Defining Out-of-Distribution Detection for EEG-BCIs

Ivo Pascal de Jong   Merlijn Mulder   Prajwal Prakash   Matias Valdenegro-Toro
Andreea Ioana Sburlea

Department of Artificial Intelligence, Bernoulli Institute – University of Groningen

university of groningen

What is Out-of-Distribution detection?

Machine Learning assumes that we do predictions on data similar to what the model is trained on, but this is often not the case. We call dissimilar samples Out-of-Distribution (OoD).

They might come from mind-wandering, off-task activities or a user being asleep. They could also come from differences in preprocessing, disconnected electrodes, or changes in environmental noise.

It is impossible to make correct classifications in these cases. Instead, we try to detect them based on the uncertainty of the model.

This has the potential to be better than standard artifact rejection, as it is specific to the model, but general for any artifact.

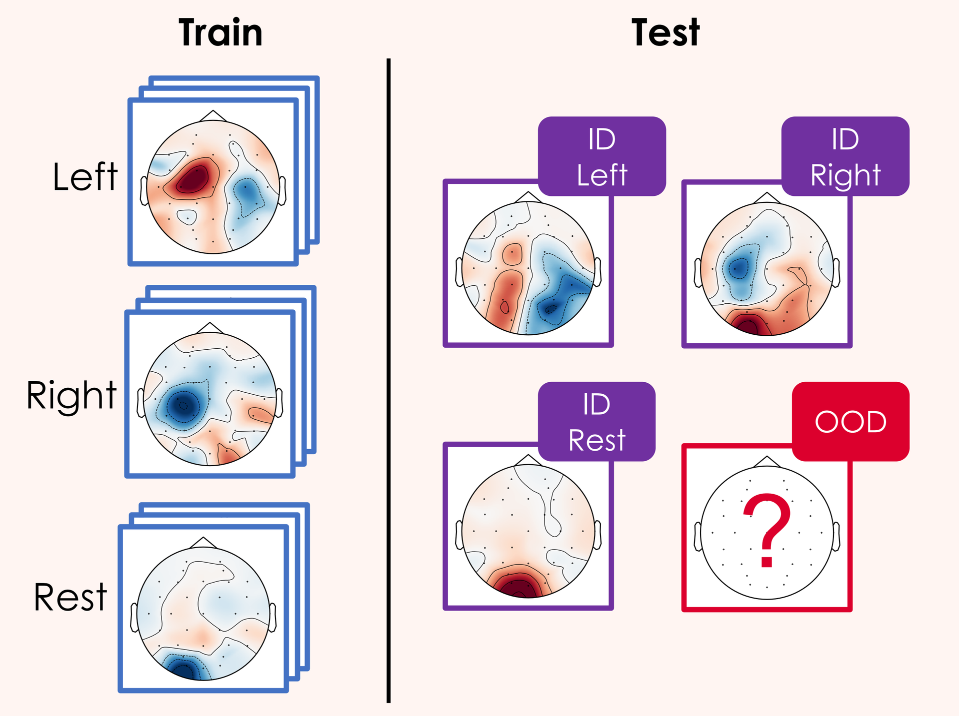


Figure 1. We aim to detect the known classes, while rejecting unfamiliar OoD samples.

