Dyadic Gender Composition and Task Classification Using fNIRS Hyperscanning Data

Liam Kruse
Department of Aeronautics and Astronautics
Stanford University
Stanford, CA 94305
lkruse@stanford.edu

***Abstract*—Functional near-infrared spectroscopy (fNIRS) is a revolutionary neuroimaging modality that provides unprecedented insights into brain activity. However, measuring the nuanced hemodynamic responses during dyadic inter-brain coherence—an fNIRS paradigm referred to as *hyperscanning*—remains largely unexplored. This work proposes a novel convolutional neural network-based approach to dyadic gender composition and task classification for an extensive hyperscanning dataset. The raw change in oxy-hemoglobin (HbO) is tested as input data against the inter-brain signal similarity computed via dynamic time warping. The proposed approach matches state-of-the-art classification accuracy, thereby providing a new avenue for exploring and understanding complex brain behavior.**

***Index Terms*—hyperscanning, convolutional neural network, dynamic time warping**

I. Introduction

Functional near-infrared spectroscopy (fNIRS) is a recently developed technology for studying brain activity and represents a low-cost, non-invasive alternative to traditional imaging modalities such as functional magnetic resonance imaging (fMRI). fNIRS uses near-infrared light to measure the hemodynamic response of the cerebral cortex to sensory, motor, and cognitive stimuli [1], [2]. This imaging modality typically consists of a lightweight sensor cap measuring hemodynamic activity across a series of channels, permitting fNIRs to be used in more ecologically valid situations and extending imaging accessibility to populations that cannot undergo fMRI scanning [3]. fNIRS has emerged as an indispensable method for researching brain computer interfaces (BCI) wherein brain signals are converted into commands for an external system [4]. BCI represent an avenue for persons with motor disabilities to express themselves and interact with the world [1].

Hyperscanning—a subdiscipline of fNIRS research wherein neural activity is simultaneously measured from two or more participants—is an area of particular interest in the neuroimaging community. Nuanced inter-brain coherence patterns emerge depending on the gender composition [5], [6] and task objectives [7] of the interacting dyad. Exploring and mapping these complex interactions more fully is essential to designing robust BCI systems, interpreting complex psychosocial interactions [8], and understanding the functionality of the human brain at a fundamental level.

In this work we propose a novel approach to dyadic gender composition and task classification using a modified form of the ubiquitous LeNet-5 convolutional neural network (CNN) architecture proposed by Lecun et al. [9]. We test two sets of input data—one set consisting of the difference in normalized oxy-hemoglobin (HbO) changes between participants, and one set consisting of inter-brain signal similarity as computed via dynamic time warping (DTW). To the best of the author's knowledge, this is the first time that CNNs have been leveraged to classify hyperscanning tasks and one of the first applications of CNNs within the broader fNIRS community. We validate our approach on an extensive hyperscanning datset of cooperative and competitive dyadic tasks obtained by the Center for Interdisciplinary Brain Sciences Research (CIBSR) at Stanford University, matching current state-of-the-art accuracy.

II. Related Work

BCI applications demand highly accurate classification of brain signals to distinguish between different user intentions, and thus fNIRS data classification remains an outstanding research question of great importance in the field of neuroimaging [4]. For example, the timely interpretation of fNIRS data could be leveraged to control prostheses [10], detect habituation in automated driving [11], or inform automated lane-change driving features [12].

Researchers have employed a multitude of machine learning techniques to classify fNIRS signals. Shamsi et al. implemented a support vector machine (SVM) with a quadratic polynomial kernel to classify movement execution tasks [4]. Peng et al. tested both a linear discriminant analysis algorithm and an SVM algorithm to classifiy motor imagery tasks wherein participants moved an on-screen object in their imagination [1]. Power et al. employed a hidden Markov model (HMM) to differentiate between mental arithmetic and music imagery, demonstrating the potential for a BCI device based on cognitive tasks rather than motor tasks [13]. HMMs have also successfully identified finger-tapping tasks [14] and detected music imagery [15] from fNIRS data. Still other researchers have conducted fNIRS classification tasks using K-nearest neighbors and Naïve Bayes [2], [10], [16]. Saadati et al. [17] and Asgher et al. [18] investigated the use of convolutional neural networks for mental workload classification; these remain the only instances of CNN-based classifiers for fNIRS data.

