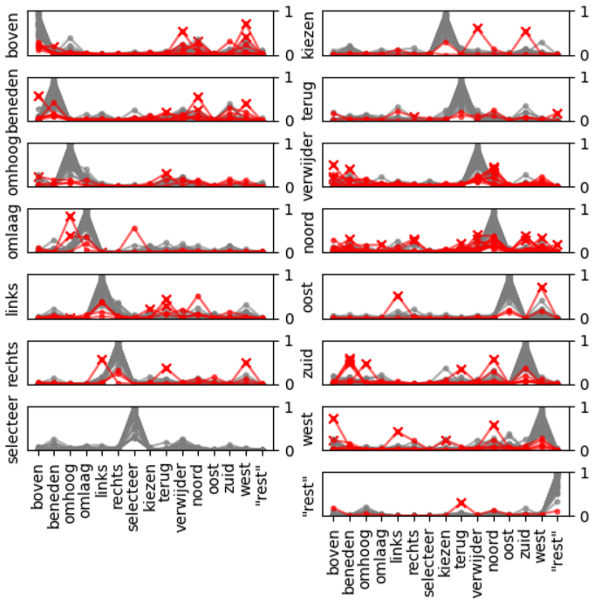

Figure 3: The probability distributions of the SVM predictions for each word and rest-trials for S1. The target words are on the y axis, the predicted words on the x axis. Correctly classified trials are plotted in gray, the misclassified trials are plotted in red, and the incorrect predictions are marked with red crosses.

Amount of Data Necessary

How much data is necessary in order to reach an acceptable decoding accuracy? For S1, 50 repetitions per word were recorded across 10 runs over the course of several days (5 word repetitions per run). When trained with the data from successive runs cumulatively, the decoding accuracy increased (Figure 4). This analysis uses mean accuracy values instead of medians since the mean provides a smoother statistic over the number of repetitions.
A mean accuracy of 68.0 ± 9.8% for S1 and 66.6 ± 11.2% for S2 was reached after the first run. A similar result was achieved for the calculations based on median values.
Notably for S1, even after the decoding accuracy reached the ceiling, the variance in performance decreased as more trials were added, suggesting a further stabilization of the decoding performance.

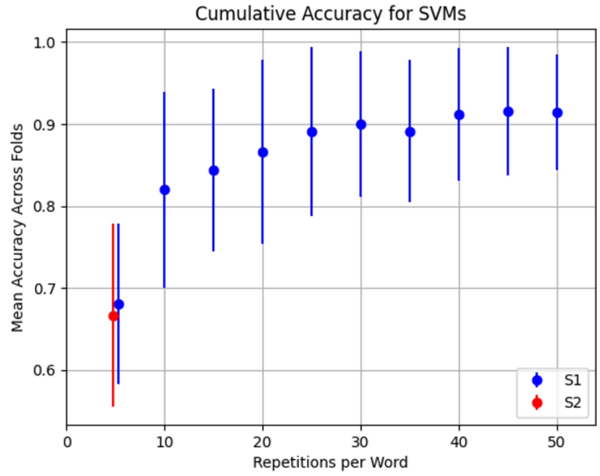

Figure 4: Cumulative accuracy for SVMs as a function of number of word repetitions.

Electrodes Relevant for Decoding

Not all of the 32 electrodes in each grid contributed to the decoding performance in the same proportion.
In Figures 5 and 6, the positions of the grids for S1 and S2 on the left hemisphere are shown. In addition, we visualize the normalized absolute SVM weights to highlight electrodes most relevant for the decoding performance. For both subjects, the most relevant electrodes were located close to the central sulcus on the ventral sensorimotor cortex, an area associated with the cortical control of articulation [14]. Importantly, electrodes with highest SVM weights were located on the postcentral gyrus.

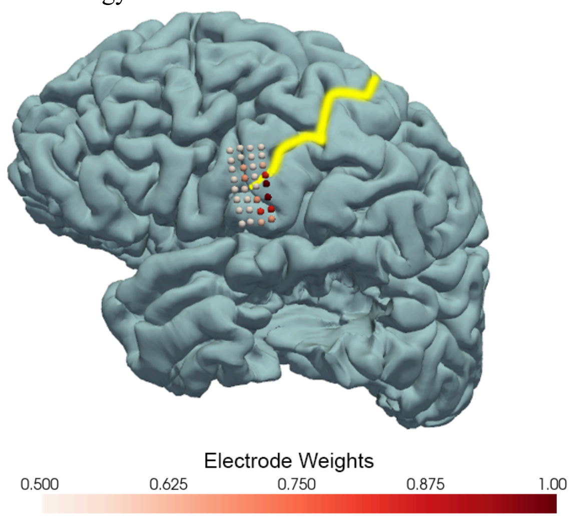

Figure 5: The electrode grid of S1, with darker colors corresponding to higher normalized SVM weights of the electrode. The central sulcus is highlighted in yellow.

