Novel and Efficient Hyperdimensional Encoding of Surface Electromyography Signals for Hand Gesture Recognition

Ancelin Salerno[a] and Sylvain Barraud[b]
CEA, LETI, Minatec campus, University Grenoble Alpes, 38000 Grenoble, France

Keywords: High-Dimensional Computing, Surface Electromyography Signal, Gesture Recognition, Encoding, Hypervectors, Ninapro Database.

Abstract: Gesture recognition has become a crucial component of human-computer interaction, with applications ranging from virtual reality to assistive technologies. This study explores Hyperdimensional Computing (HDC) as a powerful alternative to traditional machine learning techniques for real-time gesture recognition. HDC is known for its robustness and efficiency, enabling fast and accurate classification though the use of high-dimensional binary vectors. In this study, we introduce two key variants aimed at significantly improving the performance of gesture recognition: (1) an enhancement of item memory representation enabling a better gestures recognition, and (2) an advanced temporal encoding mechanism that captures the dynamic nature of gestures more efficiently. These modifications are evaluated using a benchmark dataset of surface electromyography (sEMG) signals, demonstrating significant improvements in both accuracy and computational efficiency.

1 INTRODUCTION

Recent advances in the recognition and classification of surface electromyography (sEMG) signals are opening up new opportunities in fields such as human-machine interfaces, robotic control, and augmented/virtual reality. These advances rely heavily on the accurate measurement of multi-channel surface EMG signals and the application of machine learning (ML) algorithms for gesture identification. However, deploying ML models on wearable edge devices presents both challenges and opportunities. While edge computing enables real-time processing with reduced latency and improved privacy through on-device data handling, machine learning models encounter significant challenges in addressing the variability of sEMG signals (Hudgins et al., 1993). Indeed, factors such as muscle fatigue, electrode displacement, changes in arm posture, and inter-subject/session variability can severely impact classification performance, limiting the robustness of conventional ML approaches.

Current neural network-based solutions, despite their potential, are resource-intensive, requiring large volumes of high-quality training data and incurring substantial computational and power demands. This makes their integration into embedded systems, particularly challenging for real-time gesture recognition (Hudgins et al., 1993; Benatti et al., 2014). To address these issues, we introduce **CompHD**; a novel hyperdimensional computing (HDC) framework designed for the efficient encoding and classification of sEMG signals in hand gesture recognition.

Unlike traditional HDC methods, which rely on random or continuous item memories for sequences encoding (Rahimi et al., 2016; Sgambato & Castellano, 2022), CompHD incorporates optimized hyperdimensional representations that allow for more efficient and accurate processing of gesture data. The HDC-based brain-inspired architecture offers several key advantages: it supports one-pass learning, reducing energy consumption and accelerating the learning process, while also being highly robust to noise and computational errors. Moreover, CompHD requires only a small training dataset to achieve competitive accuracy, making it well suited for low-power, real-time applications and most-importantly,

[a] https://orcid.org/0009-0008-6432-4270
[b] https://orcid.org/0000-0002-4334-9638

Salerno, A. and Barraud, S.
Novel and Efficient Hyperdimensional Encoding of Surface Electromyography Signals for Hand Gesture Recognition.
DOI: 10.5220/0013257500003911

In *Proceedings of the 18th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2025) - Volume 1*, pages 763-772
ISBN: 978-989-758-731-3; ISSN: 2184-4305

for on-subject training, as sEMG data are highly individual-specific.

For the first time, we present a comprehensive quantitative comparison between the CompHD algorithm and conventional machine learning models for sEMG-based gesture recognition, using publicly available datasets, including Ninapro—the largest and most comprehensive sEMG database. Our evaluation demonstrates that CompHD not only outperforms traditional methods in terms of robustness and computational efficiency, but also achieves competitive accuracy on this challenging benchmark.

2 DATABASE DESCRIPTION

This study utilizes several publicly available electromyography (EMG) datasets, including **Master** (Rahimi et al., 2016), **Pattern** (Lobov et al., 2018), **Ninapro DB1**, **DB4** and **DB5** (Atzori et al., 2012, 2014; Pizzolato et al., 2017; Wan et al., 2018). While these datasets are commonly used for gesture recognition tasks with various classifiers, we are the first study to systematically compare them using the same HDC model. The datasets differ in terms of the number of gestures, subjects, sampling frequencies and recording channels, providing a diverse and comprehensive foundation for developing and analyzing our HDC model for sEMG signals recognition. They offer an opportunity to explore different signal characteristics and patterns, thus facilitating the development and evaluation of machine learning models for gesture classification. The key characteristics of these datasets are summarized in Table 1.

Table 1: Number of gestures, subjects and channels considered for each dataset used in this work.

Database	Gestures number	Subjects number	Channels number
MASTER	5	5	4
PATTERN	7	37	8
NINAPRO DB1	52	27	10
NINAPRO DB4	52	10	12
NINAPRO DB5	52	10	16

The feature used in this work for gestures classification is the Mean Amplitude Value (MAV) which has been shown to be highly effective, providing both high accuracy and computationally efficiency (Scheme & Englehart, 2014).

