Rank-1 CNN for mental workload classification from EEG

Sara Sedlar [1]*, Johann Benerradi [2]*, Côme Le Breton [1]*, Rachid Deriche [1], Théodore Papadopoulo [1], Max L. Wilson[2]

[1]Inria, Université Côte d'Azur , France

[2]School of Computer Science , University o f Nottingham , United Kingdom

***authors contributed equally**

PRESENTER

Johann Benerradi

BACKGROUND

Brain signals spontaneously generated by subjects while performing various tasks can be analysed by a passive BCI system. More specifically, this type of BCI is often used with the aim to assess the mental workload of users performing tasks with different levels of mental demand, especially with electroencephalography (EEG) (Wang et al. 2015, Appriou et al. 2018, Shalchy et al. 2020). In this work we propose a convolutional neural network (CNN) model for classifying mental workload levels from the EEG signals.

METHODS

Data

Three sessions of EEG recordings of 15 participants performing the Multi-Attribute Task Battery-II (MATB-II). Each session is decomposed in blocks of different difficulties: easy, medium and difficult. The blocks are split into epochs of 2 seconds (250 Hz) giving 447 epochs for each session and participant. Difficulty labels were provided only for the 2 first sessions.

Model

The model is based on a CNN with rank-1 constraint (Dupré La Tour et al. 2018, Kim et al. 2018) and spherical spatial patterns. We denote input EEG signals X^{NxT}, where N and T are the numbers of spatial and temporal sampling points. Signals are expressed in terms of spherical harmonics as $\hat{S}=Y^{inv}X$ over spatial axis, where Y^{inv} contains inverted SH basis. Signals $\hat{S}$ are downsampled to $\hat{S}_{ds}$ by a factor of 6 over time axis. The detailed architecture of the proposed model is given in Figure 1. The model is composed of 3 convolutional layers including one with rank-1 constraint and 3 fully connected layers, resulting in a total number of trainable parameters of 342.

Procedure

For validation, the 2 labelled sessions for each of the 15 participants were assigned to the train or validation set randomly. The model was then trained and validated based on this split for all the participants at once, in a generalised manner. This approach was chosen because the dataset contains only 2 different labelled sessions per participant, making it hard for the model to generalise to an unseen session with a personalised approach.

Finally, the model was re-trained on the 2 labelled sessions in order to produce the final results consisting of a prediction for each epoch of the unlabelled test set of each participant.

