

Supplementary Information for

***This Looks Like That There*: Interpretable neural networks for image tasks when location matters**

Elizabeth A. Barnes, Randal J. Barnes, Zane K. Martin and Jamin K. Rader (2022)

Elizabeth A. Barnes

E-mail: eabarnes@colostate.edu

This PDF file includes:

Figs. S1 to S13

Table S1

Table S1. MJO use case validation accuracy of the ProtoLNet and its associated base CNN for seven different random seeds. The random seeds set the training/validation split of the different years as well as the network initialization. The bold row (random seed 30) is the ProtoLNet shown in the main text.

random seed	base CNN	ProtoLNet
28	81%	74%
29	60%	75%
30	58%	73%
31	77%	74%
32	82%	75%
33	81%	76%
34	74%	73%

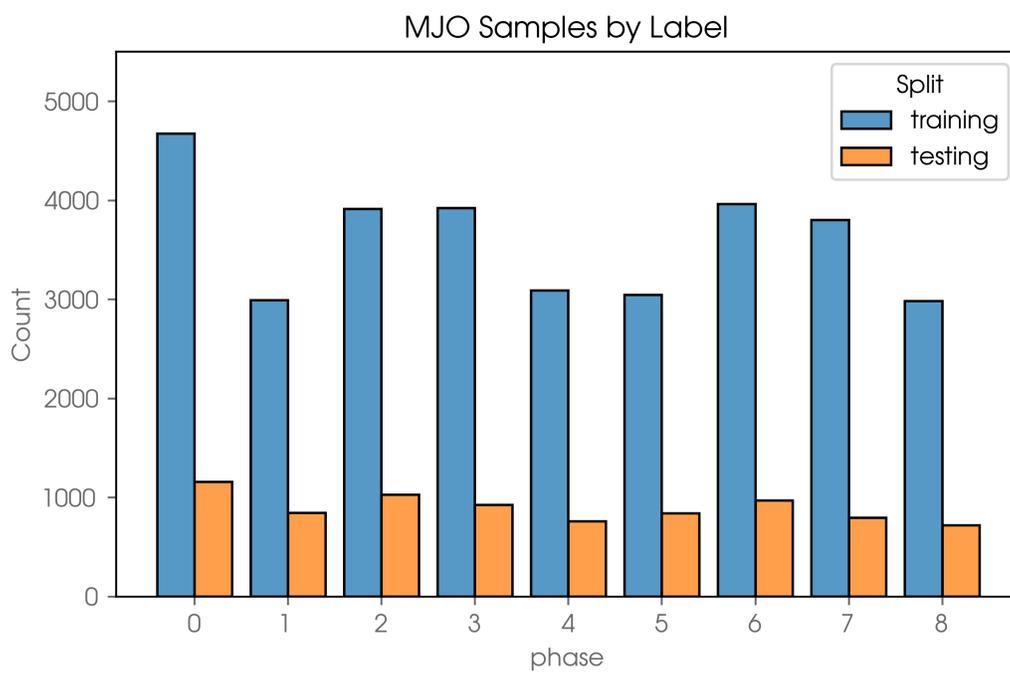


Fig. S1. Number of samples per MJO phase in training and testing sets.