Accurately classifying motor tasks and cognitive imagery is empirically challenging even in the simplest instances of binary classification. Shamsi et al. achieved an average accuracy of 70.43 % [4] and Power et al. obtained an average accuracy of 77.20 % for binary classification problems, while Shamsi et al. and Peng et al. achieved accuracies of 78.55 % and 39.98 % for five and four-class problems, respectively. The two CNN-based approaches yield some of the highest accuracies for fNIRS classification tasks, with average four-class classification accuracies of 89.00% in [17] and 83.45% in [18]. The ability of CNNs to extract the most salient features with minimal levels of *a priori* feature extraction likely contributes to the state-of-the-art classification accuracies and inspires the approach proposed in this work.

Even less research has been conducted in the nascent hyperscanning domain. Baker et al. [5] and Cheng et al. [6] conducted wavelet coherence tests on fNIRS hyperscanning data for dyadic cooperation tasks and found that inter-brain coherence is highly dependent on the gender composition of the dyad. Cui et al. explored brain activity in the superior frontal cortex during both cooperation and competition tasks [7]. Nozawa et al. studied neural synchrony between individuals engaged in cooperative conversation [19]. To the best of the author's knowledge, no researchers have employed CNNs for hyperscanning classification tasks. In this work we propose a CNN architecture loosely based on the LeNet-5 architecture to predict dyadic gender compositions given knowledge of the task type and to predict the task type given knowledge of the dyadic gender composition.

III. Hyperscanning Dataset and Features

A. Data Acquisition

We utilized the extensive hyperscanning dataset collected by the Center for Interdisciplinary Brain Sciences Research at Stanford University. The dataset is not open source; access was granted from CIBSR after the completion of a human subjects research protections course. All sample data presented in this work is sanitized of personal identifiers.

A total of 222 participants (110 females, 112 males) were recruited for the hyperscanning study. Each participant was paired with a random individual to form a dyad. Participants did not interact prior to the study and were not matched based on age or ethnicity. Each dyad performed a series of consecutive *cooperation* and *competition* tasks on computer screens; the tasks were grouped into blocks of twenty tasks and the order was shuffled to reduce bias. For cooperation tasks, participants attempted to synchronize a button press event; for competition tasks, participants raced to press a button before their partner. A continuous wave fNIRS measured cortical hemodynamic activity in each participant's right prefrontal cortex and right temporal cortex; a total of 22 channels of optical density data were nominally recorded for each participant. A full summary of the data collection process and task sequencing is discussed in [5].

B. Preprocessing

The fNIRS imaging modality is sensitive to vascular dynamics and the raw data thus contained noise artifacts; most of these artifacts were removed via a bandpass filter shortly after the initial data collection sequence. The optical density data was then converted to oxy-hemoglobin data using the modified Beer–Lambert law [5].

In the original experimental design, the duration between tasks was varied to reduce habituation bias. Distributions of the number of data points recorded for cooperation tasks is shown in Fig. 1; the distribution for competition tasks is similar and is thus omitted for brevity. The fNIRS recorded optical density data at a frequency of 7.8125 Hz; thus, a task with 60 data points had a duration of 7.68 seconds. To convert the data into a consistent input size for the proposed CNN-based approach, we trimmed the time series to only consider the last 50 data points for each task; establishing consistent timing intervals is common practice in fNIRS classification problems [1], [4].

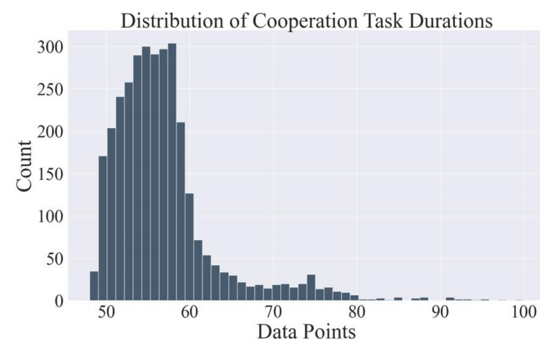


Fig. 1: Distribution of cooperation task durations.