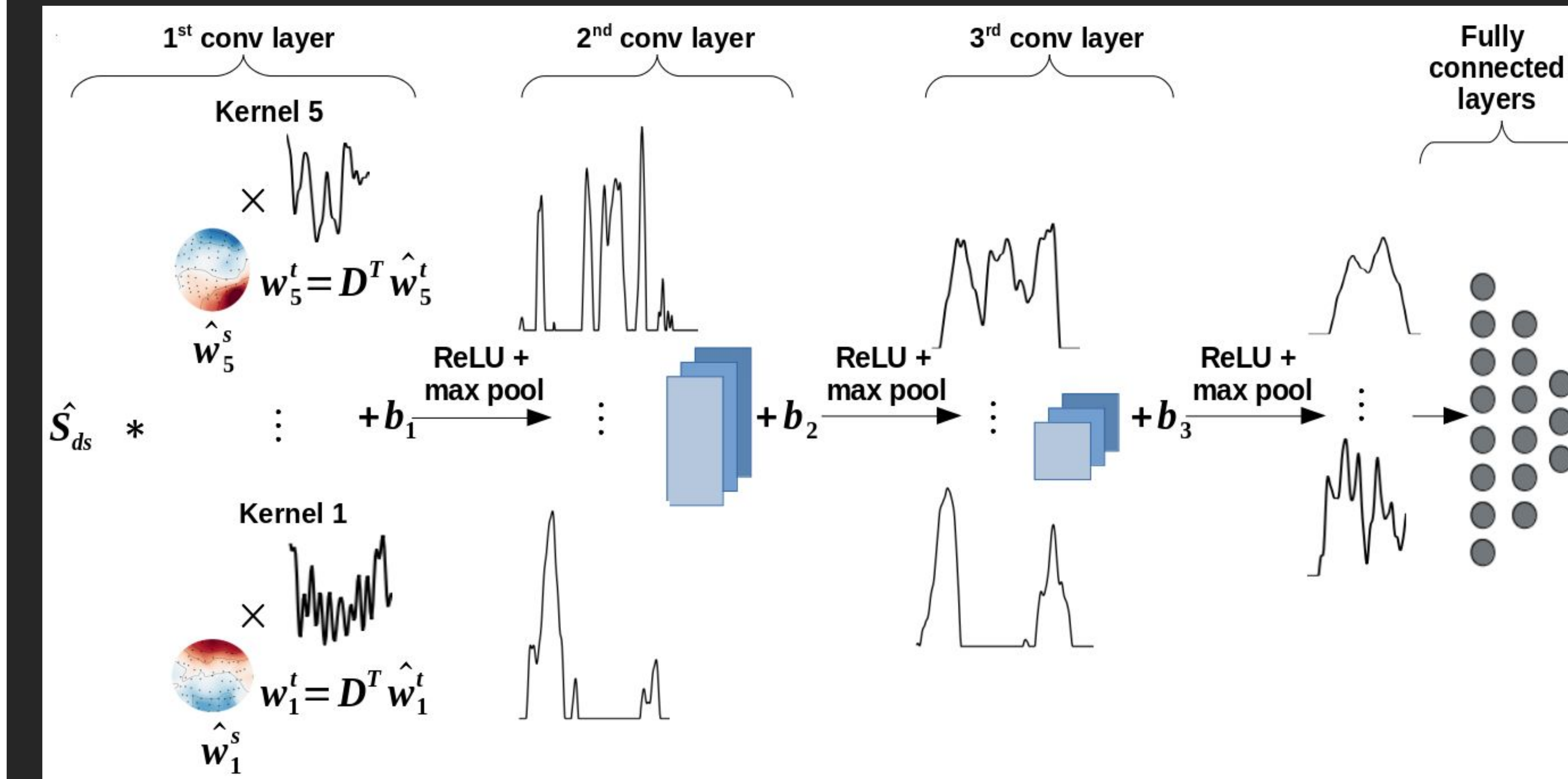


Figure 1. Rank-1 CNN for passive BCI EEG signal classification.

Model architecture and hyperparameters

- convolutional layer: 5 rank-1 kernels, 16 spherical harmonics spatial weights, 16 discrete cosine temporal weights ($\hat{w}_i^t$) transformed to signal domain as $w_i^t=D^T\hat{w}_i^t$ (D containing the discrete cosine basis), bias of 5, ReLu activation, max-pooling
- convolutional layer: 3 standard kernels of size 5x3, bias of 3, ReLu activation
- convolutional layer: 3 standard kernels of size 3x3, bias of 3, ReLu activation
- 2 fully connected layer of sizes 15x4 (bias 4), 4x4 (bias 4), ReLu activation
- last fully connected layer of size 4x3 (bias 3), softmax activation

Cross-entropy loss and Adam optimiser (learning rate 0.001).

Training performed over 25 epochs.