Simulating OoD Data

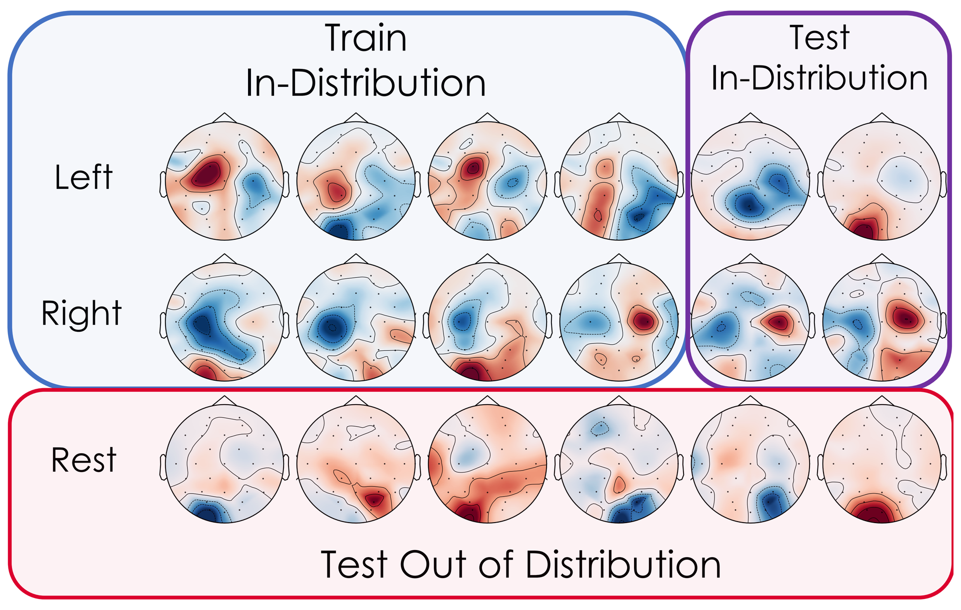
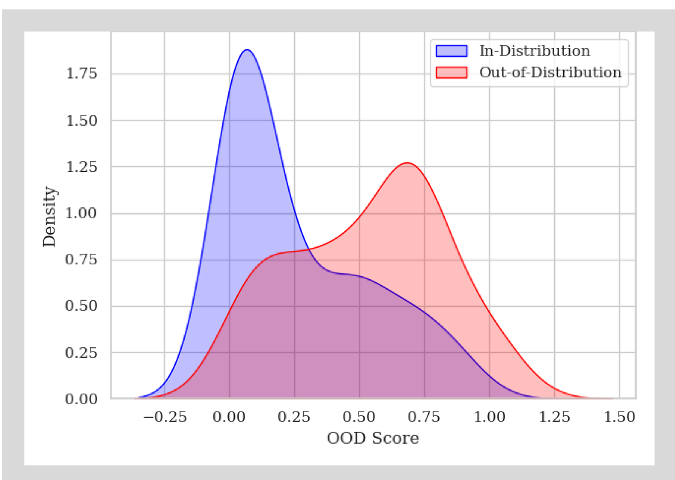
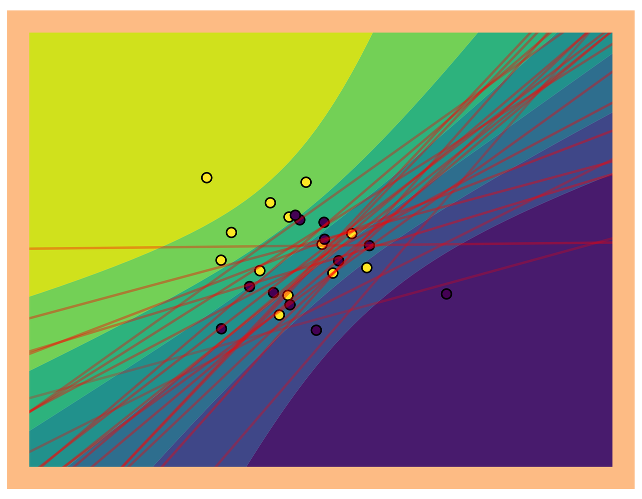


Figure 2. Leave-one-class-out OoD detection. OoD samples are created by removing one class from the training data.