Additionally, channels periodically stopped recording during the trials, resulting in intermittent columns of `NaN` values in the dataset. This could either be due to improper contact between the fNIRS cap optodes and the participant's scalp or a malfunctioning sensor. Fig. 2 shows the missing cooperation task counts for each channel. The distribution for competition tasks is virtually identical and is omitted for brevity. Channels 1, 10, and 19 did not record data during any tasks, likely indicating a malfunction; we thus removed these channels from the dataset. Furthermore, we removed channel 8 as it failed to record data on nearly ten percent of all tasks. Channels 2, 5, and 6 infrequently stopped recording during the experiments; since data is typically present for these channels, it is more likely that the occurrences of missing data represent isolated incidents of improper optode contact. Instead of removing these channels, we simply filled `NaN` values with the mean value of the adjacent channels at the given time step.

We removed all male-female dyads, keeping only male-male and female-female dyadic gender compositions to simplify the classification task; binary classification is a typical first step in fNIRS classification studies [1], [13]. Future work will extend classification to the multi-class case with all three dyadic gender compositions present in the dataset. We normalized all remaining data to complete the data cleaning process.

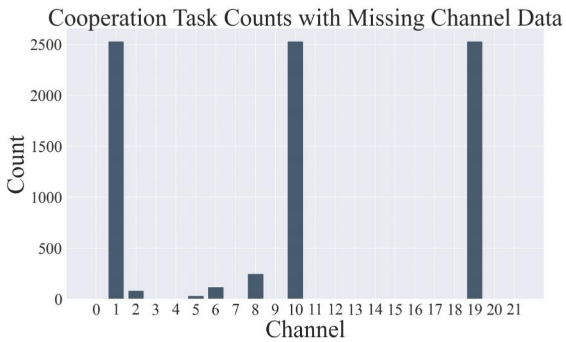


Fig. 2: Occurrences of missing cooperation task channel data.

At the conclusion of the data cleaning process, a total of 3,188 tasks remained in the dataset. Table I displays a summary of key dataset statistics.

TABLE I: Hyperscanning Dataset Statistics

Quantity	**Count**
Total Tasks	**$3,188$**
Cooperation Tasks	$1,569$
Competition Tasks	$1,619$
Total:	$3,188$
Male-Male Dyads	$1,675$
Female-Female Dyads	$1,513$
Total:	$3,188$

Fig. 3 displays post-processed sample data from the hyperscanning dataset. Normalized single channel data from one participant is plotted against the matching channel data from their dyadic partner. The vertical red lines denote the end of a given task and the start of a subsequent task.

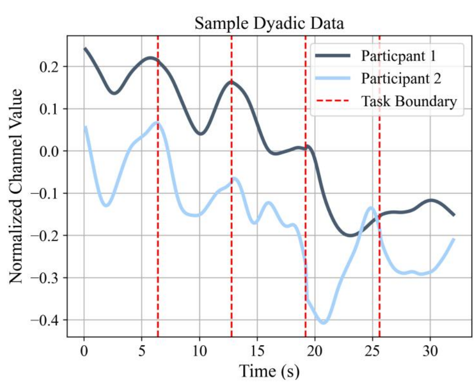


Fig. 3: Example dyadic data.

C. CNN Input Data

Convolutional neural networks are capable of extracting the salient features from input data with minimal *a priori* intervention and feature engineering—a phenomenon that drives the ubiquitous use of CNNs in image recognition and computer vision tasks. In this project we thus devote minimal attention to handcrafted feature engineering and instead test the proposed CNN-based classifier on two sets of minimally processed data:

1) Differences in normalized oxy-hemoglobin data
2) Channelwise similarity scores computed via dynamic time warping

1) Oxy-Hemoglobin Data: The raw optical density data recorded by the fNIRS cap was converted to oxy-hemoglobin data using the modified Beer–Lambert Law for scattering media; a discussion of the Beer–Lambert Law and its applications in fNIRS technology is presented in [20].

