Supplementary Material for Counterfactual Explanations for Graph Classification Through the Lenses of Density

No Author Given

No Institute Given

A   Other Results

Figure 1 shows the distribution of symmetric differences between original and counterfactual graphs ($d_\%$) and of computation times (I), for each method, for ADHDF (1st row), OHSU (2nd row), PEK (3rd row), and KKI (4th row).
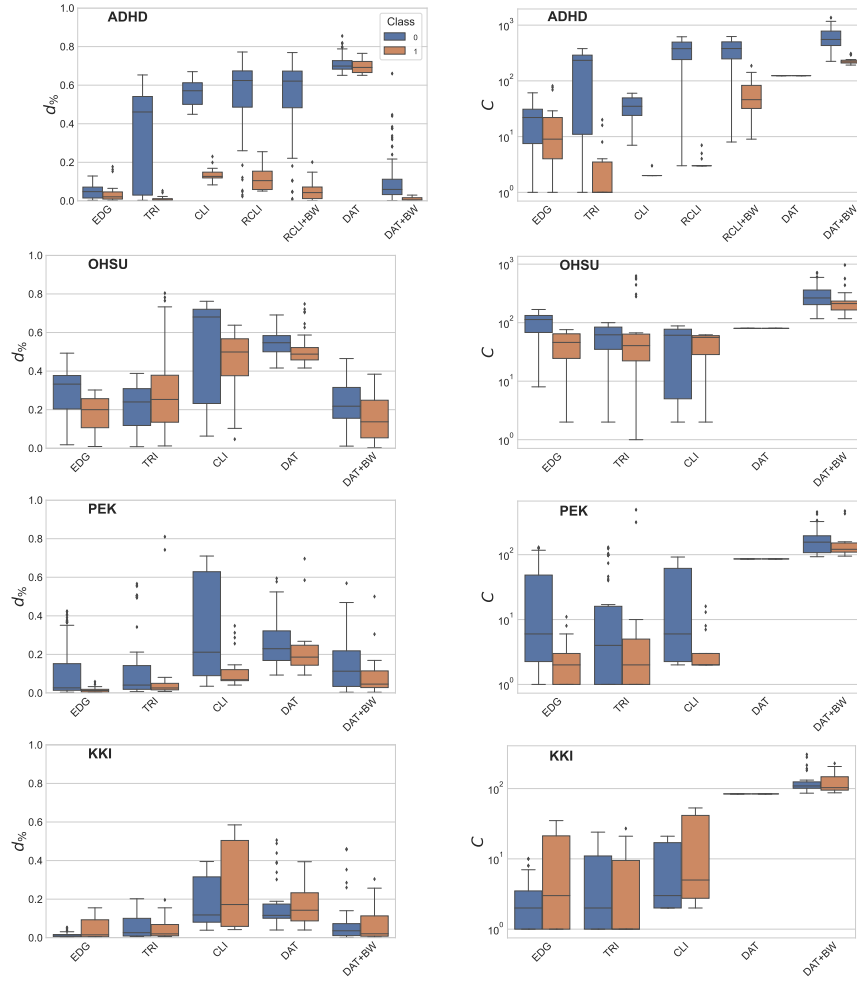
Fig. 1. Distribution of symmetric differences between original and counterfactual graphs ($d_\%$) and of computation times (I), for each counterfactual generation method, for ADHD (1st), OHSU (2nd), PEK (3rd), and KKI (4th).

Figure 2 compares the counterfactual graphs found by `EDG`, `REG`, `DAT`, and `DAT+BW`, for patient 22 in ADHD. The figure shows both the connectome of the patient with the changes highlighted in red (edges removed) and in blue (edges added), and the barplots indicating to which brain lobes the ROIs of the edges added (in blue) and removed (in red) belong (as percentages).
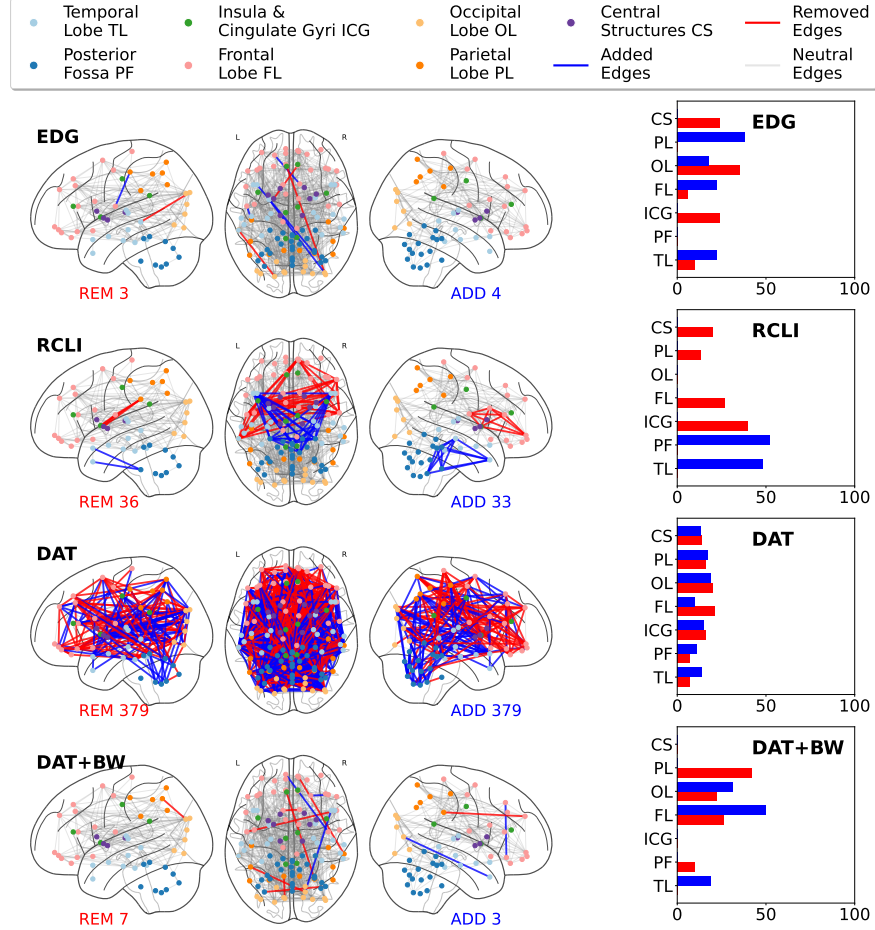

Fig. 2. Counterfactual graphs found by `EDG`, `REG`, `DAT`, and `DAT+BW`, for patient 22 in ADHD. On the left, the connectome of the patient with *connections to remove* (*REM* is the number of edges removed), and *connections to add* (*ADD* is the number of edges added) to generate the counterfactual. On the right, to which brain lobes the ROIs of the edges added (in blue) and removed (in red) belong (as percentages).