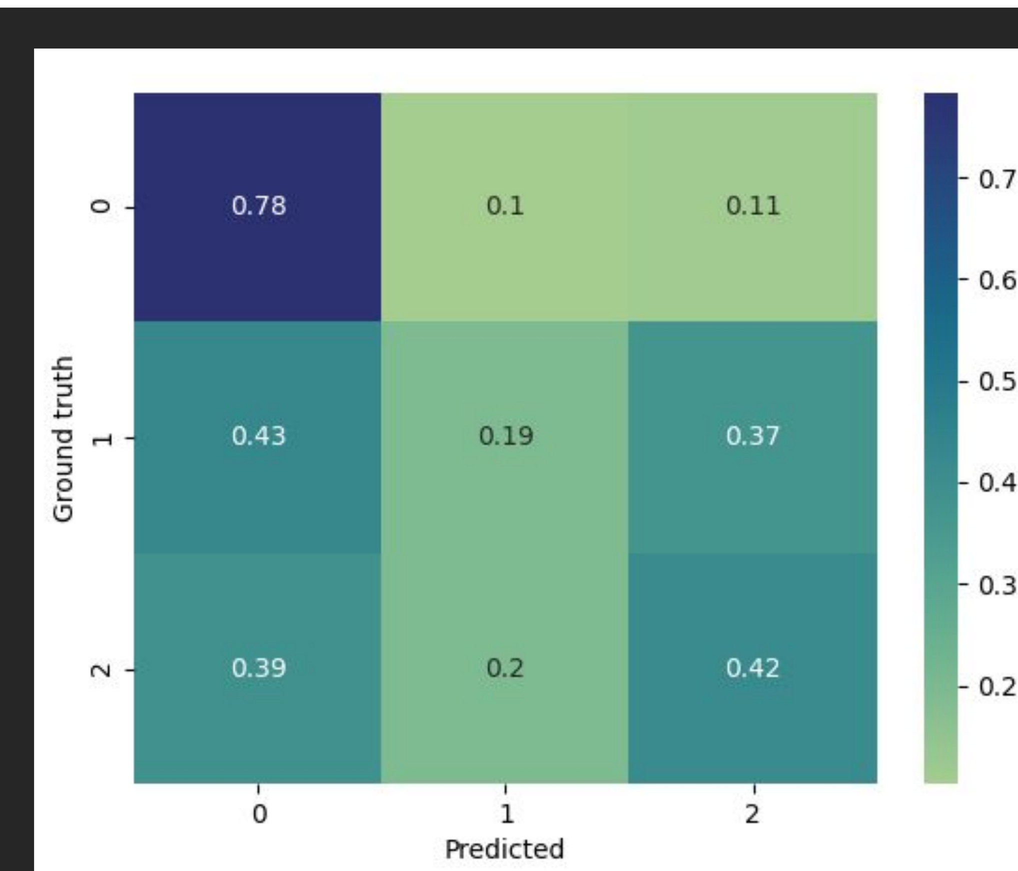


Figure 2. Overall confusion matrix of the CNN rank-1 model on 3 repetitions of the procedure with random train/validation splits (0: easy, 1: medium, 2: difficult)

