

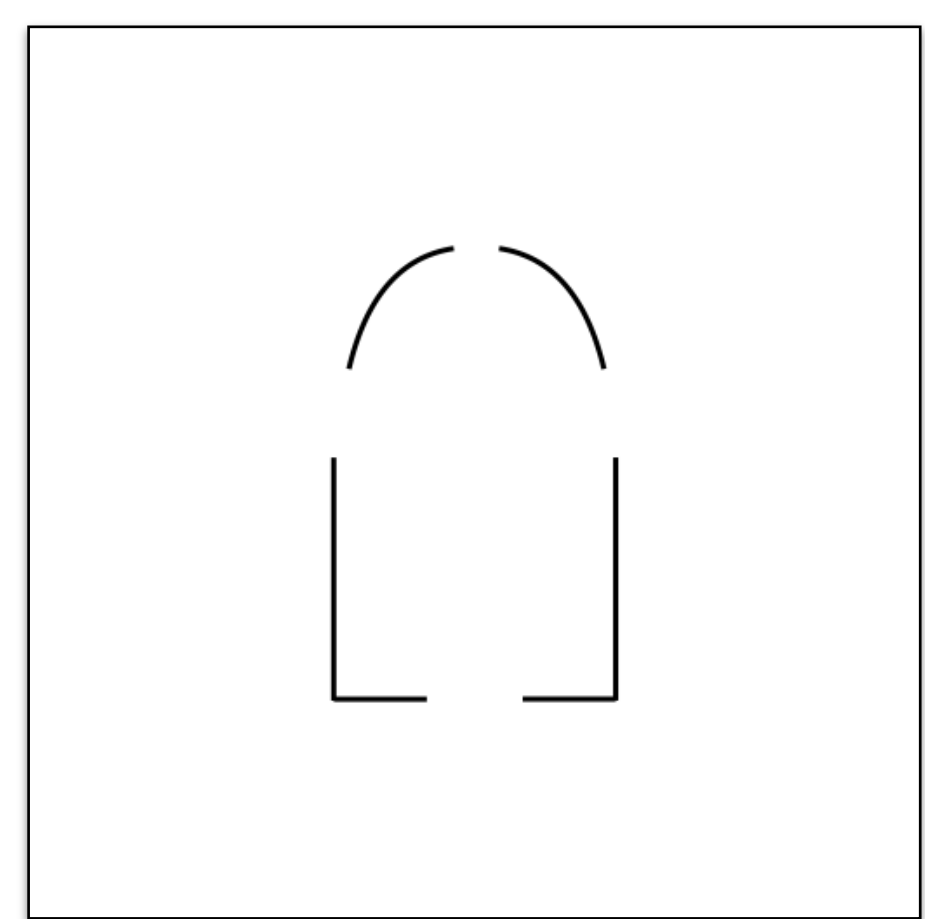


Image-generation models exhibited higher perceptual similarity but lower originality compared to humans when generating from incomplete shapes

Yaxin Liu¹, Maxwell Kay¹, Kibum Moon¹, Roger Beaty², Adam Green¹
¹Georgetown University, ²Penn State University

Motivation

Input incomplete shape¹



"Complete the drawing by incorporating the existing lines in your original doodle"

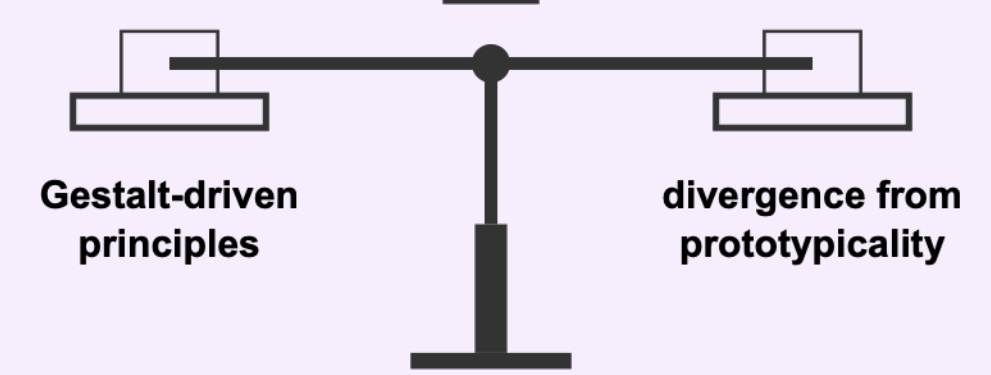
Humans "fill in the blank" by inhibiting automatic Gestalt closure