If we let H^{p1} denote the raw HbO data for the first participant in a dyad and H^{p2} denote the raw HbO data for the second participant, then we define the first set of input data as the difference in normalized HbO data between the two participants,

$$\left\|H_j^{p1}\right\|_2 - \left\|H_j^{p2}\right\|_2 \text{ for } j \text{ channels}$$

for each dyad in the dataset. Fig. 4 presents a sample plot of the difference in normalized HbO data for a sample dyad completing a competition task:

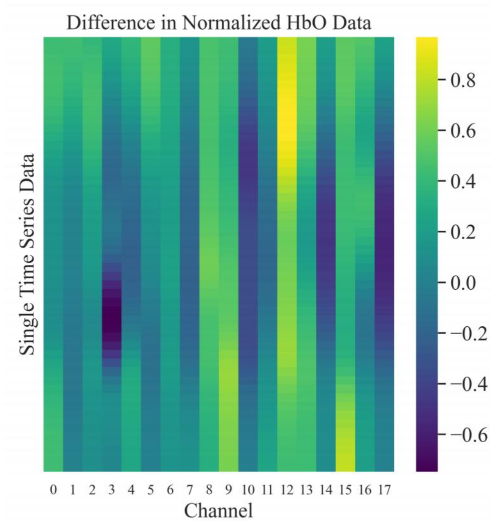


Fig. 4: Oxy-hemoglobin input data.

2) Channelwise Similarity Scores: We tested a second set of input data comprising the channelwise similarity scores computed via dynamic time warping. DTW is a technique to find an optimal alignment between two sequences of time-series data. As the name suggests, the sequences are *warped* nonlinearly to match each other [21]. Salient features in the data are then compared independent of nonlinearities in the time domain. The similarity score computed via DTW might thus more accurately describe how well two time series match each other when compared to a more conventional Euclidean distance measurement, as similar features will still be detected even if they do not line up exactly. Zhu and Najafizadeh successfully employed DTW to average fNIRS signals and localize active brain ragions in [22]. Fig. 5 shows how DTW can be used to align time series with a nonlinear time domain dependency.

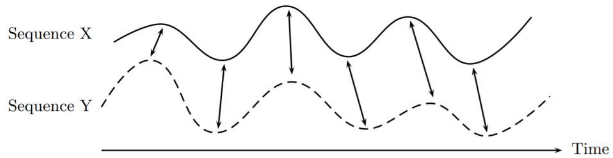


Fig. 5: Optimal alignment produced by DTW [21].

Fig. 6 presents the channelwise similarity scores for the same competition task shown in Fig. 4. Higher scores indicate a higher degree of similarity between the time series data. For example, channel 12 data for participant 1 is relatively similar to channel 15 data for participant 2 according to the DTW distance metric.

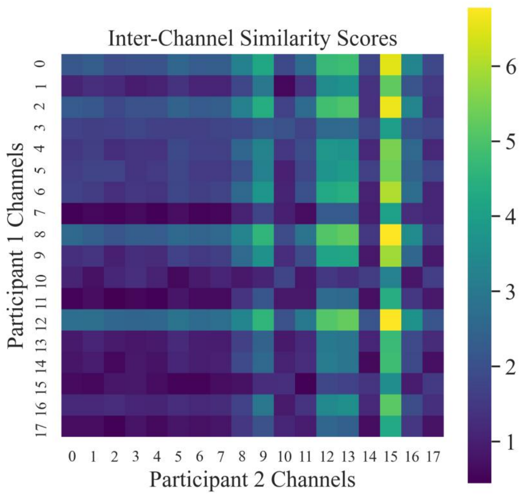


Fig. 6: Channelwise similarity score input data.

IV. Methods

A. LeNet-5 CNN

Convolutional neural networks are a class of deep neural network that has achieved tremendous success in the object detection and image classification domains. Three defining characteristics of CNNs are *local receptive fields*, *weight replication*, and *subsampling* [9].

Data is passed layer to layer in a CNN, and inputs at one layer are received from a small region of points—referred to as a *local receptive field*—in the previous layer. Local connections allow a CNN to extract the most basic visual features such as edges and corners and then assemble the features in subsequent layers to extract higher order features. Furthermore, weight vectors are held identical for particular units with receptive fields scattered across the input image—a paradigm known as *weight replication* ([23]–[25], as cited in [9]). Units with identical weights are combined into a single plane and output a *feature map*. This effectively shares feature detectors across the entire image. Finally, the concept of *subsampling* is used to reduce the resolution of the feature maps and reduce the output sensitivity to distortions and feature shifts [9]. Subsampling can be implemented using an average pooling layer in TensorFlow[1].

