A Appendix

Table 3: Text-to-Brain: Comparison between NeuroConText method, and the two baselines Text2brain and NeuroQuery. NeuroConText is performed on Mistral-7B and DiFuMo size 512. The results conclusively demonstrate that NeuroConText outperforms the baseline across all parts of the articles—Title, Abstract, and Body—underscoring its superiority in associating text to brain.

Method \ Metric [%]		recall@10	recall@100	mix&match
Title	NeuroConText (ours)	$\mathbf{9.4 \pm 2.6}$	$\mathbf{32.8 \pm 6.2}$	$\mathbf{68.7 \pm 3.6}$
	NeuroQuery	4 ± 1.6	26 ± 2.2	63.8 ± 1.3
	Text2brain	2 ± 0.6	13 ± 1.6	53.8 ± 0.5
Abstract	NeuroConText (ours)	$\mathbf{17.5 \pm 0.9}$	$\mathbf{48.9 \pm 1.9}$	$\mathbf{79.6 \pm 0.8}$
	NeuroQuery	6 ± 1.3	39 ± 2.8	75.7 ± 1.9
	Text2brain	2 ± 0.9	15 ± 0.9	55 ± 0.9
Body	NeuroConText (ours)	$\mathbf{22.6 \pm 1.4}$	$\mathbf{57.8 \pm 1.6}$	$\mathbf{84.2 \pm 0.9}$
	NeuroQuery	7 ± 0.6	38.7 ± 1.6	75.5 ± 2.5
	Text2brain	1.4 ± 0.2	12.6 ± 0.3	53.3 ± 0.2

Fig. 4: The impact of training data size on the NeuroConText performance for 10-fold cross-validation, setting 1K samples for the test set. Incorporating the articles published in recent years and expanding the data size from 10k to 19k led to an improvement in the recall@10 metric from 17.8% to 22.3%. This represents a 4.5% enhancement in associating text to the brain.