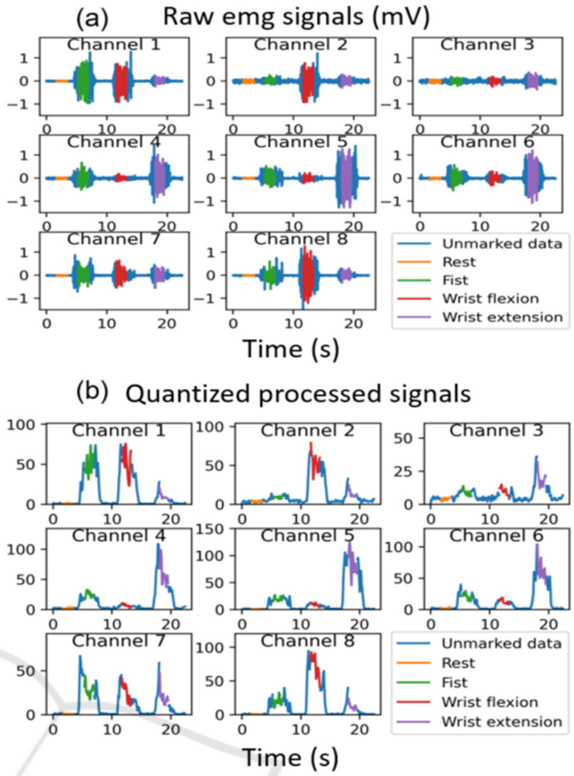


Figure 1: Example of raw (a) and preprocessed (b) sEMG signals of "Pattern" database. For further clarity, the sEMG signals measured on 8 electrodes are shown for only one repetition and only five gestures (including unmarked data).

The MAV is a widely used time-domain feature in sEMG signals analysis. It represents the average of the absolute values of the signal over a given time window. In this study, we use 100ms windows. The MAV is particularly effective at capturing muscle activation patterns and is less sensitive to noise compared to other features, making it well-suited for applications such as prosthetic control. In the context of sEMG data, where the frequency band of interest typically ranges from 10 to 500 Hz, the MAV is extracted following preprocessing steps such as filtering to remove artifacts and baseline drift. For the Ninapro DB1, this preprocessing can be performed analogically before digitization to ensure that only relevant muscle activation signals are captured. For the other datasets, MAV extraction is conducted using digital filters. This process involves summing the absolute values of the signal within the specified window and dividing by the number of samples, providing a robust measure of the signal's amplitude. The MAV is particularly effective in detecting changes in muscle activity levels, making it crucial for real-time EMG signal classification.

3 HIGH-DIMENSIONAL COMPUTING CLASSIFIER

The Hyperdimensional Computing (HDC) is an emerging computational paradigm inspired by the human brain's approach to processing information and provides a robust, efficient framework for managing data represented as high-dimensional vectors, called hypervectors (HVs) (Kanerva, 2009; Rahimi et al., 2017; Cohen & Widdows, 2015). The central concept is to encode information in high-dimensional vectors that capture rich and complex patterns. Some mathematical preserving operations are applied to process or retrieve information. As information is distributed across all dimensions of a hypervector, it is less susceptible to interference. This holistic characteristic inherently enhances the system's robustness to noise and partial information loss, as individual bit errors are unlikely to compromise the encoded meaning.

3.1 Conventional HDC Encoding

This study specifically focuses on binary HVs, as binary HDC provides significant energy efficiency benefits and is particularly well-suited for embedded device (Chen et al., 2022) or hardware-based in-memory implementations (Abhijith & Shekhar, 2019; Benatti et al., 2017; Karunaratne, et al., 2020, 2021; Li et al., 2016). Figure 2 illustrates the main components of the HDC algorithm.

The newly highlighted areas correspond to the specific blocks we have targeted to enhance the model's performance. At each stage of the spatiotemporal encoding process, binary operations are performed on binary hypervectors. These operations are dimension-preserving, ensuring that the dimensionality (**DIM**) remains consistent across both the input and output vectors. The core operations in hyperdimensional encoding — binding and bundling (Rahimi et al., 2017, Cohen & Widdows, 2015) — enable the creation and manipulation of complex symbolic structures.

- **Binding:** This operation combines two hypervectors (HVs) to create a new one that is distinct from both. Binding is performed using element-wise multiplication (or XOR for binary vectors), ensuring that the resulting hypervector is unique while still preserving information from both of the original hypervectors.

- **Bundling:** This operation combines multiple hypervectors into a single hypervector that reflects the common features of all the elements in the set.