Lecun et al. developed the iconic CNN architecture referred to as LeNet-5, shown in Fig. 7, to incorporate the three previously discussed CNN design paradigms. Two two-dimensional convolutions are performed on the input data, with intermediate subsampling layers reducing the resolution of the feature maps. The architecture concludes with fully connected layers leading to an output layer; the original LeNet-5 architecture proposed a convolutional layer fully connected to a single dense layer [9]. We employ a CNN based on the original LeNet-5 design to provide baseline results.

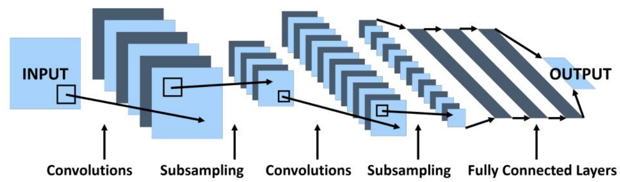


Fig. 7: LeNet-5 CNN architecture.

B. Proposed CNN Architecture

The original LeNet-5 CNN employs average pooling layers to subsample the feature maps. We propose an alternate CNN architecture that closely mirrors the LeNet-5 scheme but does not contain the average pooling layers; an architectural diagram is presented in Fig. 8. Removing the pooling layers theoretically heightens the output sensitivity to slight distortions in the feature map but maintains a high resolution throughout each layer. We suspect that the elevated resolution will enable the proposed CNN to extract nuanced features in the input fNIRS data and test this theory in the subsequent section.

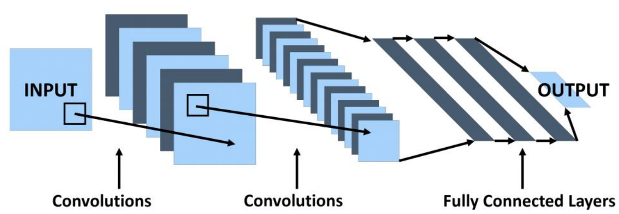


Fig. 8: Proposed CNN architecture.

V. Results and Discussion

We implemented a modified version of the LeNet-5 CNN in TensorFlow with two sets of alternating two-dimensional convolution and average pooling layers, followed by three dense layers with 128 units each. We also implemented our proposed CNN that does not include the average pooling layers. The architectures were tested on the series of four binary classification tasks described in Table II.

[1]https://www.tensorflow.org/

TABLE II: Classification Tasks

- **MM Task Prediction:** Given male-male dyadic data, predict if task is cooperation or competition
- **FF Task Prediction:** Given female-female dyadic data, predict if task is cooperation or competition
- **Coop Gender Prediction:** Given cooperation task data, predict if dyad is male-male or female-female
- **Comp Gender Prediction:** Given competition task data, predict if dyad is male-male or female-female

We conducted the four listed tests with both sets of input data—normalized differences in HbO and channelwise similarity scores—on both CNN architectures. We used three-fold cross-validation to estimate the accuracy of each method. The CNNs were trained for 20 epochs with a batch size of 32 to balance the bias-error tradeoff. Results for the HbO input data are presented in Table III and results for the DTW similarity score input data are presented in Table IV.

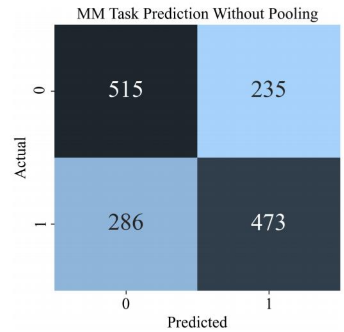
TABLE III: Results with Raw Input Data

Classification Task	Pooling	Accuracy (%)
MM Task Prediction	Yes	53.88
	No	55.53
FF Task Prediction	Yes	54.48
	No	57.20
Coop Gender Prediction	Yes	62.99
	No	66.95
Comp Gender Prediction	Yes	62.07
	No	65.29

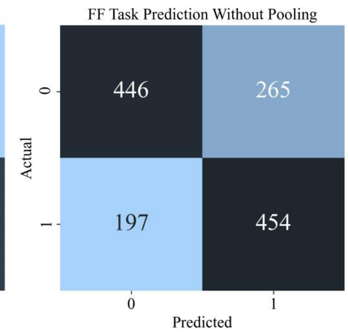
TABLE IV: Results with DTW Input Data

Classification Task	Pooling	Accuracy (%)
MM Task Prediction	Yes	55.86
	No	65.47
FF Task Prediction	Yes	60.28
	No	66.08
Coop Gender Prediction	Yes	75.80
	No	80.61
Comp Gender Prediction	Yes	72.84
	No	78.88

