How bio-inspired attention affects task performance in visual and auditory models

Bogeng Song¹, Grace Lindsay^{1,2} ¹Department of Psychology, New York University; ²Center of Data Science, New York University

Summary:

In the visual model, a gain-based attention mechanism can effectively improve task performance, especially when applied at higher levels. However, for the auditory model, no similar results were found.

BACKGROUND

Previous studies show that visual attention mechanisms in earlier layers do not improve task performance as much as higher-level regions¹

For the auditory task, The Cocktail Party effect elegantly illustrates humans' ability to "tune in" to one conversation in a noisy scene. Selective attention also plays a role in this essential cognitive capacity²

However, it is currently unclear whether attentional mechanisms

Attention Mechanism

Get tuning curves for each of the feature maps at a layer:

- 1. Take the spatial average of the feature map in response to images.
- 2. Use response values to calculate normalized tuning curves.





Calculate the mean value of the kth

to all Images N.

feature map in the lth layer in response

Calculate the spatial average of each feature map



Calculate the tuning value of the kth feature map in the lth layer for category c.



improve task performance in auditory tasks in a similar way to visual

attention.

METHOD

Detection Task

Visual Detection task: ImageNet Merged Image



Report if a specific category is present

Auditory Detection task: **AudioMNIST** Merged Voice

Report if a specific speaker is present

Modulate activity according to tuning values:

1. Scale neural activity for an entire feature map by its tuning value for the attended category (and an overall strength parameter)

 $x_{ij}^{lk} = \left(1 + \beta f_c^{lk}\right) \left[\left(I_{ij}^{lk}\right) \right]_+$ Activity

Activity with attention (of Overall unit at the i,j spatial strength location of the kth feature map in the lth parameter layer)

Tuning value of without the kth feature attention map in the lth layer for the attended category c



Activity Layer (ReLu function), Change the slope based on tunning value

Manipulate the model with Attention Mechanism:

We add the attention mechanism to one layer of the neural network each time.

RESULT

Readout Task

Visual Readout task: Array Image



Read out the position of specific category

Auditory Readout task: Merged Voice



Read out the number the specific speaker says

Detection Task

0.78

0.76

0.74

0.72

0.70





Model Selection

Visual Model:

VGG16 ³ With 13 Convolution layers with 3 full linear layers.





Auditory Model:





In visual tasks, the attention mechanism effectively improves the task performance. When the attention mechanism is added to a higher layer, the task performance improve more.

In the auditory task, the attention mechanism did not improve task performance in both tasks...

CONCLUSIONS

1. We verified that the visual attention mechanism can effectively improve task

performance, but the auditory attention mechanism does not improve task performance.

2. In both visual tasks, when the attention mechanism is added to the higher layer, the

References

1. Lindsay, G. W., & Miller, K. D. 2018









improvement of task performance is better than when the attention mechanism is added