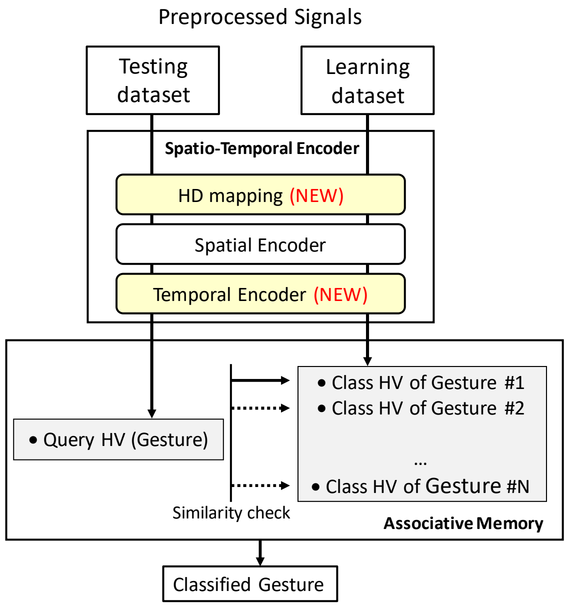


Figure 2: Overview of HDC classifier for gesture recognition accounting for the new HD mapping and temporal encoder proposed in this work.

Bundling captures similarities among different vectors representing related information and is implemented using an element-wise majority function across the set of hypervectors.

The bundling operation is essential for linking temporal information across multiple hypervectors from different timestamps. The hypervector resulting from bundling n consecutive temporal hypervectors is referred to as n-gram. In hyperdimensional computing-based algorithms, n is a hyperparameter commonly referred to as **NGRAM**. Both **DIM** and **NGRAM** are critical hyperparameters that significantly influence the model's performance.

NGRAM captures contextual relationships by encoding local dependencies, with larger n-grams offering richer contextual information at the expense of increased complexity. **DIM**, on the other hand, determines the dimensionality of the hypervectors used to represent data in high-dimensional space. It typically ranges from several thousand to as many as 10,000 elements. This high dimensionality ensures that randomly generated vectors are almost orthogonal, minimizing overlap and enhancing distinguishability. A higher **DIM** typically offers greater capacity for encoding and distinguishing patterns. However, it requires more computational resources. Striking the right balance between these parameters is key to optimizing accuracy, efficiency, and generalization in hyperdimensional computing systems.

3.2 Our Spatio-Temporal Encoder

The first modification we propose involves the spatiotemporal encoder. In surface electromyography (sEMG) signal processing, the sequence order of input data is generally less critical than in tasks like text or speech recognition, where the position of each element in a sequence significantly contributes to its meaning. sEMG signals capture the electrical activity of muscles during contractions, which typically produce patterns based on muscle activation levels rather than strict temporal order. As a result, models for sEMG analysis can often focus more on feature extraction and less on temporal dependencies compared to language processing tasks. This characteristic allows for the use of simpler preprocessing techniques and makes certain machine learning approaches, such as convolutional neural networks (CNNs), particularly effective for interpreting sEMG data. Instead of using the binding operation to encode sequences—preserving both the values and their order—we propose employing a bundling operation to compute the mean vector over a temporal window. This method enhances the model's robustness to noisy sEMG signals, where amplitude variability across time samples could otherwise impair sequence encoding and compromise signal interpretation.

3.3 Our Novel Composite Encoding (CompHD)

In text recognition (Abhijith & Shekhar, 2019; Cohen & Widdows, 2015; Rahimi et al., 2017), mapping data to high-dimensional hypervectors that are orthogonal ensures that letters or symbols are equidistant from each other, preventing any bias in the representation of specific characters. This orthogonality is especially advantageous for discrete classification tasks, where each character is treated as a distinct entity. However, for continuous data such as surface electromyographic (sEMG) signals, it is crucial to capture subtle variations in signal amplitude and frequency. In this case, projecting the data into continuous hypervectors offers a more effective representation of the continuous nature of sEMG amplitudes (Cohen & Widdows, 2015; Rahimi et al., 2016/2017; Salerno & Barraud, 2024).

In this case, the distance between hypervectors should reflect the magnitude of the data, allowing for smoother transitions and a more nuanced encoding of continuous signal variations. This continuous projection is better suited for tasks that require fine-grained discrimination between data points, such as analyzing muscle activation patterns in sEMG. Thus, while orthogonal hypervectors are optimal for discrete data like text, continuous hypervectors are more appropriate for representing dynamic, continuous signals like sEMG.

Mapping data into hypervectors using a combination of random (i.e., orthogonal) hypervectors and continuous hypervectors (Figure 3) offers significant advantages for pattern recognition, particularly in tasks where data points are highly similar to previously encountered examples. This approach leverages the high-dimensional properties of hypervectors to encode information in a way that enables fine distinctions between closely related data, while still preserving the ability to generalize to broader patterns.

By associating the distance between hypervectors with differences in their index positions, this method enhances the recognition of subtle variations in data, making it highly effective for tasks that require both precise identification of known patterns and flexibility in adapting to new, unseen data. This dual capability is especially valuable in applications like biomedical signal processing, where small variations can be critical, yet robust generalization is essential to handle variability across subjects or conditions

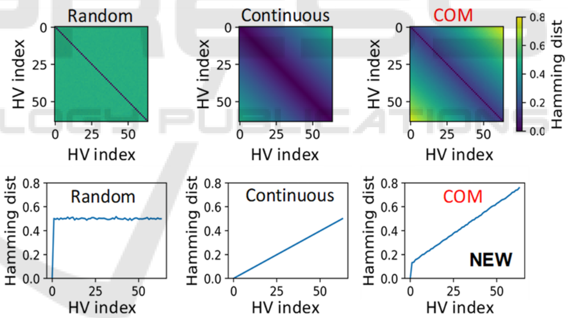


Figure 3: HD mapping scheme used in this work. A new HD mapping (COM) is proposed to encode the preprocessed and quantized sEMG signals into binary hypervectors. The hamming distance between two HVs is calculated and compared to conventional Random and Continuous mapping.