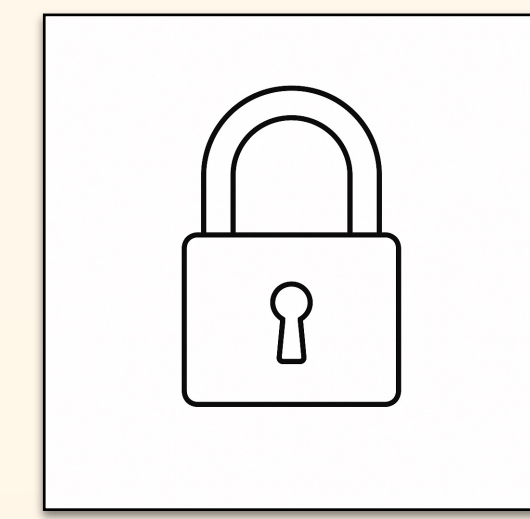
Canonical/prototypical



Divergent/unique



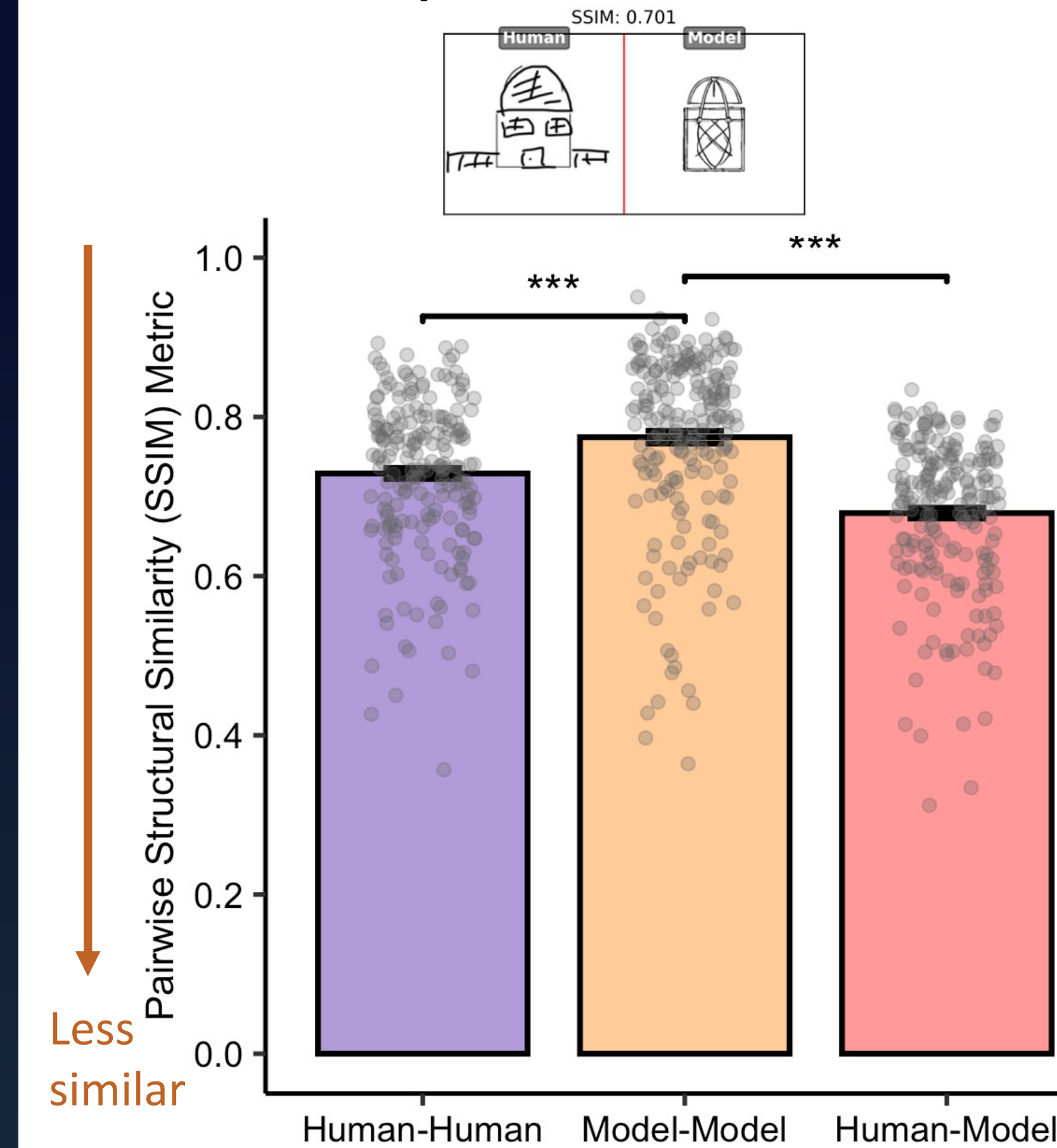
Machines 'fill in the blank' by optimizing/inhibiting statistical patterns



