Aligning Neuronal Coding of Dynamic Visual Scenes with Foundation Vision Models (Supplementary Materials)

Rining Wu[1,2], Feixiang Zhou[3], Ziwei Yin[2], and Jian K. Liu[1,2]

[1] University of Leeds ml20r2w@leeds.ac.uk
[2] University of Birmingham {z.yin.3, j.liu.22}@bham.ac.uk
[3] Lancaster University f.zhou3@lancaster.ac.uk

1 SD-KL Metric

For considering the duration of spikes, we designed a metric called Spike Duration — Kullback-Leibler Divergence (SD-KL). The pseudocode of SD-KL is demonstrated in Algorithm 1.

Algorithm 1 Pseudocode of the SD-KL

Input: $\hat{y} \in \mathbb{R}^{N \times F}, y \in \mathbb{R}^{N \times F}, \alpha = 0.3, \beta = 1.0$
Output: score

$\hat{\mathcal{D}} \leftarrow$ **peak widths**$(\min(\max(0, \hat{y}), \beta)), \hat{\mathcal{D}} \subset \mathbb{R}$
$\mathcal{D} \leftarrow$ **peak widths**$(\min(\max(0, y), \beta)), \mathcal{D} \subset \mathbb{R}$
$Var_{\cup} \leftarrow \frac{\alpha}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2, x_i \in \{D, \hat{D}\}$
$\mathcal{L}_{\cup} \leftarrow \left\{ (lower_{\cup} - 3 * Var_{\cup}) + i \cdot \frac{(upper_{\cup} + 3 * Var_{\cup}) - (lower_{\cup} - 3 * Var_{\cup})}{199} \mid i = 0, 1, \ldots, 199 \right\}$
$pdf_{\mathcal{D}} \leftarrow$ **KDE**$(\mathcal{D}, \alpha)$
$pdf_{\hat{\mathcal{D}}} \leftarrow$ **KDE**$(\hat{\mathcal{D}}, \alpha)$
$P_{\mathcal{D}} = \left\{ \left(x_i, \frac{pdf_{\mathcal{D}}(x_i)}{\sum_{j=1}^{N} pdf_{\mathcal{D}}(x_j)} \right) \mid x_i \in \mathcal{L}_{\cup} \right\}$
$P_{\hat{\mathcal{D}}} = \left\{ \left(x_i, \frac{pdf_{\hat{\mathcal{D}}}(x_i)}{\sum_{j=1}^{N} pdf_{\hat{\mathcal{D}}}(x_j)} \right) \mid x_i \in \mathcal{L}_{\cup} \right\}$
score $\leftarrow D_{KL}(P_{\hat{\mathcal{D}}} || P_{\mathcal{D}})$
score $\leftarrow \min(\max(0, \text{score}), 1000)$
return score

Where α is the bandwidth of the kernel density estimation (KDE), set to 0.3, β is the high cut of the spike duration, set to 1.0, N is the number of RGCs, F is the number of frames, and $y, \hat{y}$ are the ground truth and predicted FR, respectively. Furthermore, the *peak widths* calculates durations of each spike then returns the set of durations [1]. The KDE algorithm calculates the probability density function (PDF) of the input set of durations with a given bandwidth α.