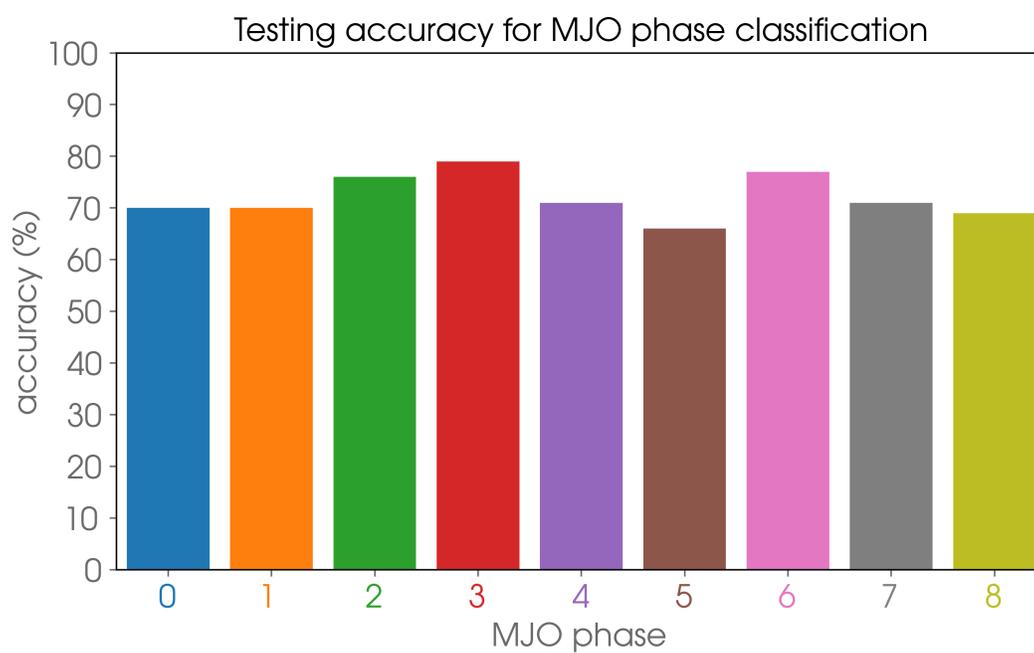


Fig. S2. Testing accuracy as a function of MJO phase.

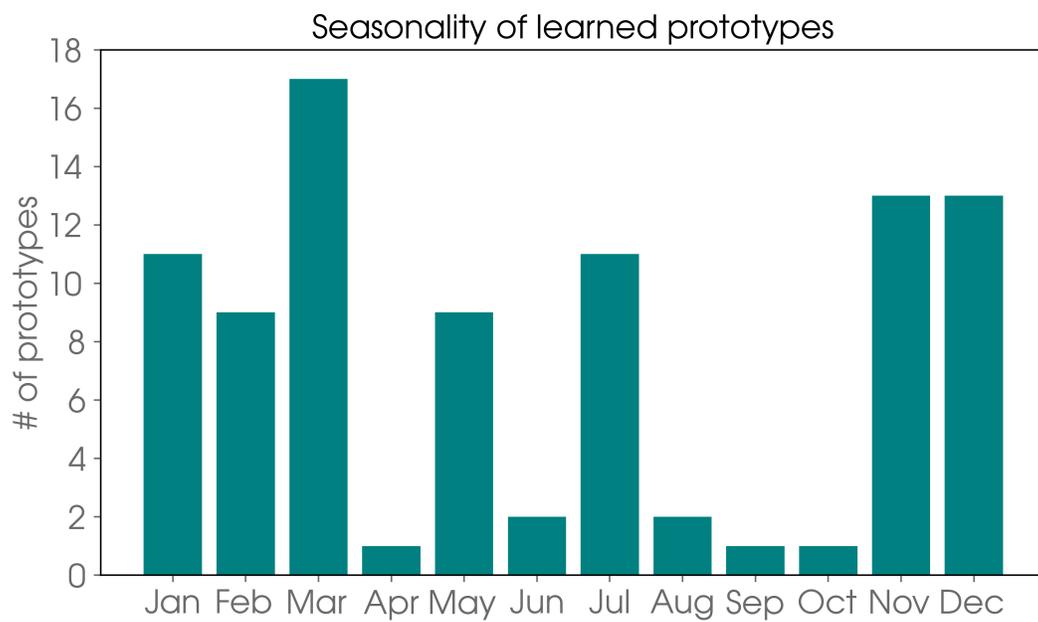


Fig. S3. Number of learned prototypes (out of 90 total) binned by month of the year of the training sample from which the prototype was drawn.