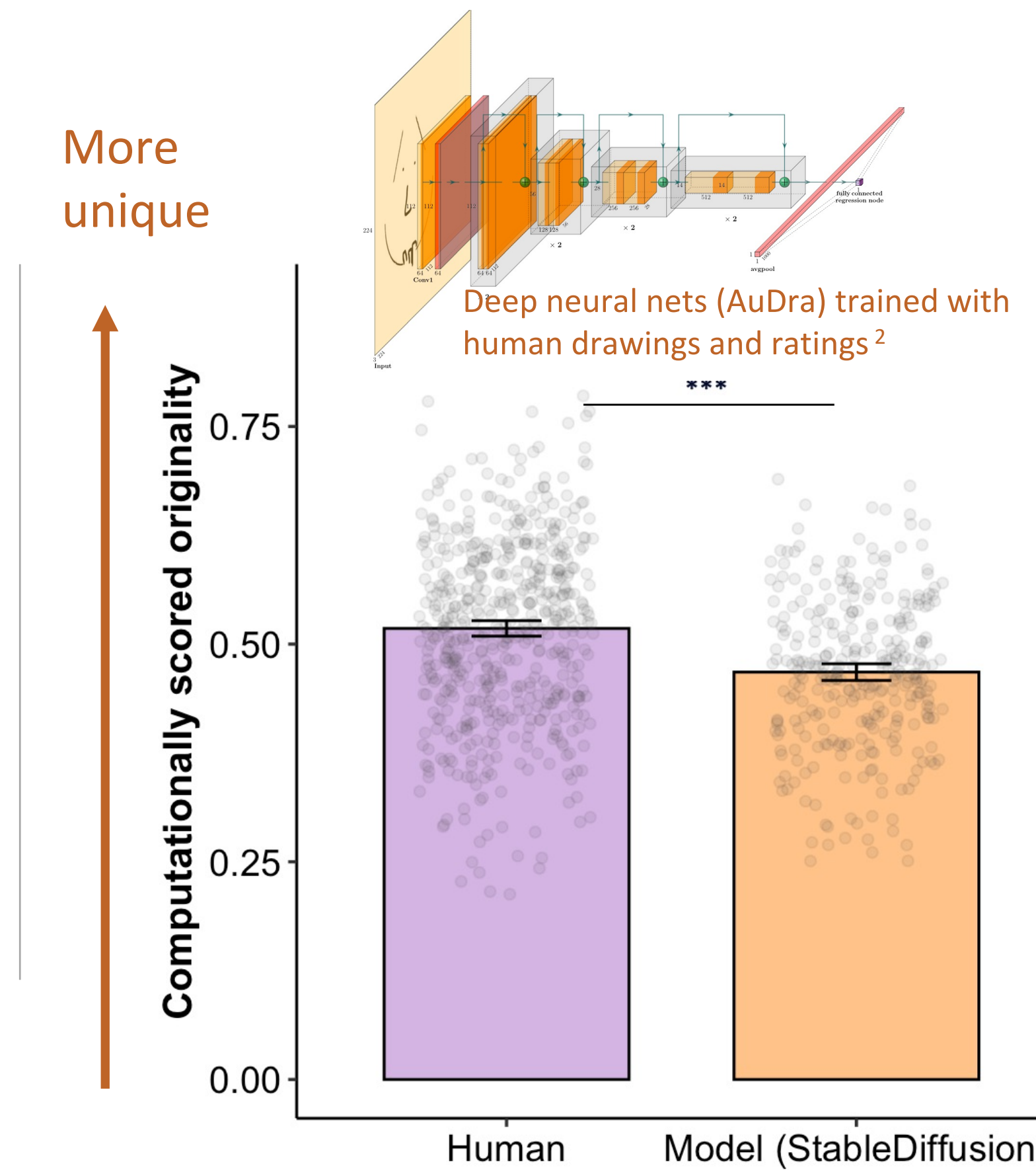
- How do human sketches and machine sketches differ?
- How might controlling global/local attention mechanisms affect machine-generated images?