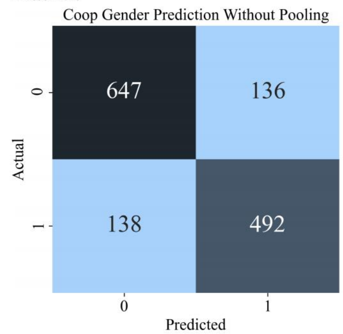
The tabulated results indicate it is easier to classify dyadic gender composition given the task type than it is to classify the task type given the gender composition. Furthermore, the results indicate that removing the pooling layers gives more predictive power to the CNN. The CNN achieves a higher classification accuracy when receiving channelwise similarity scores as the input data. The classification accuracies obtained with similarity score inputs and pooling layers removed matches the fNIRS binary classification state-of-the-art accuracies discussed in Section II. Confusion matrices displaying cumulative predictions across all cross-validation folds are shown in Figs. 9a-9d. For the sake of brevity, only results for the experiments with similarity score inputs and pooling layers removed are displayed.



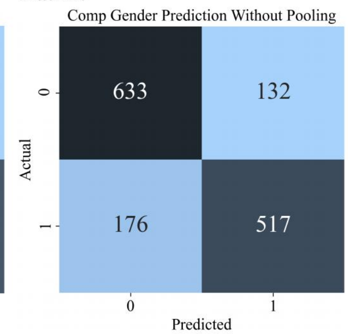
(a) MM task prediction confusion matrix.



(b) FF task prediction confusion matrix.



(c) Coop task prediction confusion matrix.



(d) Comp task prediction confusion matrix.

Fig. 9: Confusion matrices for experiments with similarity score inputs and pooling layers removed.

VI. Conclusion

Functional near-infrared spectroscopy is a nascent neuroimaging modality that is receiving increasing interest from the medical community. However, nuanced hemodynamic behaviors recorded via fNIRS hyperscanning procedures during dyadic inter-brain coherence is not yet fully understood. In this work we presented a CNN-based classifier for predicting the gender composition and task type of dyads performing cooperative and competitive tasks. We computed channelwise similiarity scores for each dyad using dynamic time warping and obtained state-of-the-art classification accuracy on binary fNIRS classification tasks. To the best of the author's knowledge, this is the first use of a convolutional neural network to classify fNIRS signals, thereby providing neuroimaging researchers with a novel strategy for analyzing dyadic inter-brain coherence. In future work we will explore alternate CNN architectures and extend the binary classification tasks to multi-class classification tasks.

Acknowledgment

The author thanks the Center for Interdisciplinary Brain Sciences Research for dataset access and Dr. Stephanie Balters for useful discussions.

References

[1] H. Peng, J. Chao, S. Wang, J. Dang, F. Jiang, B. Hu, and D. Majoe, "Single-trial classification of fnirs signals in four directions motor imagery tasks measured from prefrontal cortex," *IEEE transactions on nanobioscience*, vol. 17, no. 3, pp. 181–190, 2018.

[2] H. Ayaz, P. A. Shewokis, S. Bunce, M. Schultheis, and B. Onaral, "Assessment of cognitive neural correlates for a functional near infrared-based brain computer interface system," in *International Conference on Foundations of Augmented Cognition*. Springer, 2009, pp. 699–708.

[3] N. Liu, X. Cui, D. M. Bryant, G. H. Glover, and A. L. Reiss, "Inferring deep-brain activity from cortical activity using functional near-infrared spectroscopy," *Biomedical optics express*, vol. 6, no. 3, pp. 1074–1089, 2015.

[4] F. Shamsi and L. Najafizadeh, "Multi-class classification of motor execution tasks using fnirs," in *2019 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*. IEEE, 2019, pp. 1–5.

[5] J. M. Baker, N. Liu, X. Cui, P. Vrticka, M. Saggar, S. H. Hosseini, and A. L. Reiss, "Sex differences in neural and behavioral signatures of cooperation revealed by fnirs hyperscanning," *Scientific reports*, vol. 6, no. 1, pp. 1–11, 2016.