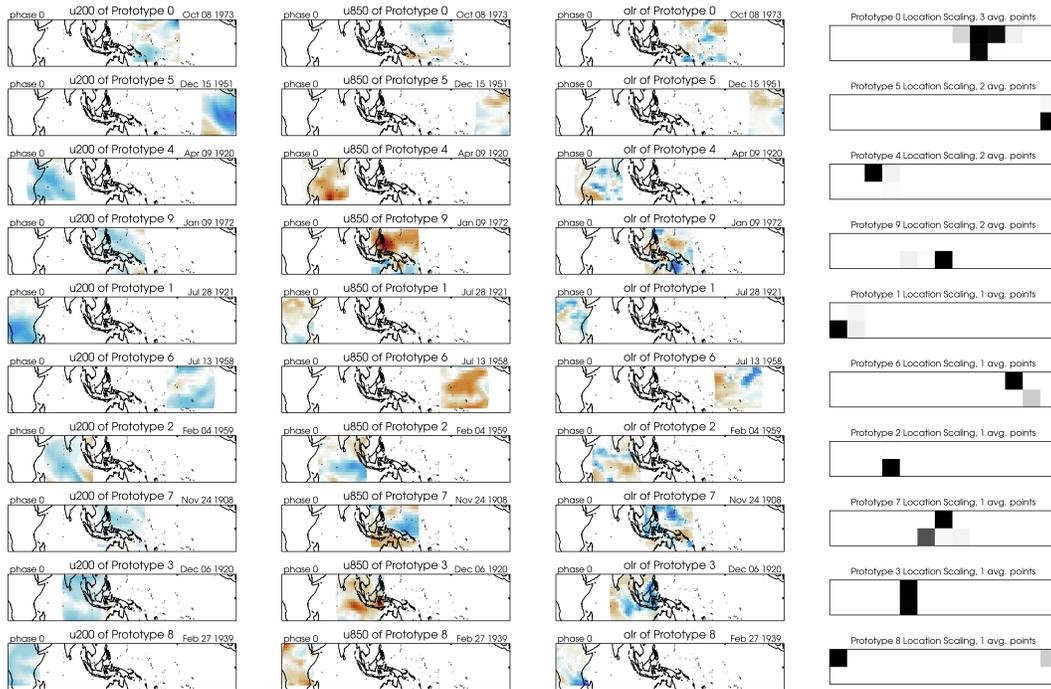


Fig. S4. All prototypes for MJO phase 0, ordered by the average number of points contributed across correctly classified testing samples. The average number of points and the percent time the prototype is the winning prototype are printed in the upper-left and upper-right corners of each location scaling grid, respectively.