Human vs. Generative model

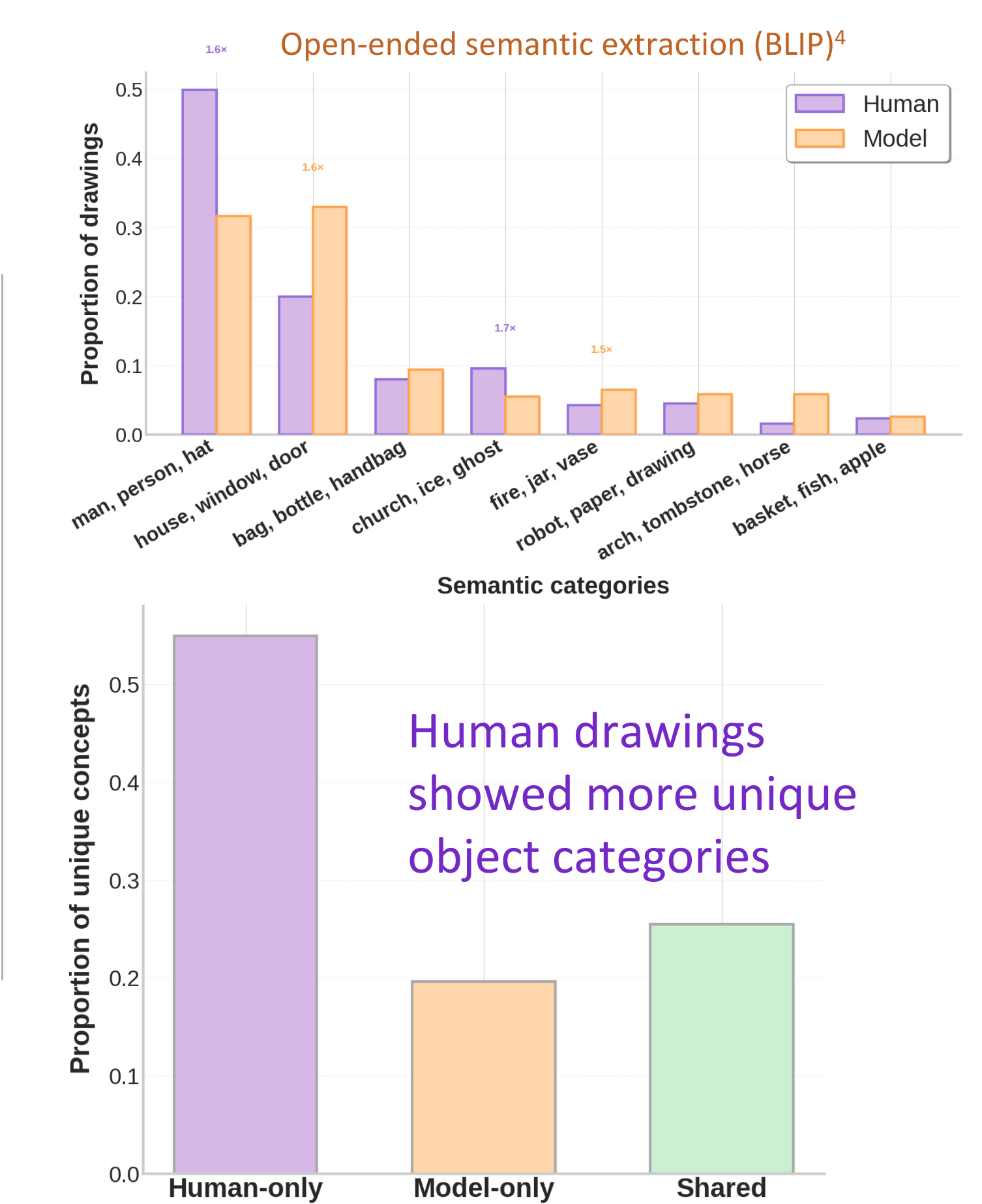
Perceptual Similarities