3.4 Leave-P-Groups-out Cross-Validation

In this section, we outline the specific training and testing procedures used for CompHD, including dataset partitioning, cross-validation techniques, and performance evaluation metrics to assess model accuracy. To evaluate and compare the gesture recognition rates for different training sizes while maintaining temporal integrity of gestures repetitions, the Leave-P-Groups-Out Cross-Validation (LPGO) is an excellent choice. LPGO provides a robust approach for evaluating machine-learning models in scenarios where data is structured into distinct groups, such as gesture repetitions in human activity recognition. This technique systematically partitions the dataset into smaller groups, such as repetitions, ensuring that the temporal structure of the data is preserved during validation. During the inference phase, the model trained on the remaining groups is tested on P small groups, known as the left-out groups. The LPGO cross-validation process is illustrated in Figure 4.

In the context of finite sets of gesture repetitions, LPGO ensures that the model is evaluated exclusively on unseen repetitions, as it is reinitialized at each iteration. Thus, LPGO closely simulates real-world variability in user performance. This technique is particularly valuable for applications like prosthetic control or gesture recognition, where consistent performance across repetitions is crucial. By assessing how well the model generalizes across different instances of the same gesture, LPGO provides insights into its robustness.

Figure 4: Protocol used for learning and testing phases. Leave-P-groups-Out Cross-Validation method is applied to use all the repetitions for both training and testing. In this example P=1, which corresponds to a simple 5-fold Cross-Validation.

4 EMG-BASED GESTURES RECOGNITION

Using this training and testing protocol, we evaluated the performance of three state-of-the-art hypervector memory models: Random Item Memory (RIM), Continuous Item Memory (CIM), and our composite model, CompHD (COM), across the databases listed in Table 1. This comparison allows us to assess the strengths and limitations of each approach in the context of gesture recognition, highlighting how CompHD leverages the advantages of both RIM and CIM to achieve enhanced performance. The results are presented in order from the simplest (Master) to the most complex database (Ninapro). This organization allows us to clearly demonstrate the model's effectiveness and robustness, highlighting its adaptability across different gesture recognition scenarios.

In Table 2, we present a summary of the database gesture information and the optimal NGRAM value, one of the key hyperparameters in our model. The tuning of these hyperparameters involves identifying a set of optimal parameter values for the HDC model, aimed at maximizing performance (i.e., recognition accuracy) while minimizing memory resource allocation. Higher-dimensional hypervectors are critical as they help prevent data loss during encoding. Several hyperparameters in HDC require fine-tuning, including the NGRAM value, which determines the number of bundled hypervectors during temporal encoding, and the dimensionality (DIM) of the hypervectors.

To optimize these hyperparameters, we conducted an extensive Grid search across different NGRAM values and dimensionalities. The tuning process involved computing the average performance score for various combinations of NGRAM and DIM. Specifically, we evaluated these parameters using a 2D heatmap (see ***Appendix***), which allowed us to visualize the relationship between the dimensionality and the NGRAM value. From this analysis, we identified the optimal parameter set that maximized the performance score for a given database, ensuring the best possible recognition rate. These selected values were then applied consistently across all subjects in the database.

Table 2: Average gesture duration, optimal NGRAM and dimension of HVs used for each database.

Database	Average gesture duration (a.u.)	Optimal NGRAM (a.u.)	Dimension of HVs
MASTER	26.5	25	8192
PATTERN	17	9	16384
NINAPRO DB1	25.08	41	16384
NINAPRO DB4	37.06	51	16384
NINAPRO DB5	25.75	43	16384

4.1 Master Database

The first database, known as the Master database, contains ten repetitions of five gestures performed by four subjects (with one subject excluded due to inconsistent repetitions). In our study, the CompHD model consistently outperformed both CIM and RIM in classification accuracy across all training sizes. Notably, CompHD achieved higher accuracy rates at each incremental training size, demonstrating its robustness on both small and large datasets. For some subjects, CompHD's classification accuracy reached an impressive 99.9%, further highlighting its precision in delivering accurate classifications across subjects. This consistent advantage at every stage of training establishes CompHD as a significantly more effective model than CIM and RIM for achieving high-accuracy classifications under various training conditions. Remarkably, CompHD requires only 40% of the training data to outperform both CIM and RIM, even when larger training sizes (up to 90%) are used. This underscores CompHD's superior efficiency and effectiveness in gesture recognition tasks.