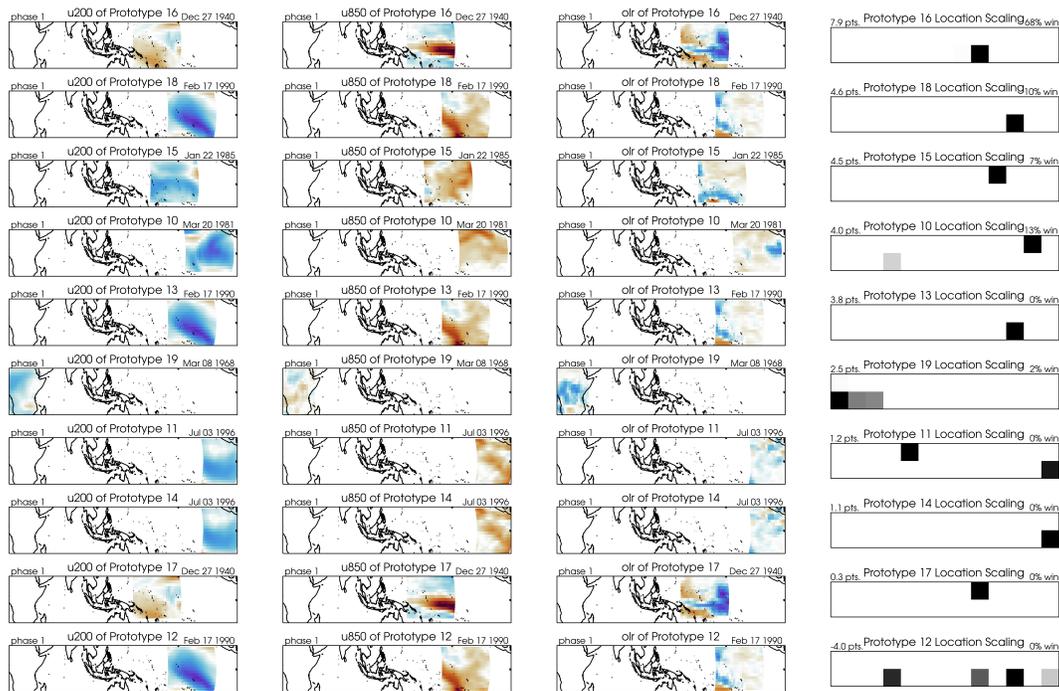


Fig. S5. All prototypes for MJO phase 1, ordered by the average number of points contributed across correctly classified testing samples. The average number of points and the percent time the prototype is the winning prototype are printed in the upper-left and upper-right corners of each location scaling grid, respectively.

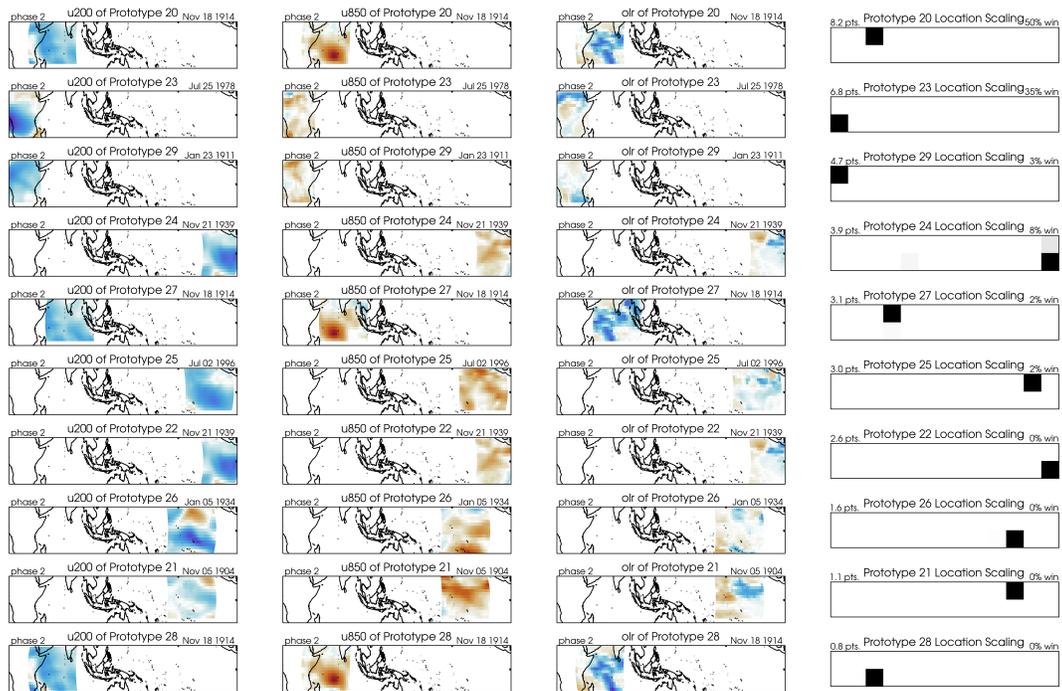


Fig. S6. All prototypes for MJO phase 2, ordered by the average number of points contributed across correctly classified testing samples. The average number of points and the percent time the prototype is the winning prototype are printed in the upper-left and upper-right corners of each location scaling grid, respectively.