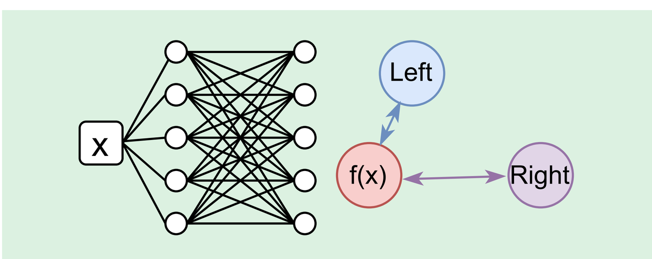
Methods for detecting OoD Data



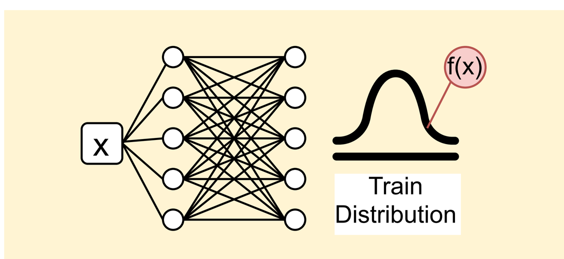
Class confidence (entropy) – Softmax



Model uncertainty – Deep Ensembles & MC Dropout



Distance in latent space – DUQ & KNN



Density in latent space – DDU & Energy score

Other sources of OoD

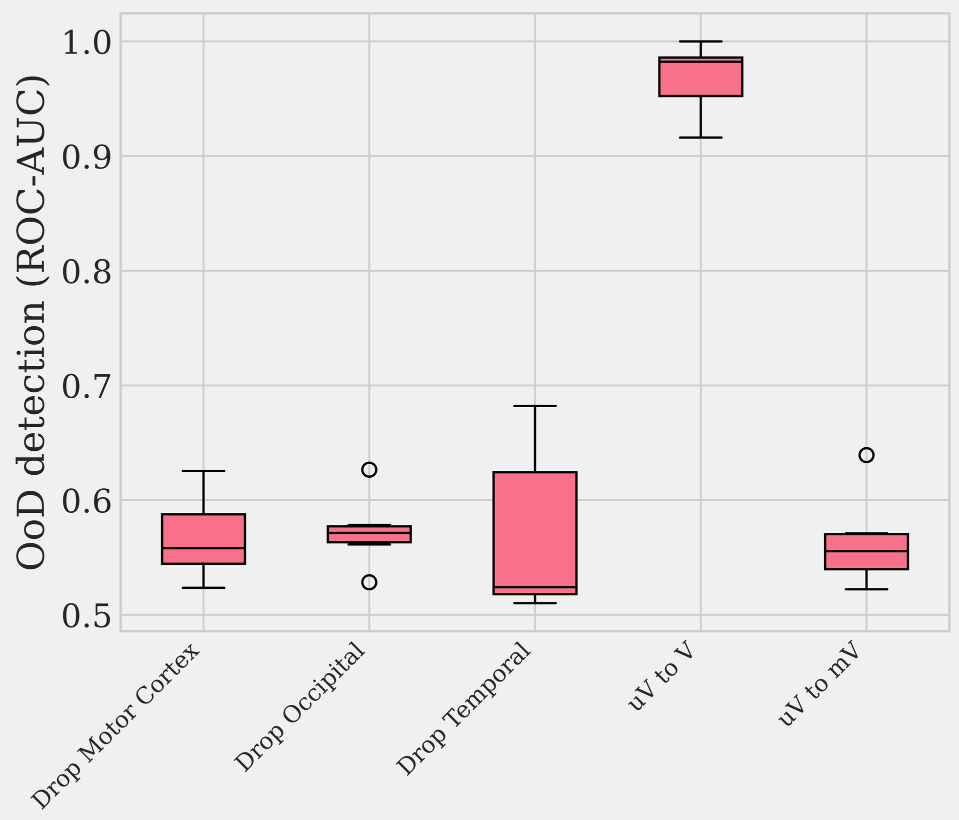
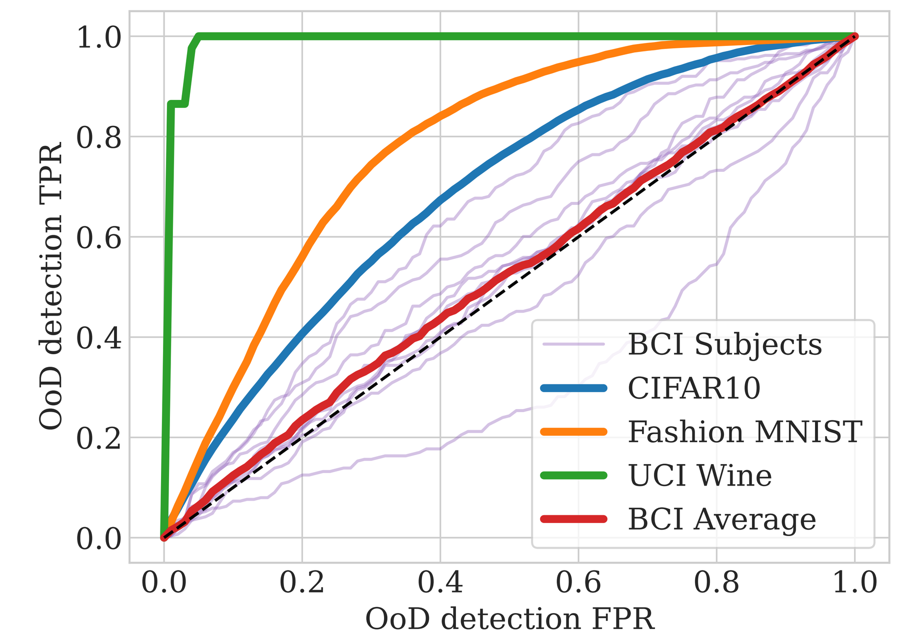
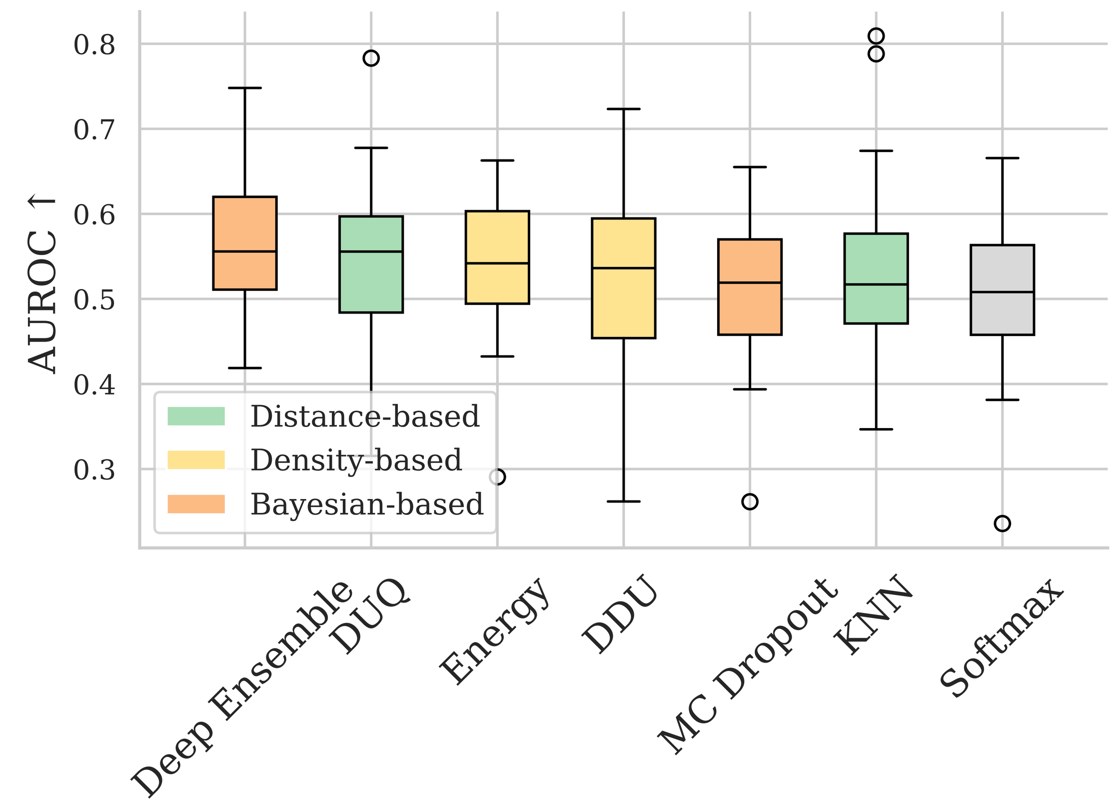