Uniqueness Scoring



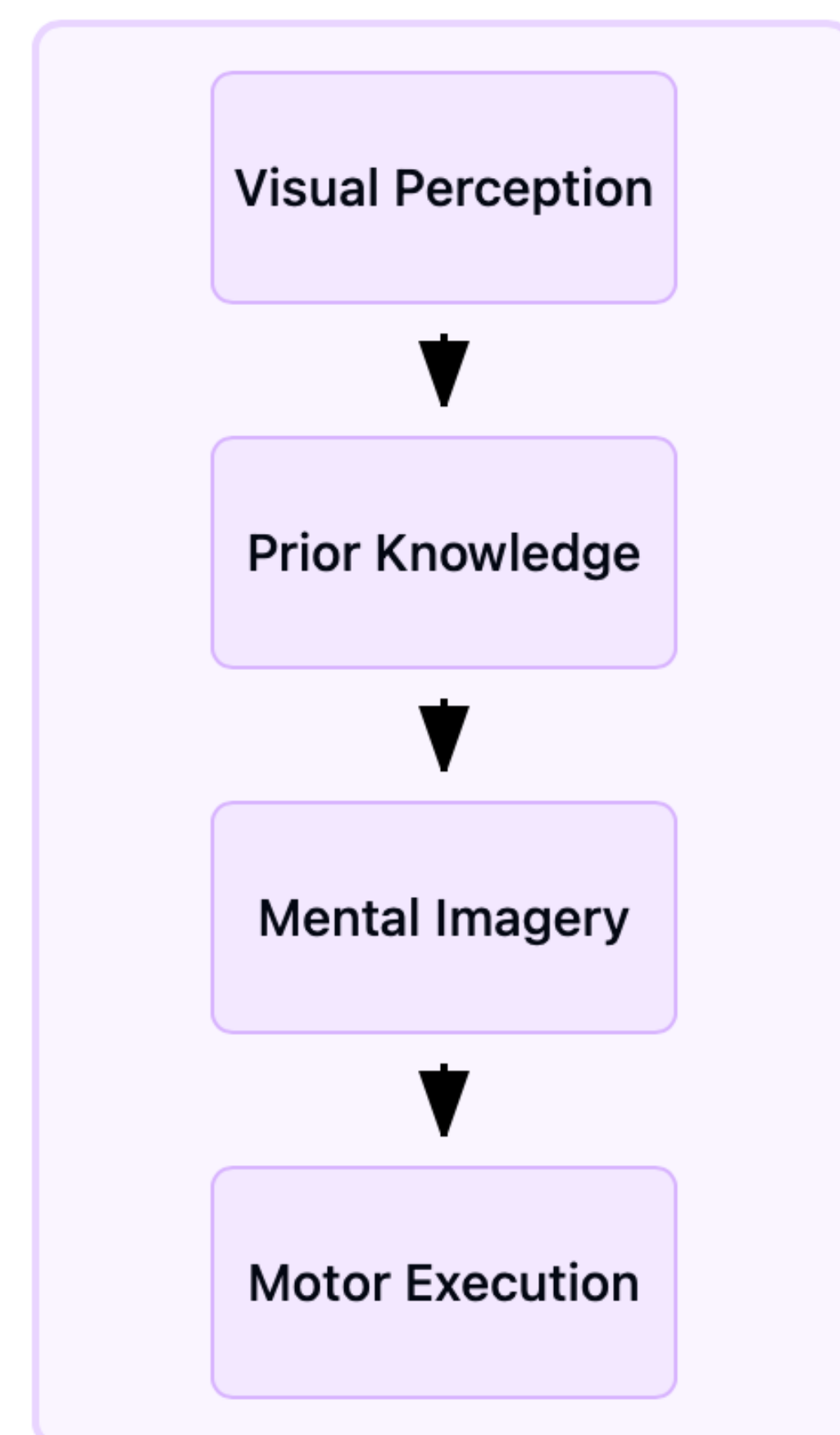
Semantic Similarities



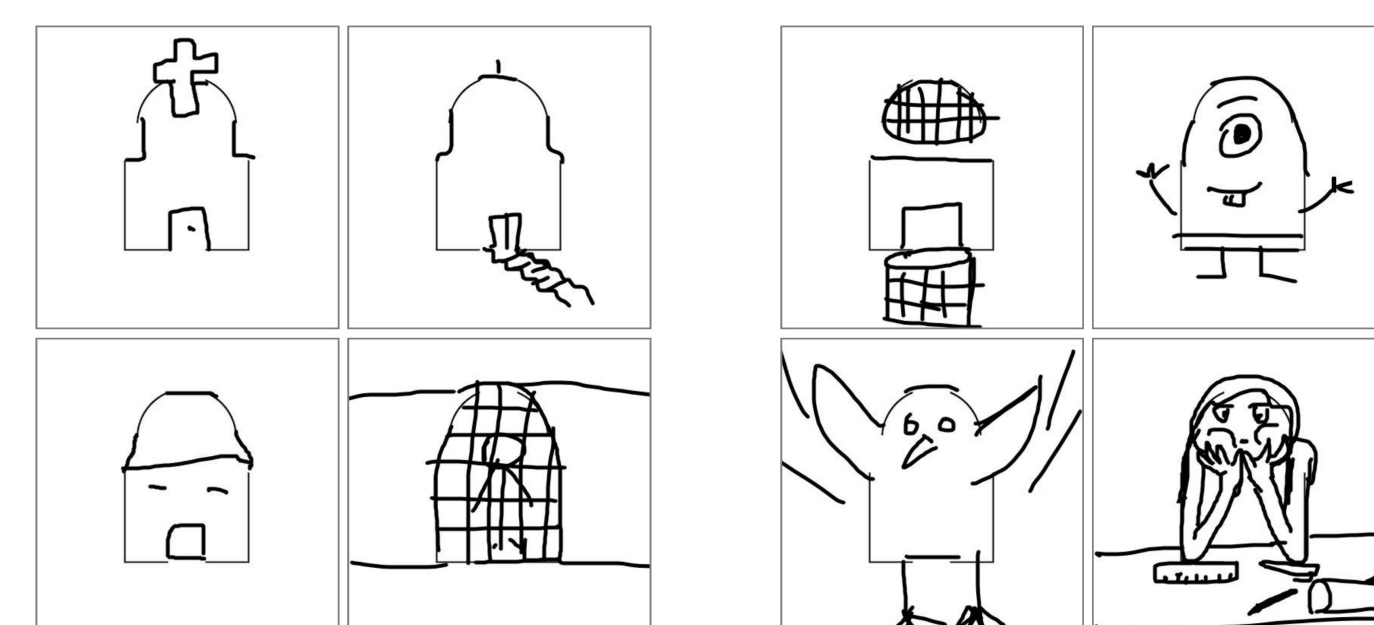