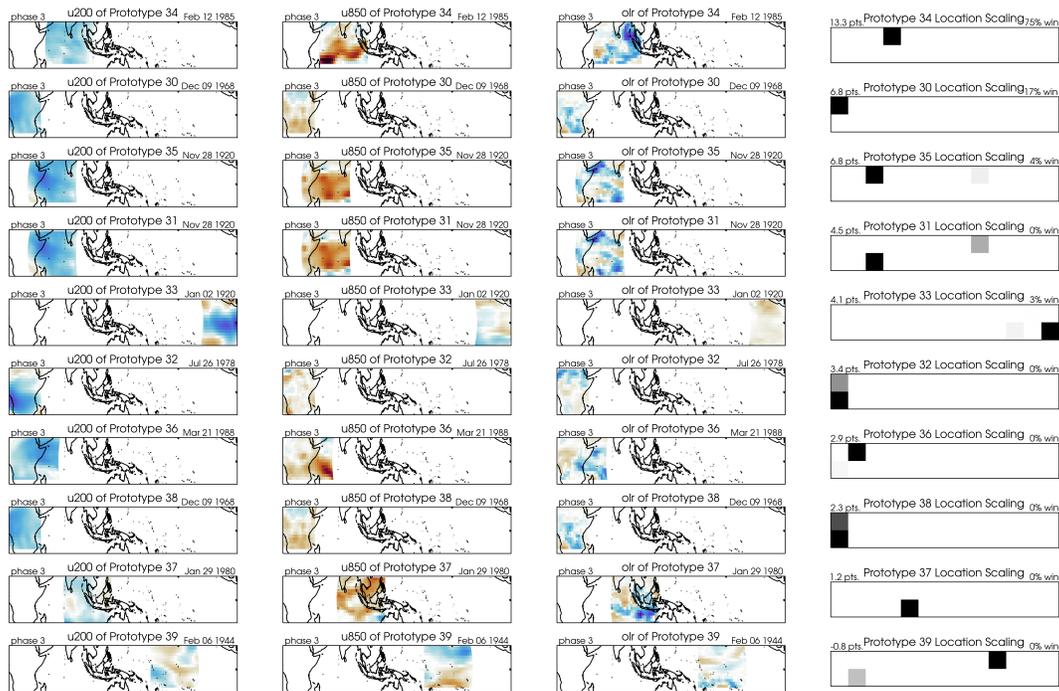


Fig. S7. All prototypes for MJO phase 3, ordered by the average number of points contributed across correctly classified testing samples. The average number of points and the percent time the prototype is the winning prototype are printed in the upper-left and upper-right corners of each location scaling grid, respectively.



Fig. S8. All prototypes for MJO phase 4, ordered by the average number of points contributed across correctly classified testing samples. The average number of points and the percent time the prototype is the winning prototype are printed in the upper-left and upper-right corners of each location scaling grid, respectively.