Figure 4. We can simulate different kinds of OoD. Some easier to detect than others. What do we need for robust BCIs?

Results

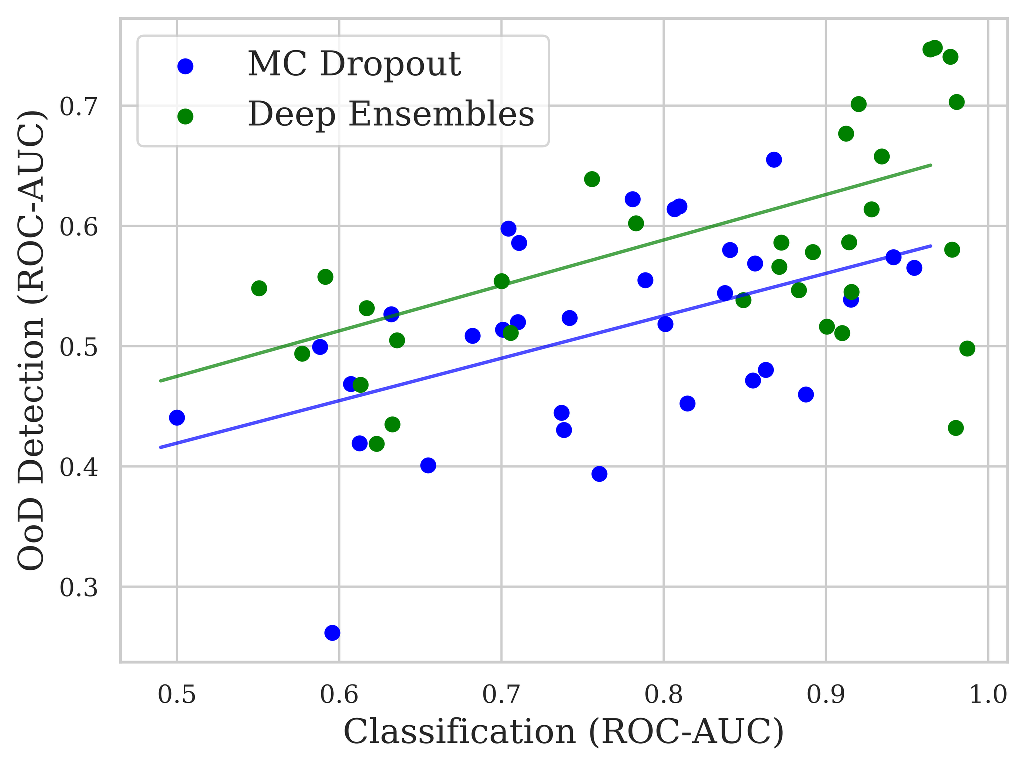


(a) ROC-curves for leave-one-class-out OoD detection. On all other datasets this works well, but for Motor-Imagery BCI this is nearly random.

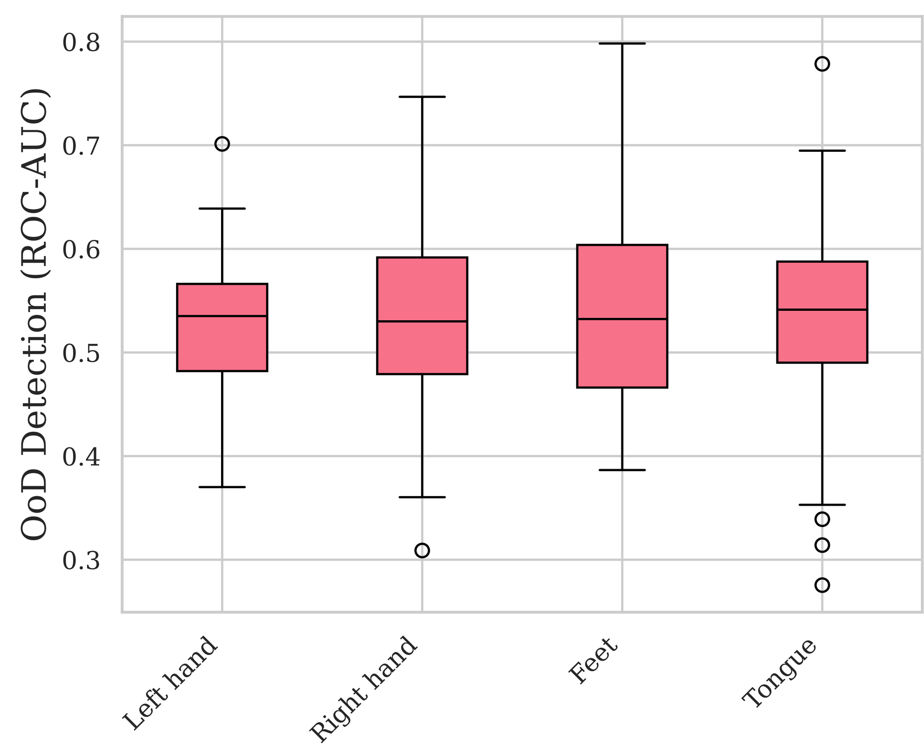


(b) Results of different OoD detection methods. Leave-one-class-out OoD detection is difficult for BCIs!

Figure 5. Primary results. Leave-one-class out OoD detection works in other fields, but is difficult for BCIs!



(a) Participants with good classifications also get better OoD detection. $\rho = 0.5, p < 0.01$



(b) Is one class more difficult to OoD detect than the other? No!

Figure 6. Further analysis. Bad OoD detection cannot be attributed to a specific class, but does correlate with classification performance.

Conclusion

- Out-of-Distribution detection is critical for real-life BCIs
- Currently available OoD detection methods are not sufficient for detecting off-task thoughts
- What kind of Out-of-Distribution you use matters a lot
- We need to define a set of Out-of-Distribution detection tasks that appropriately represent real-life scenarios.

What do you think should be detected with OoD detection?