Computational experiment

Human Sketches

College students completed drawings digitally²



Examples

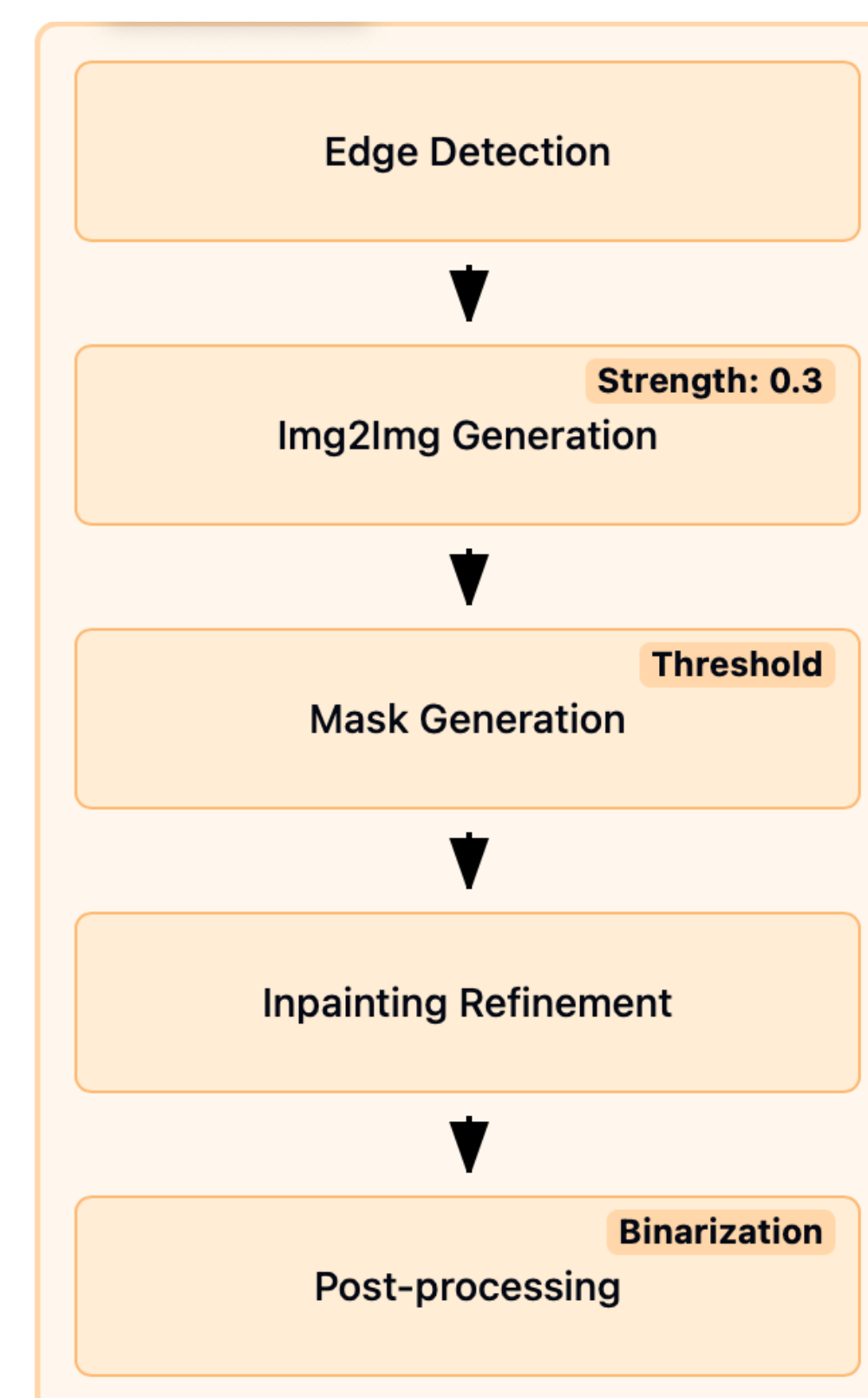


Low uniqueness High uniqueness