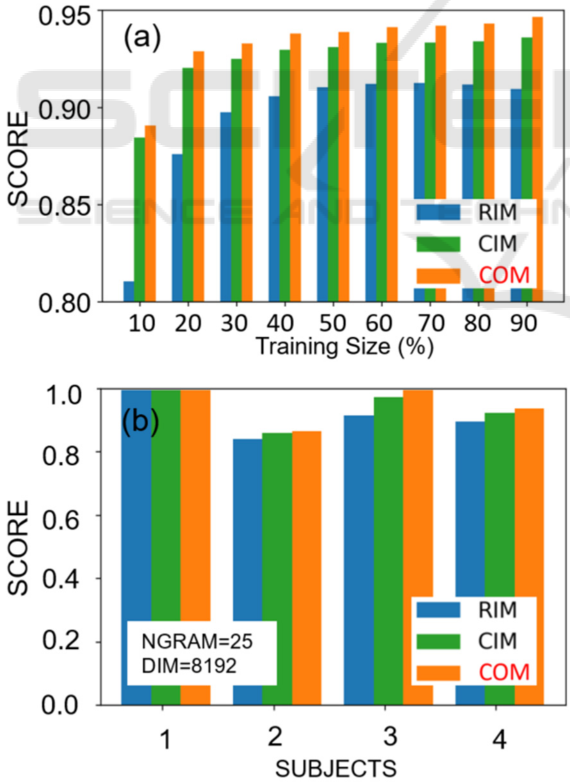


Figure 5: **Master database**. (a) Classification accuracy was averaged across all subjects and gestures for different training sizes. (b) Averaged accuracy achieved per subject across all five gestures using the largest training size.

4.2 Pattern Database

The second database, known as the Pattern database, includes four repetitions of seven gestures performed by 36 subjects. This dataset was recorded using eight electrodes, providing a rich data source for analyzing gesture recognition performance across diverse subjects.

As with the previous database, the CompHD model consistently outperformed both CIM and RIM in classification accuracy across all training sizes. Notably, CompHD achieved higher accuracy rates at each incremental training size, demonstrating its robustness on both smaller and larger datasets. For some subjects, CompHD's classification accuracy reached an impressive 100%, further emphasizing its effectiveness in delivering precise classifications across different subjects.

This consistent advantage at every stage of training reinforces CompHD as a significantly more effective model than CIM and RIM for achieving high-accuracy classifications under varied training conditions.

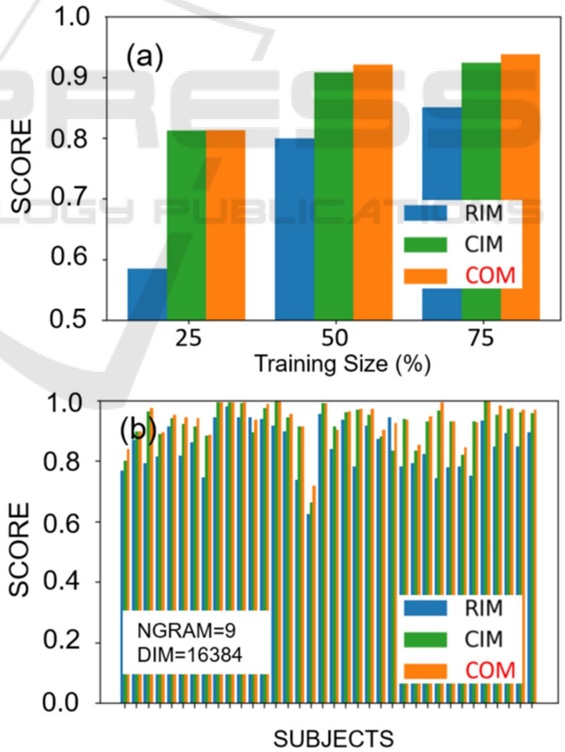


Figure 6: **Pattern database**. (a) The classification accuracy, averaged across all subjects and gestures, was evaluated for different training sizes. (b) Averaged accuracy achieved per subject across all seven gestures using the largest training size.

4.3 Ninapro Database 1

The third database tested, Ninapro DB1, includes ten repetitions of 52 gestures performed by 27 subjects. This dataset was recorded using 8 electrodes.

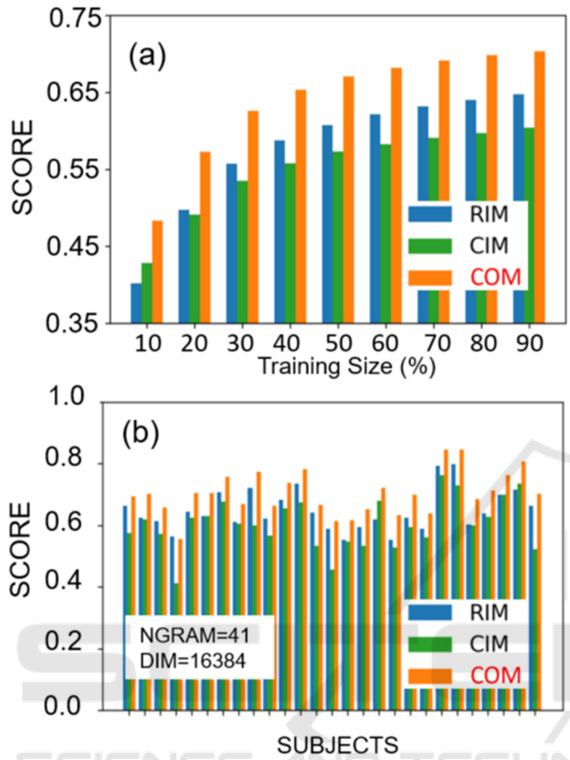


Figure 7: **Ninapro DB1**. (a) Classification accuracy was averaged across all subjects and gestures for different training sizes. (b) Averaged accuracy achieved by each subject across all 52 gestures using the largest training size.