Published by
Verlag der Technischen Universität Graz

10.3217/978-3-99161-014-4-011

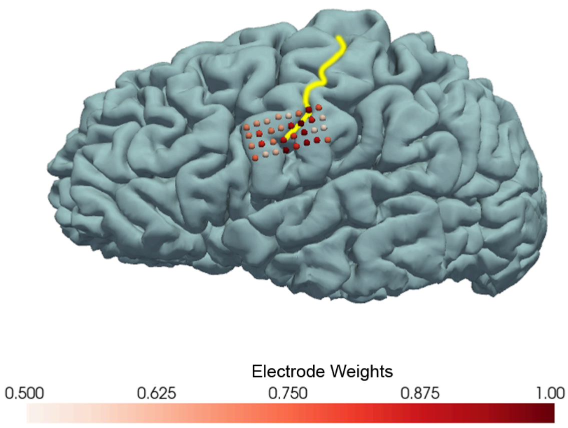

Figure 6: The electrode grid of S2, with darker colors corresponding to higher normalized SVM weights of the electrode. The central sulcus is highlighted in yellow.

DISCUSSION

In this study, we set out to answer three questions concerning decoding of individual words from HD-ECoG, namely: what is the highest decoding accuracy, what is the minimal amount of data necessary and what are the locations of the most informative electrodes.

We were able to reach a median accuracy of 93.3% for decoding 14 words with 50 repetitions per word. Upon closer inspection, we saw that in over 98% of all trials, the correct words were among the top three SVM predictions. This result could be used to improve predictive performance even further, for example by giving the user a list of top-ranking alternatives or combining top predictions with statistics of previous use to enable quick corrections by the user in case of misclassification.

For controlling external devices, an accuracy of command identification of at least 90% has been determined as the acceptable threshold in a survey among ALS-patients [15]. This threshold was reached after 6 runs, corresponding to 30 repetitions per word, see Figure 4.

For both subjects, we obtained a mean accuracy of over 66% with only five repetitions per word, which is well above chance. While this accuracy does not yet approach the threshold desired for long-term use, a model trained on only five repetitions could already be used to provide immediate feedback to the user while continuing to update the decoder in the background. In our experience, participants find tasks with feedback more engaging, leading to a higher quantity of data collection. The concrete influence of such early feedback on both motivation and performance could be further investigated in a future work.

The most relevant electrodes in both subjects were located on the dorsal part of the ventral sensorimotor cortex, with the highest weights found closest to the central sulcus. This mirrors the results of another ECoG-study for keyword decoding [6]. Interestingly, most contributing electrodes seemed to be located on the postcentral gyrus – the somatosensory area of the brain. Since our subjects were able-bodied people, one might attribute this result to the sensory feedback from mouth movements. However, motor decoding from the somatosensory cortex has previously been shown both for amputees [16] and people with paralysis due to ALS [17], suggesting that there is information about movement in somatosensory areas even in the absence of direct sensory feedback. This is in line with work on an efference copy of voluntary movements in the somatosensory cortex [18, 19]. It remains to be seen whether comparable decoding performance from the sensorimotor cortex can be achieved in locked-in individuals using a BCI.

One important difference between our study subjects and future locked-in users is that the voice onset times will not be available as ground truth for training the classifiers. As an alternative, different methods of extracting activity onset directly from brain signals have already been proposed and used [4,6].

CONCLUSION

In the present study, we achieved a high accuracy of decoding 14 individual words from HD-ECoG brain activity recorded from the ventral sensorimotor cortex of two able-bodied subjects. For a subject with 50 repetitions per word, 30 repetitions per word were sufficient to reach a decoding accuracy of over 90%, and the most informative electrodes for both patients were located in the ventral postcentral gyrus.

ACKNOWLEDGEMENTS

This publication is part of the project Dutch Brain Interface Initiative (DBI2) with project number 024.005.022 of the research programme Gravitation, which is financed by the Dutch Ministry of Education, Culture, and Science (OCW) via the Dutch Research Council (NWO).

In addition, this project is funded by the European Union's HORIZON-EIC-2021-PATHFINDER CHALLENGES program under grant agreement No 101070939 and by the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract number 22.00198. Finally, the authors were supported by the National Institutes of Health under award number UH3NS114439 (NINDS).

