# Suggestion for NLDL Winter School 2023

Anders Sildnes
University of Tromsø
Postboks 6050 Langnes
anders.sildnes@uit.no

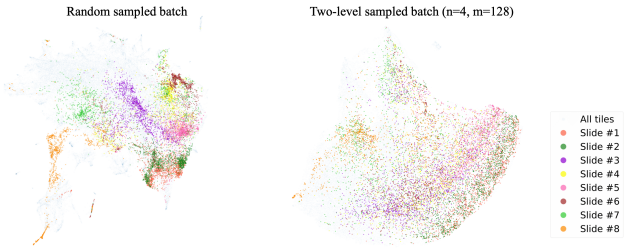


Figure 1. Figure taken entirely from [2]. Both pictures show a 2D UMAP projection of tile representations. On the left a training algorithm with randomly sampled tiles from slides, showing that features learned in the same slide cluster together. Using a different sampling algorithm, the picture on the right shows a more even distribution w.r.t slides

## 1. Suggestion

[2] uses self-supervised learning to learn morphological features of Lung Squamous Cell Carcinoma (LSCC). In training they use public TCGA and CPTAC datasets of H&E-stained biopsy slides. They observe from previous work that other training algorithms cluster together tiles from the same slides, creating a batch effect where the algorithm learns features that are more slide specific than feature specific. This indicates that the network might do poorly on generalized input. They visualize this using 2D UMAP, shown in Fig. 1.

I wish to evaluate [2] and quantify the quality of *the explanation* using metrics inspired by Quantus [1]. This means 1) finding relevant qualitative metrics for the given batch effect plots and 2) testing how the metrics change with model, batch and image adjustments. In particular, I am curious how well the metrics show batch/clustering effects for few (1-3) slides. This can be a useful indication early during a training process.

## References

[1] Anna Hedström, Leander Weber, Dilyara Bareeva, Franz Motzkus, Wojciech Samek, Sebastian Lapuschkin, and Marina M.-C. Höhne. Quantus: An explainable ai toolkit for responsible evaluation of neural network explanations. 2022. 1

[2] Weicheng Zhu, Carlos Fernandez-Granda, and Narges Razavian. Interpretable prediction of lung squamous cell carcinoma recurrence with self-supervised learning, 2022. 1