While RIM outperforms CIM with larger training sizes, the reverse holds true for smaller sample sizes. Despite this, the CompHD model consistently demonstrated superior classification performance across all training sizes compared to both CIM and RIM. While the average score amongst subjects is approximately 70%.

For some subjects, CompHD's classification accuracy reached an impressive 85%, further highlighting its precision in delivering accurate classifications across subjects, even with more complex datasets. Notably, CompHD requires only 40% of the training data to outperform both CIM and RIM, even when larger training sizes (up to 90%) are used.

4.4 Ninapro Database 4

The fourth database tested, Ninapro DB4, includes six repetitions of 52 gestures performed by 10 subjects, recorded using 10 electrodes. As with the previous databases, the CompHD model consistently outperformed both CIM and RIM in classification accuracy across all training sizes.

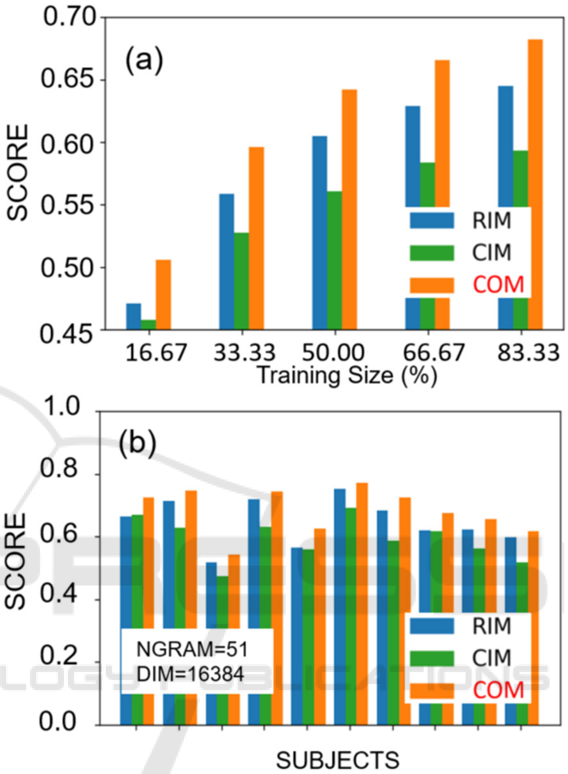


Figure 8: **Ninapro DB4**. (a) Classification accuracy was averaged across all subjects and gestures for different training sizes. (b) Averaged accuracy achieved by each subject across all 52 gestures with the largest training size.

For some subjects, CompHD's classification accuracy reached nearly 80%, further emphasizing its ability to deliver precise classifications across different subjects. Notably, CompHD requires only 50% of the training data to achieve similar accuracies as both CIM and RIM, even when larger training sizes (up to 83.33%) are used.

The random Item memory (RIM) outperforms the Continuous one (CIM) only on the Ninapro database 1 and 4. However, our new Composite Item Memory (COM) is robust across all database and outperforms both of the traditional item memories used.

4.5 Ninapro Database 5

The fifth database: Ninapro DB5 includes six repetitions of 52 gestures performed by 10 subjects and was recorded using 16 electrodes.

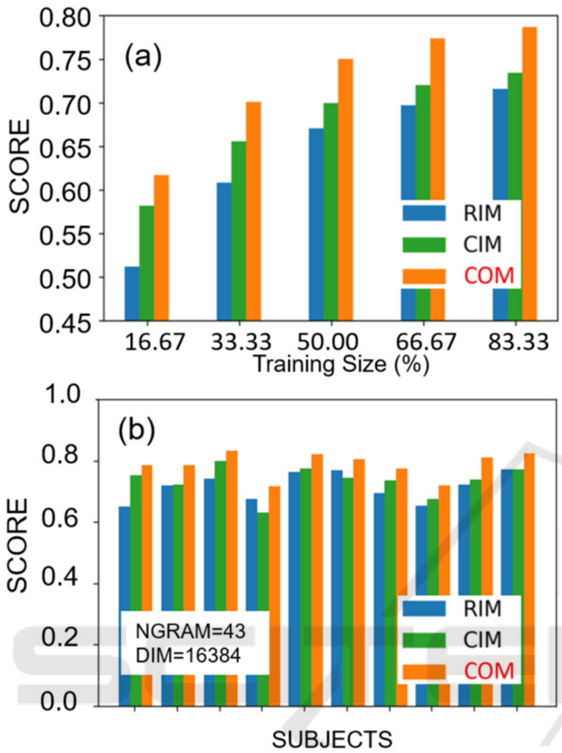


Figure 9: **Ninapro DB5**. (a) Classification accuracy was averaged across all subjects and gestures for different training sizes. (b) Averaged accuracy achieved by each subject across all 52 gestures using the largest training size.