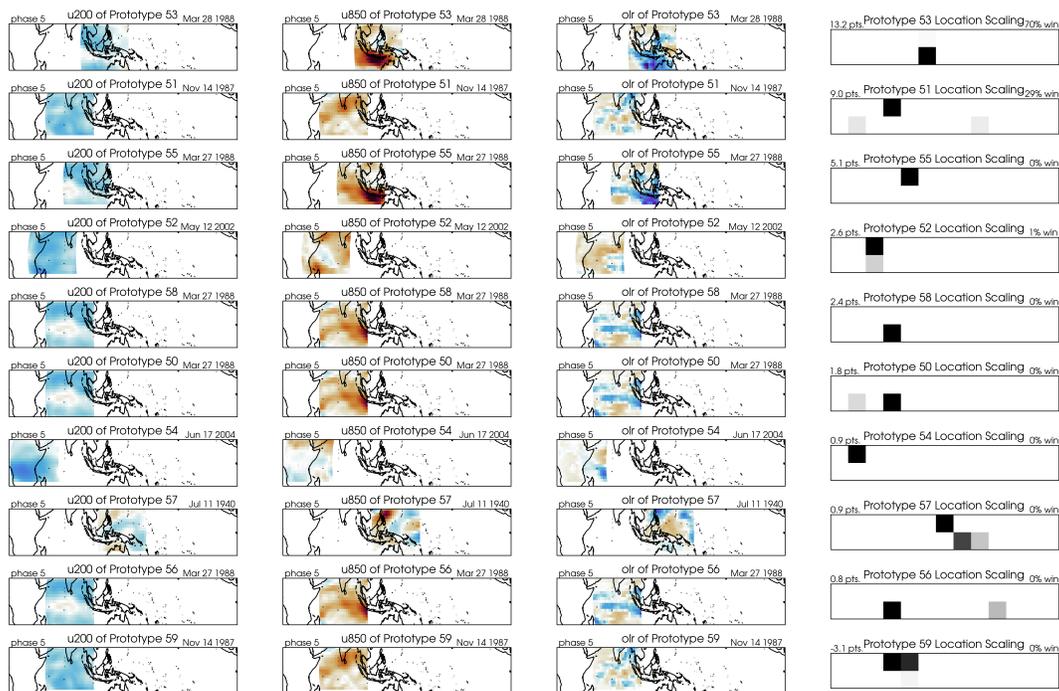


Fig. S9. All prototypes for MJO phase 5, ordered by the average number of points contributed across correctly classified testing samples. The average number of points and the percent time the prototype is the winning prototype are printed in the upper-left and upper-right corners of each location scaling grid, respectively.

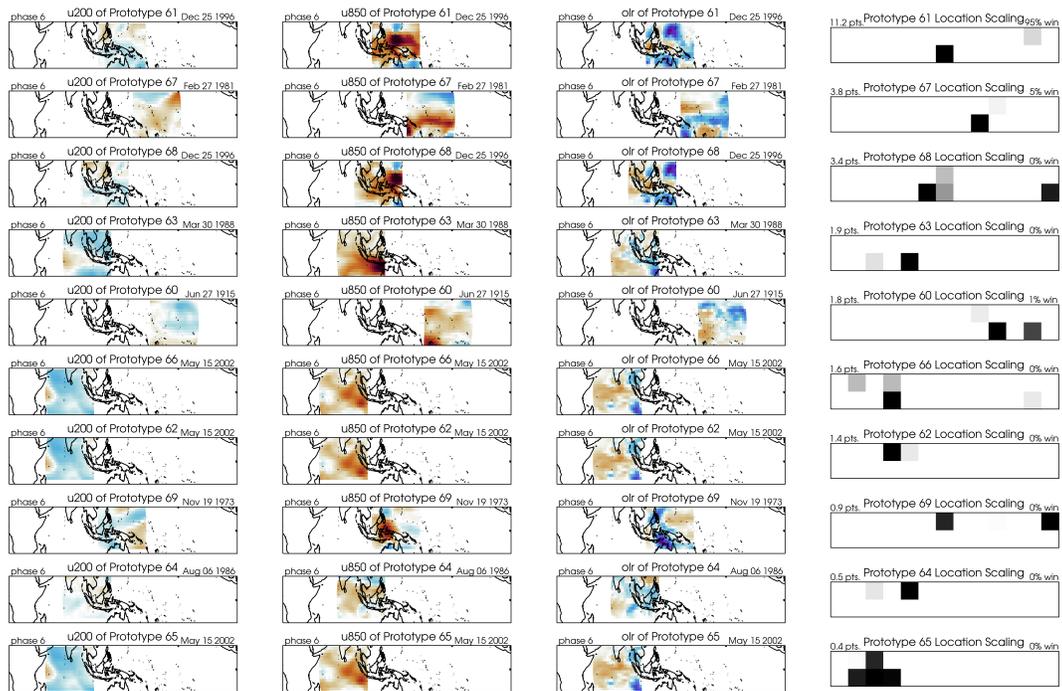


Fig. S10. All prototypes for MJO phase 6, ordered by the average number of points contributed across correctly classified testing samples. The average number of points and the percent time the prototype is the winning prototype are printed in the upper-left and upper-right corners of each location scaling grid, respectively.