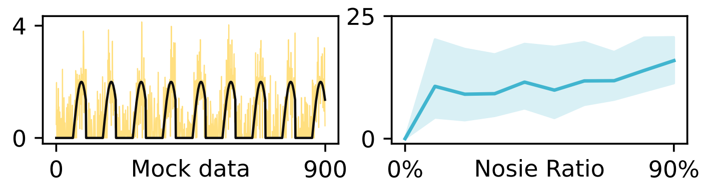
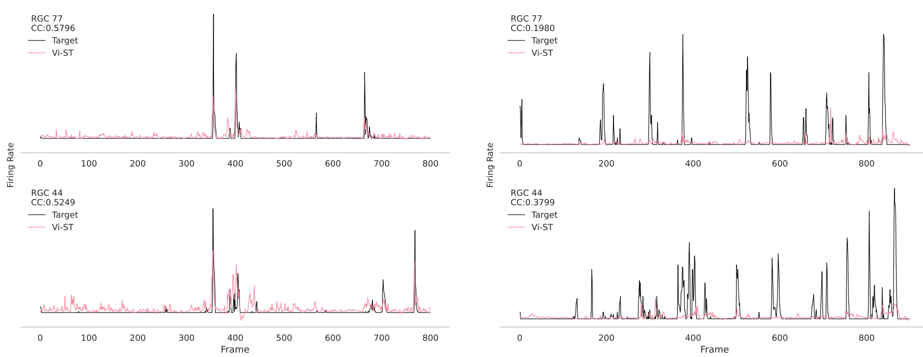


Fig. 1: SD-KL Metric Baseline

We selected several signals simulated using *the sine function* and added random noise at different ratios to obtain the baseline of the SD-KL metric, as Fig. 1.

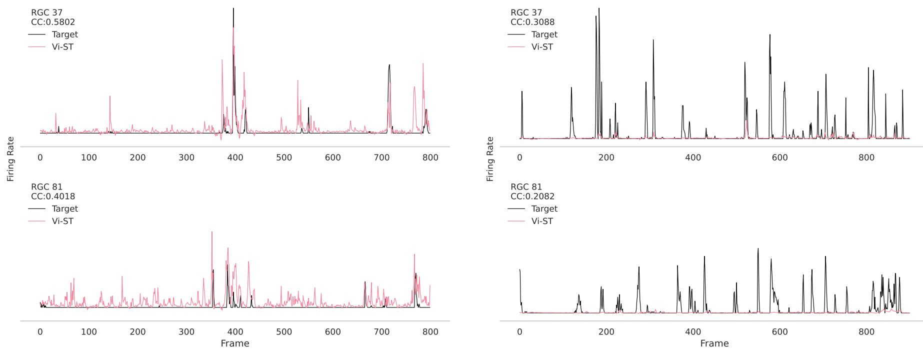
2 Experimental Results

We choose the best, middle, and worst cases of the CC to demonstrate the real predictions of the Vi-ST.