Once again, the CompHD model demonstrated superior classification performance across all training sizes compared to both CIM and RIM. For some subjects, CompHD's classification accuracy reached nearly 80%, further highlighting its efficacy in delivering precise classifications across subjects. Notably, CompHD requires only 50% of the training data to outperform both CIM and RIM, even when larger training sizes (up to 83.33%) are used.

Moreover, CompHD not only achieves better accuracy than CIM and RIM, but the latter two models also exhibit significant inconsistencies across multiple databases. In contrast, CompHD has proven its robustness across various datasets and training sizes, consistently outperforming both CIM and RIM in every case.

Beyond its impressive accuracy and robustness, HDC (Hyperdimensional Computing) also demonstrates notable resilience to errors, making it a compelling choice for applications requiring fault tolerance. This error resilience enhances HDC's suitability for real-world conditions in in-memory computing, where environmental and electrical variability, along with data imperfections, are common challenges.

5 CONCLUSIONS

In this study, CompHD consistently outperformed both state-of-the-art CIM and RIM, demonstrating superior classification accuracy across all training sizes while requiring significantly fewer samples. This efficiency opens the door to embedded, on-device training, reducing reliance on large, centralized datasets and enabling more adaptive, resource-efficient deployments.

CompHD's robustness goes beyond performance; it also exhibits resilience to data variability and encoding errors, which is a critical advantage for real-world applications that deal with imperfect data or noisy environments. These qualities, combined with its compatibility with low-complexity and massively parallel operations, position CompHD as a highly effective choice for embedded systems. Furthermore, its high accuracy, efficiency, and fault tolerance make it a promising candidate for in-memory computing applications.

ACKNOWLEDGEMENTS

This project is supported by the "HDC" exploratory Carnot project.

REFERENCES

Hudgins, B., Parker, P., Scott, R. N. (1993). A new strategy for multifunction myoelectric control. *IEEE Transactions on Biomedical Engineering*, vol.40(1), 82-94.

Benatti, S., Farella, E., Gruppioni, E., Benini, L. (2014). Analysis of robust implementation of an EMG pattern recognition based control. *International Conference on Bio-inspired Systems and Signal Processing 2014* BIOSIGNAL.

Rahimi, A., Benatti, S., Kanerva, P., Benini, L., Rabaey, J. M. (2016). Hyperdimensional biosignal processing: A case study for EMG-based hand gesture recognition. *Proceedings of the 2016 IEEE International Conference on Rebooting Computing (ICRC),* San Diego, CA, USA, pp. 1-8.

Sgambato, B., Castellano, G. (2022). Performance comparison of different classifiers applied to gesture recognition from sEMG signals. In Bastos-Filho, T. F., de Oliveira Caldeira, E. M., Frizera-Neto, A. (Eds.), *XXVII Brazilian Congress on Biomedical Engineering. CBEB 2020.* IFMBE Proceedings, Vol. 83. Springer, Cham.

Salerno, A., Barraud, S. (2024). Evaluation and implementation of High-Dimensionnal Computing for gesture recognition using sEMG signals. *Proceedings of the 2024 International Conference on Control, Automation and Diagnosis (ICCAD)*

Lobov, S., Krilova, N., Kastalskiy, I., Kazantsev, V., Makarov, V. A. (2018). Latent factors limiting the performance of sEMG-interfaces. *Sensors*, 18(4), 1122.

Atzori, M., Gijsberts, A., Castellini, C., et al. (2014). Electromyography data for non-invasive naturally-controlled robotic hand prostheses. *Scientific Data*, 1, 140053

Atzori, M., et al. (2012). Building the Ninapro database: A resource for the biorobotics community. *Proceedings of the 2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, Rome, Italy, pp. 1258-1265.

Pizzolato, S., Tagliapietra, L., Cognolato, M., Reggiani, M., Müller, H., et al. (2017). Comparison of six electromyography acquisition setups on hand movement classification tasks. *PLOS ONE*, 12(10), e0186132.

Wan, Y., Han, Z., Zhong, J., Chen, G. (2018). Pattern recognition and bionic manipulator driving by surface electromyography signals using convolutional neural network. *International Journal of Advanced Robotic Systems*, 15(5).

Scheme, E., Englehart, K. (2014). On the robustness of EMG features for pattern recognition based myoelectric control: A multi-dataset comparison. *Proceedings of the 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Chicago, IL, USA, pp. 650-653.

Kanerva, P. (2009). Hyperdimensional Computing: An Introduction to Computing in Distributed Representation with High-Dimensional Random Vectors. *Cogn Comput* 1, 139–159.

Widdows, D., Cohen, T. (2015). Reasoning with vectors: A continuous model for fast robust inference. *Logic Journal of the IGPL*, 23(2), 141-173.

Rahimi, A., et al. (2017). High-dimensional computing as a nanoscalable paradigm. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 64(9), 2508-2521.