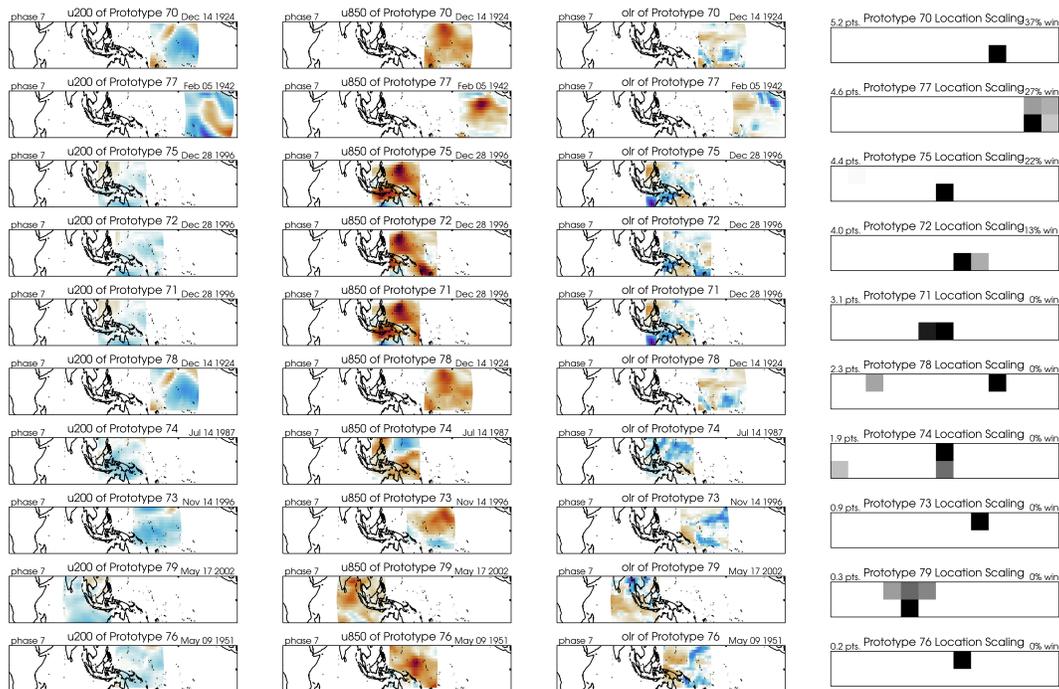


Fig. S11. All prototypes for MJO phase 7, ordered by the average number of points contributed across correctly classified testing samples. The average number of points and the percent time the prototype is the winning prototype are printed in the upper-left and upper-right corners of each location scaling grid, respectively.