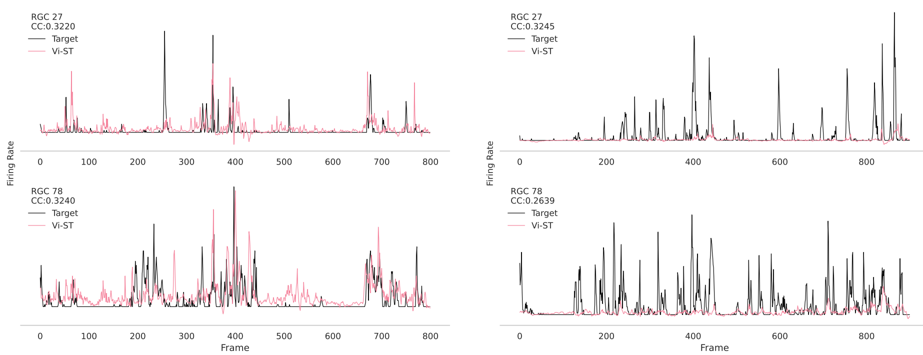
(a) *Mov1→Mov2* **(b)** *Mov2→Mov1*

Fig. 2: Real Predictions of good CC



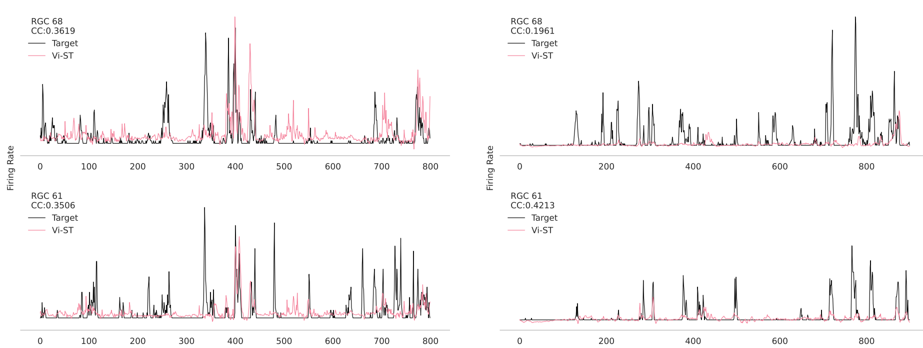
(a) *Mov1→Mov2* **(b)** *Mov2→Mov1*

Fig. 3: Real Predictions of good CC



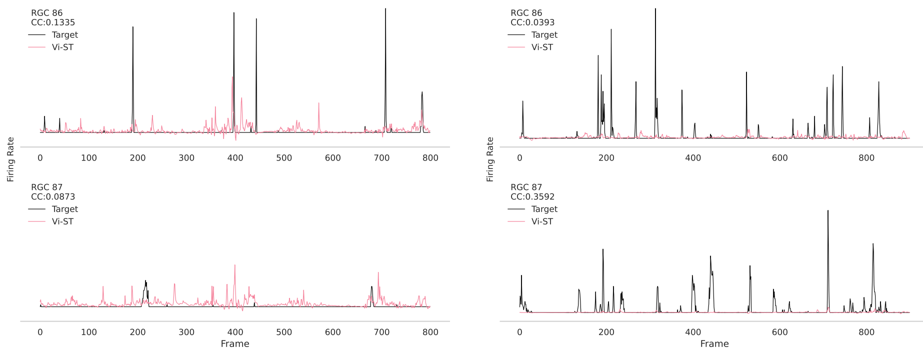
(a) *Mov1→Mov2* **(b)** *Mov2→Mov1*

Fig. 4: Real Predictions of middle CC



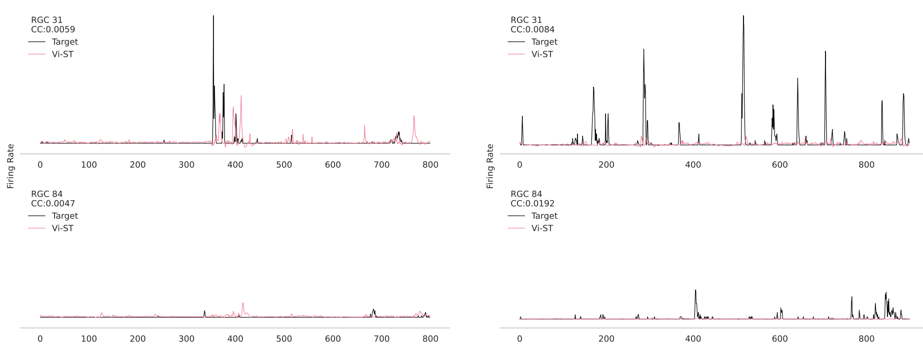
(a) *Mov1→Mov2* **(b)** *Mov2→Mov1*

Fig. 5: Real Predictions of middle CC



(a) *Mov1→Mov2* **(b)** *Mov2→Mov1*

Fig. 6: Real Predictions of worst CC



(a) *Mov1→Mov2* **(b)** *Mov2→Mov1*

Fig. 7: Real Predictions of worst CC

References

1. Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S.J., Brett, M., Wilson, J., Millman, K.J., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Carey, C.J., Polat, İ., Feng, Y., Moore, E.W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E.A., Harris, C.R., Archibald, A.M., Ribeiro, A.H., Pedregosa, F., van Mulbregt, P., SciPy 1.0 Contributors: SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods **17**, 261–272 (2020). https://doi.org/10.1038/s41592-019-0686-2