Supplementary Materials: BUNDL

Algorithm 1 BUNDL Uncertainty Computation from MCMC Samples

Require: $f_\theta(.)$, x from current epoch
Ensure: no gradient computation
  $N \leftarrow 20$ ▷ Number of MCMC samples
  $n \leftarrow 1$
  $\bar{f} \leftarrow 0 \in \mathbb{R}^N$
  while $n \leq N$ **do**
    $\bar{f}^n \leftarrow f_\theta(x)$ ▷ Network output with dropout
  end while
  $\overline{f_\theta(x)} = \frac{1}{N}\sum_{n=1}^{N} \bar{f}^n$ ▷ Average
  $z_x^{\hat{y}} = \frac{-1}{0.6932N}\sum_{n=1}^{N} \bar{f}^n \log \bar{f}^n + (1-\bar{f}^n)\log(1-\bar{f}^n)$ ▷ Expected entropy
  if truncate **then**
  $z_x^0 = 0.001$ ▷ Oversegmented seizure
  end if

Algorithm 2 BUNDL training framework

Require: θ, learning rate α, number of epochs M, $\{x, p_g\} \in$ batched training data
  $\theta \leftarrow \theta_{pre}$ ▷ Pretrained with clean samples and $z_x^{\hat{y}} = 0$
  $m \leftarrow 1$
  while $m \leq N$ **do**
    for $x^i, p_g^i \in traindata$ **do**
      Predict $f_\theta(x^i)$
      $\overline{f_\theta(x)}, z_x^{\hat{y}} \leftarrow$ from Algorithm 1
      $p_{yc} = p_g(z_x^1 \cdot \overline{f_\theta(x)} + p_g(1-z_x^1)) + (1-p_g)(z_x^0 \cdot \overline{f_\theta(x)} + p_g(1-z_x^0))$
      $\mathcal{L} = p_{yc} \cdot log(f_\theta(x)) + (1-p_{yc}) \cdot log(1-f_\theta(x))$
      $\theta \leftarrow \theta - \alpha\nabla\mathcal{L}$
    end for
  end while

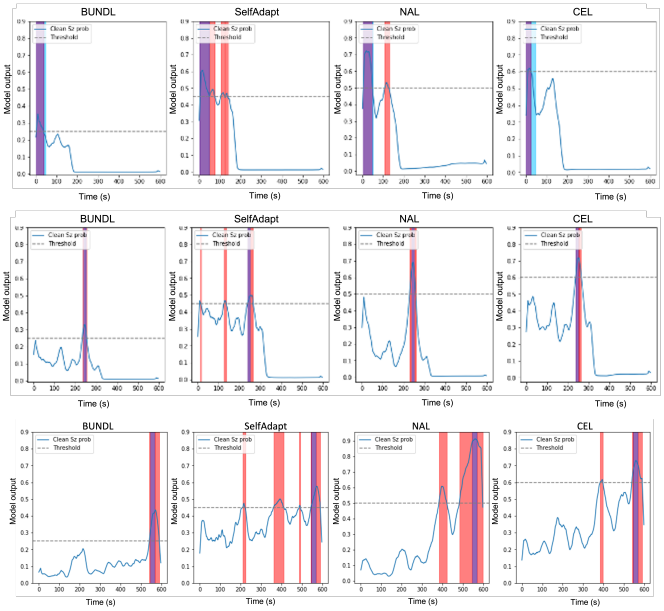

Fig. 1: Visualization of three seizures predicted by BUNDL and Baseline algorigthms. Blue line is predicted clean seizure probability p_{yc}. We denote given seizure intervals in blue, and predicted seizure intervals in red, and their intersection (correct seizure predictions) in purple.

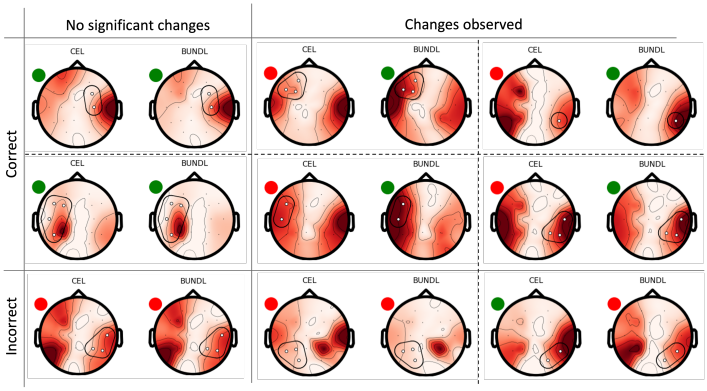

Fig. 2: SOZ localization by (a) DeepSOZ-BUNDL (b) DeepSOZ-CEL. Predicted SOZ is given by red heatmap. True SOZ is marked in black. Correct SOZ is indicated by green dot and incorrect SOZ is indicated by red dot in upper left.