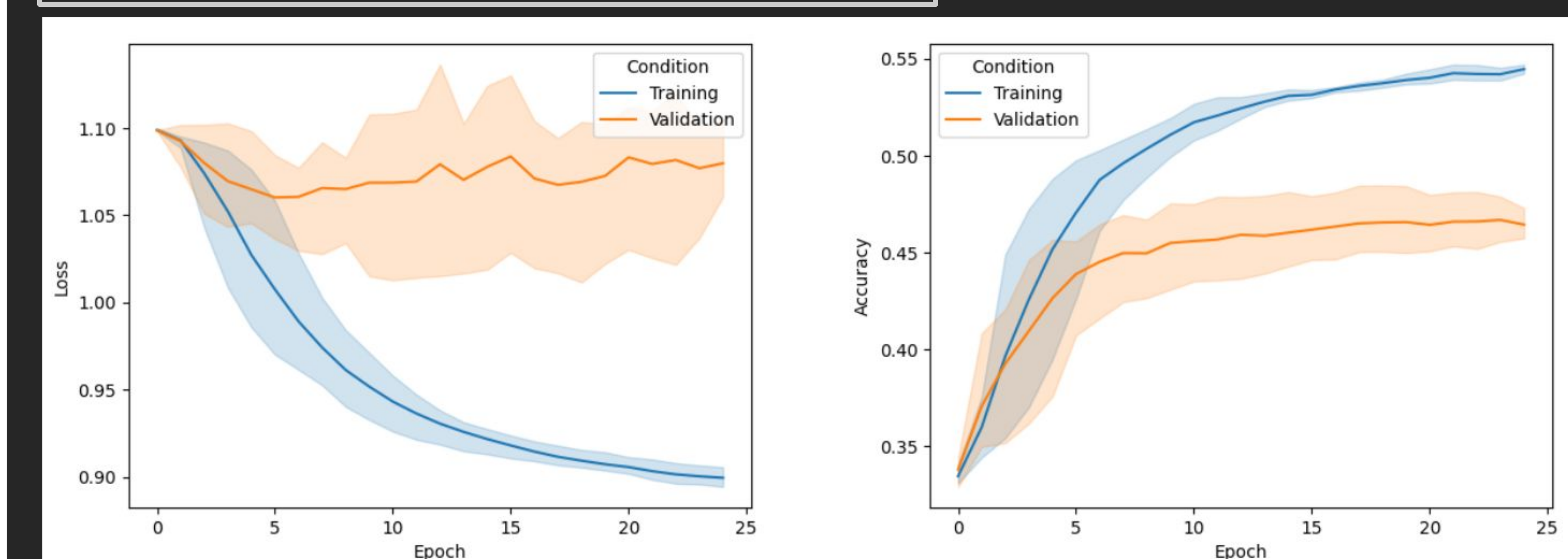


Figure 3. Training graphs of the CNN rank-1 model (loss on the left, accuracy on the right), using one labelled session for the train set and the other labelled session for the validation set (assigned randomly for each participant). The bands represent the 95 % confidence intervals on 3 repetitions of the procedure with random train/validation splits.

RESULTS

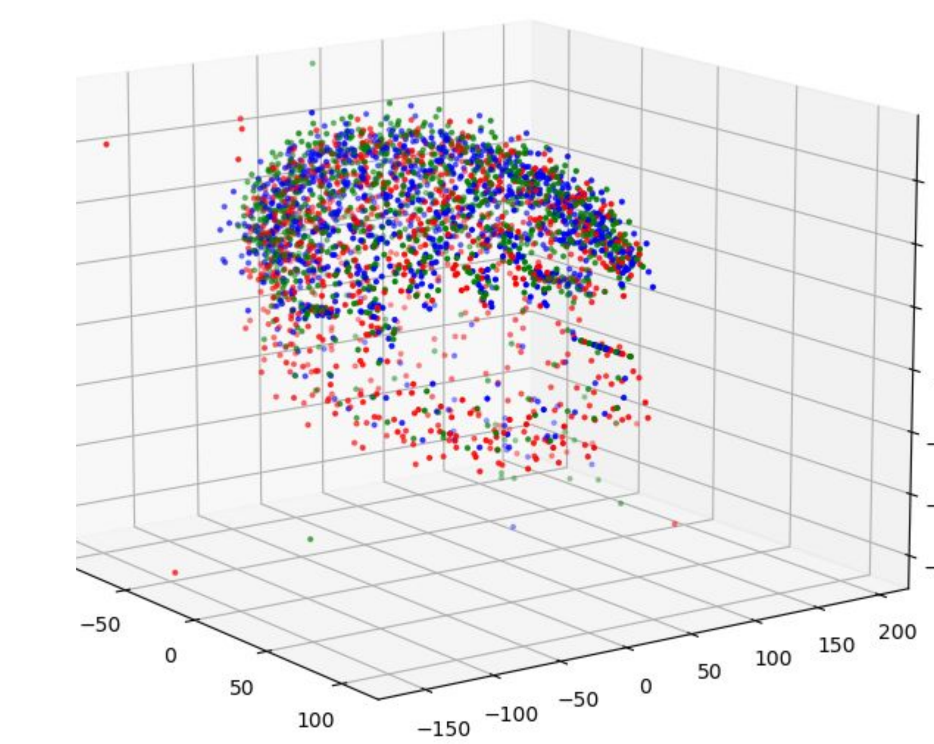
The accuracy on the validation set averaged over 3 experiments with random train/validation split was 46.46%.