REFERENCES

[1] Felgoise SH, Zaccheo V, Duff J, Simmons Z. Verbal communication impacts quality of life in patients with amyotrophic lateral sclerosis. Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration. 2016;17(3-4):179-183

[2] Körner S, Siniawski M, Kollewe K, Rath KJ, , Krampfl K, Zap, A et al. Speech therapy and

Published by
Verlag der Technischen Universität Graz

10.3217/978-3-99161-014-4-011

communication device: Impact on quality of life and mood in patients with amyotrophic lateral sclerosis. Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration. 2013;14(1):20-25.
[3] Willett FR, Kunz EM, Fan C, Avansino DT, Wilson GH, Choi EY et al. A high-performance speech neuroprosthesis. Nature.2023;620(7976):1031-1036.
[4] Metzger SL, Littlejohn KT, Silva AB, Moses DA, Seaton MP, Wang R et al. A high-performance neuroprosthesis for speech decoding and avatar control. Nature. 2023;620(7976):1037-1046.
[5] Pels EGM, Aarnoutse EJ, Leinders S, Freudenburg ZV, Branco MP, van der Vijgh BH et al. Stability of a chronic implanted brain-computer interface in late-stage amyotrophic lateral sclerosis. Clinical Neurophysiology. 2019;130(10):1798-1803.
[6] Luo S, Angrick M, Coogan C, Candrea DN, Wyse-Sookoo K, Shah S, et al. Stable Decoding from a Speech BCI Enables Control for an Individual with ALS without Recalibration for 3 Months. Advanced Science. 2023;10(35):2304853.
[7] Moses DA, Metzger SL, Liu JR, Anumanchipalli GK, Makin JG, Sun PF et al. Neuroprosthesis for decoding speech in a paralyzed person with anarthria. New England Journal of Medicine. 2021;385(3):217-227.
[8] Metzger SL, Liu JR, Moses DA, Dougherty ME, Seaton MP, Littlejohn KT et al. Generalizable spelling using a speech neuroprosthesis in an individual with severe limb and vocal paralysis. Nat Commun. 2022;13(1):6510.
[9] Berezutskaya J, Freudenburg ZV, Vansteensel MJ, Aarnoutse EJ, Ramsey NF, Van Gerven MAJ. Direct speech reconstruction from sensorimotor brain activity with optimized deep learning models. J Neural Eng. Published online July 19, 2023.
[10] Gramfort A, Luessi M, Larson E, Engeman DA, Strohmeier D, Brodbeck C, et al. MEG and EEG data analysis with MNE-Python. Front Neurosci. 2013;7.
[11] Combrisson E, Jerbi K. Exceeding chance level by chance: The caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy. Journal of Neuroscience Methods. 2015;250:126-136.
[12] Akiba T, Sano S, Yanase T, Ohta T, Koyama M. Optuna: A Next-generation Hyperparameter Optimization Framework. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. ACM; 2019:2623-2631.
[13] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O et al. Scikit-learn: Machine Learning in Python. Published online 2012. 12(85):2825−2830, 2011.
[14] Bouchard KE, Mesgarani N, Johnson K, Chang EF. Functional organization of human sensorimotor cortex for speech articulation. Nature. 2013;495(7441):327-332.
[15] Huggins JE, Wren PA, Gruis KL. What would brain-computer interface users want? Opinions and priorities of potential users with amyotrophic lateral sclerosis. Amyotrophic Lateral Sclerosis. 2011;12(5):318-324.
[16] Bruurmijn MLCM, Pereboom IPL, Vansteensel MJ, Raemaekers MAH, Ramsey NF. Preservation of hand movement representation in the sensorimotor areas of amputees. Brain. 2017;140(12):3166-3178.
[17] Leinders S, Vansteensel MJ, Piantoni G, et al. Using fMRI to localize target regions for implanted brain-computer interfaces in locked-in syndrome. Clinical Neurophysiology. 2023;155:1-15.
[18] Umeda T, Isa T, Nishimura Y. The somatosensory cortex receives information about motor output. Sci Adv. 2019;5(7):eaaw5388.
[19] Christensen MS, Lundbye-Jensen J, Geertsen SS, Petersen TH, Paulson OB, Nielsen JB. Premotor cortex modulates somatosensory cortex during voluntary movements without proprioceptive feedback. Nat Neurosci. 2007;10(4):417-419.

Published by
Verlag der Technischen Universität Graz