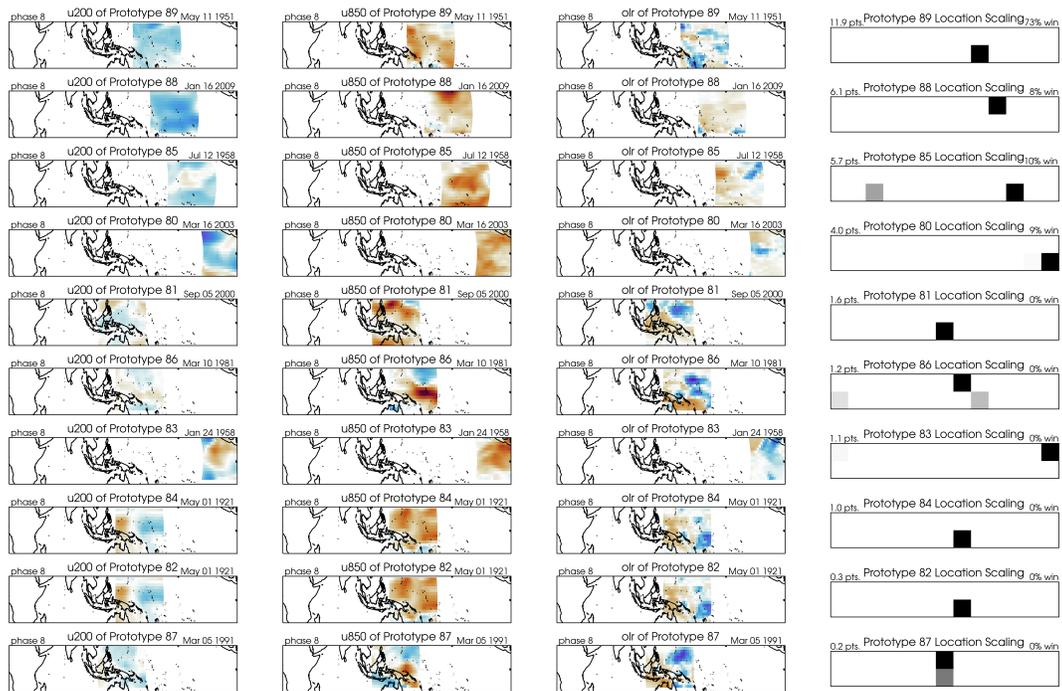


Fig. S12. All prototypes for MJO phase 8, ordered by the average number of points contributed across correctly classified testing samples. The average number of points and the percent time the prototype is the winning prototype are printed in the upper-left and upper-right corners of each location scaling grid, respectively.

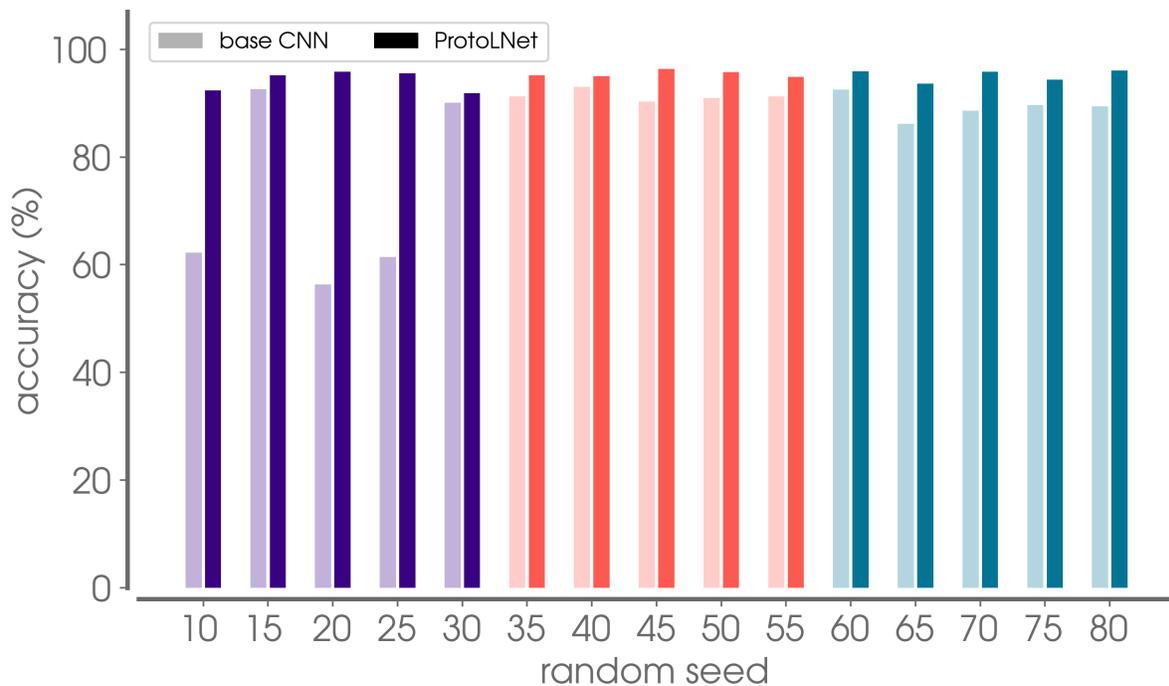


Fig. S13. Testing accuracy comparison for the idealized quadrants use case with a reduced training size of 1,400 samples. When training the base CNN, random seeds 10-30 (purple) use a dropout rate of 0.0 on the fully connected layer, random seeds 35-55 (peach) use a dropout rate of 0.2, and random seeds 60-80 (teal) use a dropout rate of 0.5. Dropout is not used when training the associated ProtoLNet. In all instances, the ProtoLNet exhibits improved accuracy over the base CNN when evaluated on 3,000 testing samples.