## 6 Supplementary Material

**1. Visual Comparison of CutMix and BAR** Synthetic MRIs generated by CutMix and BAR is shown in Fig 2.

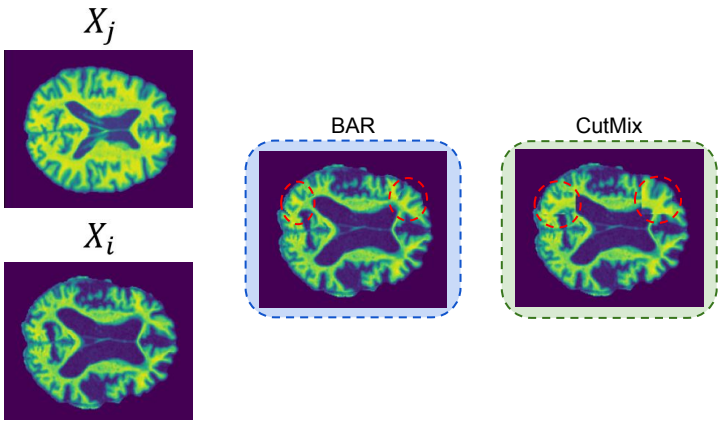


**Fig. 2.** Comparison between a sample generated with BAR and CutMix given an anchor $X_i$ and a random MRI $X_j$ on axial view. For BAR, two random regions are selected from the AAL atlas (*Superior frontal gyrus, medial orbital* and *Superior frontal gyrus, dorsolateral*). These regions are taken from $X_j$ and replaced in the same parts of $X_i$. For CutMix, square-regions are selected from similar regions on $X_j$ and replaced in $X_i$ in the same fashion (marked by red circles). Notice how BAR produces more realistic-looking synthetic MRIs as random patches often are too bulky and cutting/replacing regions from lateral ventricle.

**2. Attention Visualization of CutMix and BAR** We used Attention Rollout [24], which yields averaged attention weights across all layers and heads. We analysed two cases, an AD sample, and a CN sample. The average attention outputs are shown in Fig. 3 and Fig. 4, respectively for AD and CN (40th-70th slices are shown with increments of 10 for all views.). In both cases CutMix based model erroneously classified the given sample and BAR based model made a correct prediction. BAR attends regions more globally (Fig. 3) for the AD case and also correctly chooses not to focus on non-AD atrophy (Fig. 4).

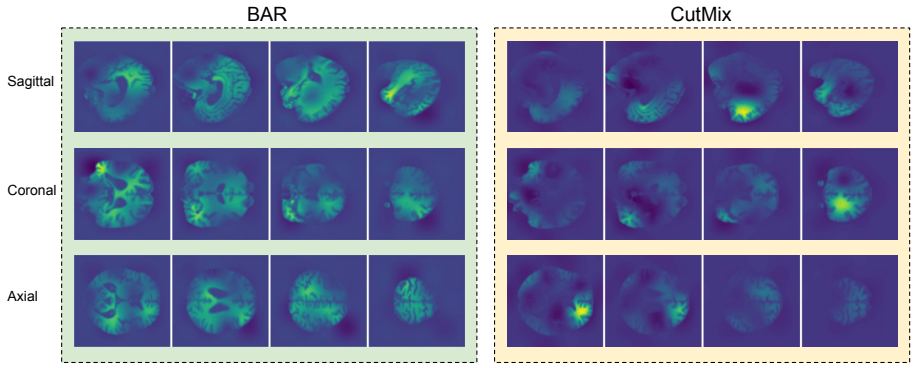


**Fig. 3.** Attention Rollout results for the AD case, slices between 40th-70th are shown with increments of 10 for all views.

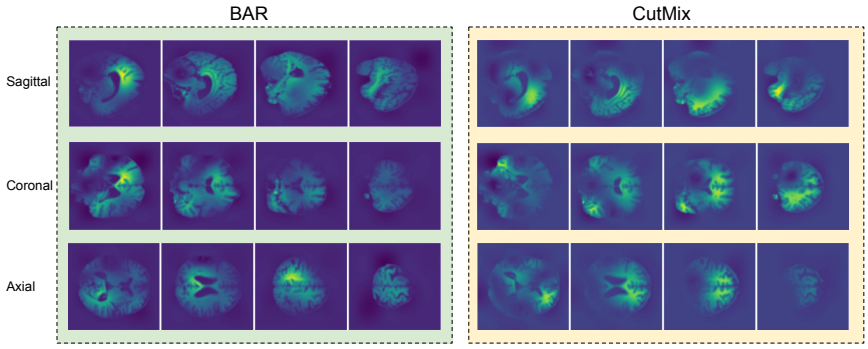


**Fig. 4.** Attention Rollout results for the CN case, slices between 40th-70th are shown with increments of 10 for all views.