Chen H., Najafi, M. H., Sadredini, E., Imani, M. (2022). Full Stack Parallel Online Hyperdimensional Regression on FPGA. *Proceedings of the* 2022 IEEE 40th International Conference on Computer Design (ICCD), Olympic Valley, CA, USA, 2022, pp. 517-524

Benatti, S., et al. (2017). A sub-10mW real-time implementation for EMG hand gesture recognition based on a multi-core biomedical SoC. *Proceedings of the 2017 7th IEEE International Workshop on Advances in Sensors and Interfaces (IWASI)*, Vieste, Italy, pp. 139-144.

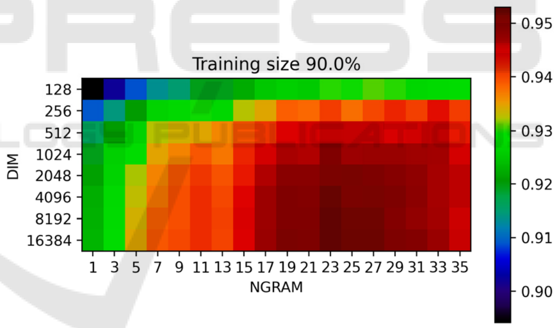
Karunaratne G., Rahimi, A., L. Gallo, M., Cherubini, G., Sebastian A. (2021). Real-time Language Recognition using Hyperdimensional Computing on Phase-change Memory Array," *IEEE 3rd International Conference on Artificial Intelligence Circuits and Systems (AICAS),* Washington DC, DC, USA, 2021, pp. 1-1

Li, H., et al. (2016). Hyperdimensional computing with 3D VRRAM in-memory kernels: Device-architecture co-design for energy-efficient, error-resilient language recognition. *Proceedings of the 2016 IEEE International Electron Devices Meeting (IEDM)*, San Francisco, CA, USA, pp. 16.1.1-16.1.4.

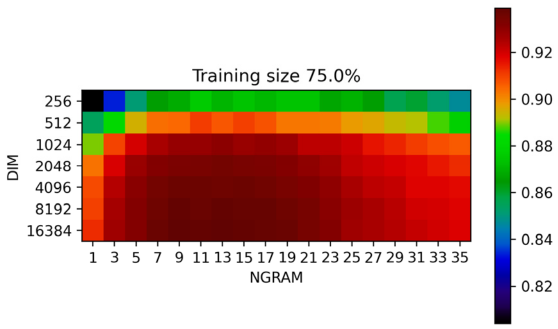
Karunaratne, G., Le Gallo, M., Cherubini, G., Benini, L., Rahimi, A., Sebastian, A. (2020). In-memory hyperdimensional computing. *arXiv preprint,* 1906.01548.

Abhijith, M., Shekhar, G. (2019). Language classification technique using in-memory high dimensional computing (HDC). *Proceedings of the 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT)*, Kannur, India, pp. 293-298.

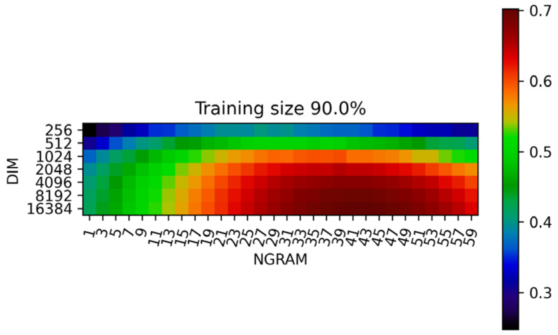
APPENDIX



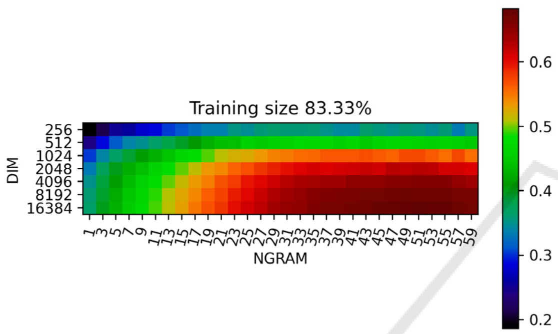
Master DB: Heatmap showing the accuracy of our HDC model for different HV dimensions (DIM) and NGRAM.



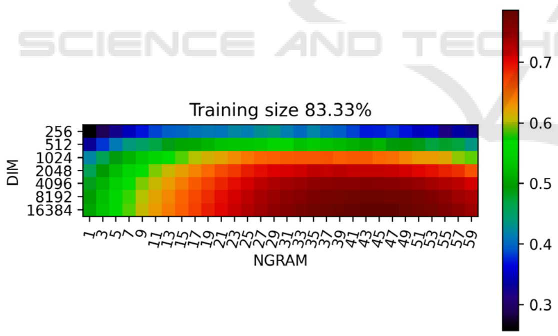
Pattern DB: Heatmap showing the accuracy of our HDC model for different HV dimensions (DIM) and NGRAM.



Ninapro DB1: Heatmap showing the accuracy of our HDC model for different HV dimensions (DIM) and NGRAM.



Ninapro DB4: Heatmap showing the accuracy of our HDC model for different HV dimensions (DIM) and NGRAM.



Ninapro DB5: Heatmap showing the accuracy of our HDC model for different HV dimensions (DIM) and NGRAM.