Project Proposal

Extended reality (XR) pushes the frontier of intuitive and high-bandwidth human-machine interaction. However, existing XR interfaces often use standardized gesture sets which can be both less effective than a user defining custom gestures (Nacenta et al., 2013) and inaccessible to those with motor impairments (Yamagami et al., 2023). Training personal gesture classifiers is difficult because they require either training subject-specific models which can be burdensome to calibrate for every novel user, or collecting data from thousands of participants to sufficiently generalize across novel users (Kaifosh & Reardon, 2025). We aim to train a personalized EMG gesture classification model that can generalize well with little calibration data. Meta-learning algorithms enable models that can generalize across users via user-specific adaptation with only a small set of training samples. The performance of these models can often be improved through selective meta-augmentation of training data, where data with new classes are created from preexisting data (Rajendran et al., 2023). However, which augmentation methods are optimal for gesture recognition with electromyogram (EMG) and inertial measurement unit (IMU) have not been well investigated. Additionally, data augmentation methods are normally performed with image data, not time series data. In this project, we aim to address this by investigating different augmentations to create new gesture classes. We will evaluate the effects on performance in a meta-learning gesture classification model and analyze which augmentations offer the greatest benefits.

We will use a small gesture dataset (Yamagami et al., 2024) with 32 users and 10 gesture classes per user, recorded with 16 channels of EMG and 72 channels of IMU. We will train and evaluate an existing meta-learning model architecture on our augmented datasets. We will first meta-augment our dataset by creating symmetric (i.e., left-right) mirroring and combining one-handed gestures to create new two-handed gestures. Our project will explore different ways of mirroring and combining gestures, similar to Shahi et al. (2024), while applying additional transformations like rotating IMU data and activity-aware masking. We will also determine how noise-based perturbation, similar to Um et al. (2017), affects newly created and existing gestures. Multiple methods of noise perturbation, such as additive noise, temporal warping, and sensor wobble will be tested. The specific augmentation method used, which could be either a single method or combination of methods, will create a new augmented dataset to train our meta-learning model, then evaluate on testing data to give a performance metric. Our project will additionally focus on creating infrastructure to enable arbitrary augmentations of the base training data. As the meta-learning gesture classification model takes extensive resources to train, we will work with Rice's high-performance computing (HPC) cluster NOTS to benchmark our augmented datasets. We will adapt the pre-existing classification model, which includes a method of training and testing, to launch and aggregate results with our augmented data using SLURM, an HPC software.

References

Kaifosh, P., & Reardon, T. R. (2025). A generic non-invasive neuromotor interface for human-computer interaction. *Nature*. https://doi.org/10.1038/s41586-025-09255-w

Nacenta, M. A., Kamber, Y., Qiang, Y., & Kristensson, P. O. (2013). Memorability of pre-designed and user-defined gesture sets. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. https://doi.org/10.1145/2470654.2466142

Rajendran, J., Irpan, A., & Jang, E. (2020). Meta-Learning Requires Meta-Augmentation. *ArXiv (Cornell University)*, *33*, 5705–5715.

Shahi, S., Mollyn, V., Tymoszek Park, C., Kang, R., Liberman, A., Levy, O., Gong, J., Bedri, A., & Laput, G. (2024). Vision-Based Hand Gesture Customization from a Single Demonstration. *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*, 1–14. https://doi.org/10.1145/3654777.3676378

Um, T. T., Pfister, F. M. J., Pichler, D., Endo, S., Lang, M., Hirche, S., Fietzek, U., & Kulić, D. (2017). Data augmentation of wearable sensor data for parkinson's disease monitoring using convolutional neural networks. Proceedings of the 19th ACM International Conference on Multimodal Interaction. https://doi.org/10.1145/3136755.3136817

Yamagami, M., Mitchell, C. L., Portnova-Fahreeva, A. A., Kong, J., Mankoff, J., & Wobbrock, J. O. (2024). Customized Mid-Air Gestures for Accessibility: A $B Recognizer for Multi-Dimensional Biosignal Gestures. ArXiv.org. https://arxiv.org/abs/2409.08402

Yamagami, M., Portnova-Fahreeva, A. A., Kong, J., Wobbrock, J. O., & Mankoff, J. (2023). How Do People with Limited Movement Personalize Upper-Body Gestures? Considerations for the Design of Personalized and Accessible Gesture Interfaces. PubMed. https://doi.org/10.1145/3597638.3608430