The performance is consistently higher than the chance level of 33.33% as seen in Figure 3, however this remains far from being robust, highlighting the difficulty of classification of unseen sessions. The confusion matrix presented Figure 2 also highlight the fact that the accuracy is very different from class to class, with a tendency towards classifying signals into the easy task difficulty category.

The accuracy on the test set with hidden labels was 48.2%.

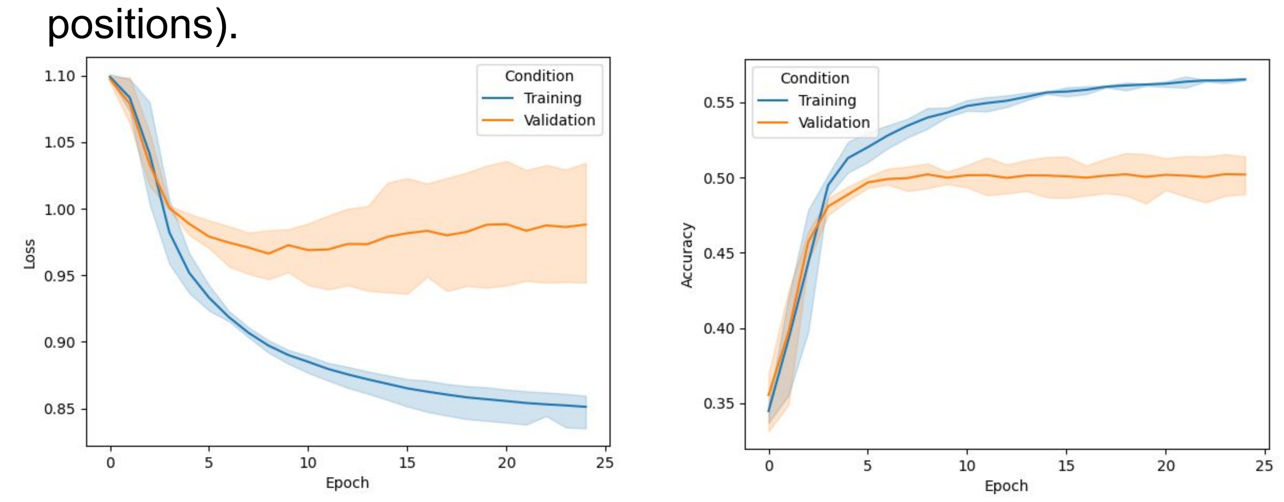
DISCUSSION

It appears that electrode orientations are not consistent, and some are either misplaced on the scalp or mislocated.



We fixed the axis inversions, however some inconsistencies remain, and further investigation regarding the registration of electrode position needs to be performed.

Regardless, we noticed that training the same model again using standard electrode positions makes it more stable when reproducing the experiment 3 times with random train/validation splits, as can be seen on the loss and accuracy graphs below. With this electrode positioning, the model averages a validation accuracy of 50.6% (as opposed to 46.46 % with the provided positions).



FUTURE WORK

Future work could focus on hyperparameter tuning with more exhaustive techniques such as grid search and per-subject tuning after generalised training.

ACKNOWLEDGEMENTS

Work supported by the ERC under the European Union's Horizon 2020 research and innovation program (ERC Advanced Grant agreement No 694665:CoBCoM).

Inria · erc · University of Nottingham UK | CHINA | MALAYSIA · 3IA Côte d'Azur Interdisciplinary Institute for Artificial Intelligence · UNIVERSITÉ CÔTE D'AZUR