[6] X. Cheng, X. Li, and Y. Hu, "Synchronous brain activity during cooperative exchange depends on gender of partner: A fnirs-based hyperscanning study," *Human brain mapping*, vol. 36, no. 6, pp. 2039–2048, 2015.

[7] X. Cui, D. M. Bryant, and A. L. Reiss, "Nirs-based hyperscanning reveals increased interpersonal coherence in superior frontal cortex during cooperation," *Neuroimage*, vol. 59, no. 3, pp. 2430–2437, 2012.

[8] T. Koike, H. C. Tanabe, and N. Sadato, "Hyperscanning neuroimaging technique to reveal the "two-in-one" system in social interactions," *Neuroscience research*, vol. 90, pp. 25–32, 2015.

[9] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[10] J. Shin and J. Jeong, "Multiclass classification of hemodynamic responses for performance improvement of functional near-infrared spectroscopy-based brain–computer interface," *Journal of biomedical optics*, vol. 19, no. 6, p. 067009, 2014.

[11] S. Balters, S. Sibi, M. Johns, M. Steinert, and W. Ju, "Learning-by-doing: Using near infrared spectroscopy to detect habituation and adaptation in automated driving," in *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 2017, pp. 134–143.

[12] S. Sibi, S. Baiters, B. Mok, M. Steiner, and W. Ju, "Assessing driver cortical activity under varying levels of automation with functional near infrared spectroscopy," in *2017 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2017, pp. 1509–1516.

[13] S. D. Power, T. H. Falk, and T. Chau, "Classification of prefrontal activity due to mental arithmetic and music imagery using hidden markov models and frequency domain near-infrared spectroscopy," *Journal of neural engineering*, vol. 7, no. 2, p. 026002, 2010.

[14] R. Sitaram, H. Zhang, C. Guan, M. Thulasidas, Y. Hoshi, A. Ishikawa, K. Shimizu, and N. Birbaumer, "Temporal classification of multichannel near-infrared spectroscopy signals of motor imagery for developing a brain–computer interface," *NeuroImage*, vol. 34, no. 4, pp. 1416–1427, 2007.

[15] T. H. Falk, K. Paton, S. Power, and T. Chau, "Improving the performance of nirs-based brain-computer interfaces in the presence of background auditory distractions," in *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2010, pp. 517–520.

[16] V. Gottemukkula and R. Derakhshani, "Classification-guided feature selection for nirs-based bci," in *2011 5th International IEEE/EMBS Conference on Neural Engineering*. IEEE, 2011, pp. 72–75.

[17] M. Saadati, J. Nelson, and H. Ayaz, "Convolutional neural network for hybrid fnirs-eeg mental workload classification," in *International Conference on Applied Human Factors and Ergonomics*. Springer, 2019, pp. 221–232.

[18] U. Asgher, K. Khalil, Y. Ayaz, R. Ahmad, and M. J. Khan, "Classification of mental workload (mwl) using support vector machines (svm) and convolutional neural networks (cnn)," in *2020 3rd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*. IEEE, 2020, pp. 1–6.

[19] T. Nozawa, Y. Sasaki, K. Sakaki, R. Yokoyama, and R. Kawashima, "Interpersonal frontopolar neural synchronization in group communication: an exploration toward fnirs hyperscanning of natural interactions," *Neuroimage*, vol. 133, pp. 484–497, 2016.

[20] L. Kocsis, P. Herman, and A. Eke, "The modified beer–lambert law revisited," *Physics in Medicine & Biology*, vol. 51, no. 5, p. N91, 2006.

[21] M. Müller, "Dynamic time warping," *Information retrieval for music and motion*, pp. 69–84, 2007.

[22] L. Zhu and L. Najafizadeh, "Dynamic time warping-based averaging framework for functional near-infrared spectroscopy brain imaging studies," *Journal of biomedical optics*, vol. 22, no. 6, p. 066011, 2017.

[23] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propagation," California Univ San Diego La Jolla Inst for Cognitive Science, Tech. Rep., 1985.

[24] K. Fukushima and S. Miyake, "Neocognitron: A new algorithm for pattern recognition tolerant of deformations and shifts in position," *Pattern recognition*, vol. 15, no. 6, pp. 455–469, 1982.

[25] Y. LeCun *et al.*, "Generalization and network design strategies," *Connectionism in perspective*, vol. 19, pp. 143–155, 1989.