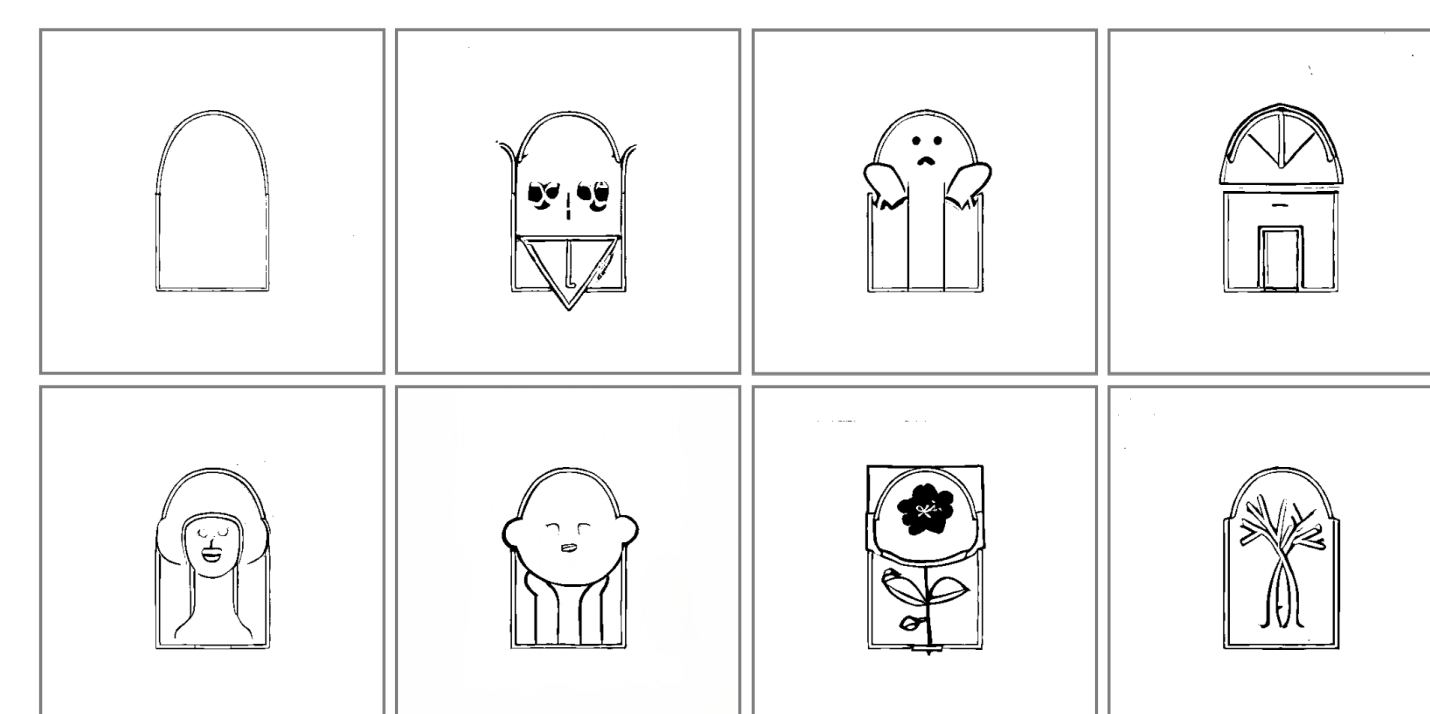
Off-the-shelf Generative Models

StableDiffusion model³

Processing pipeline



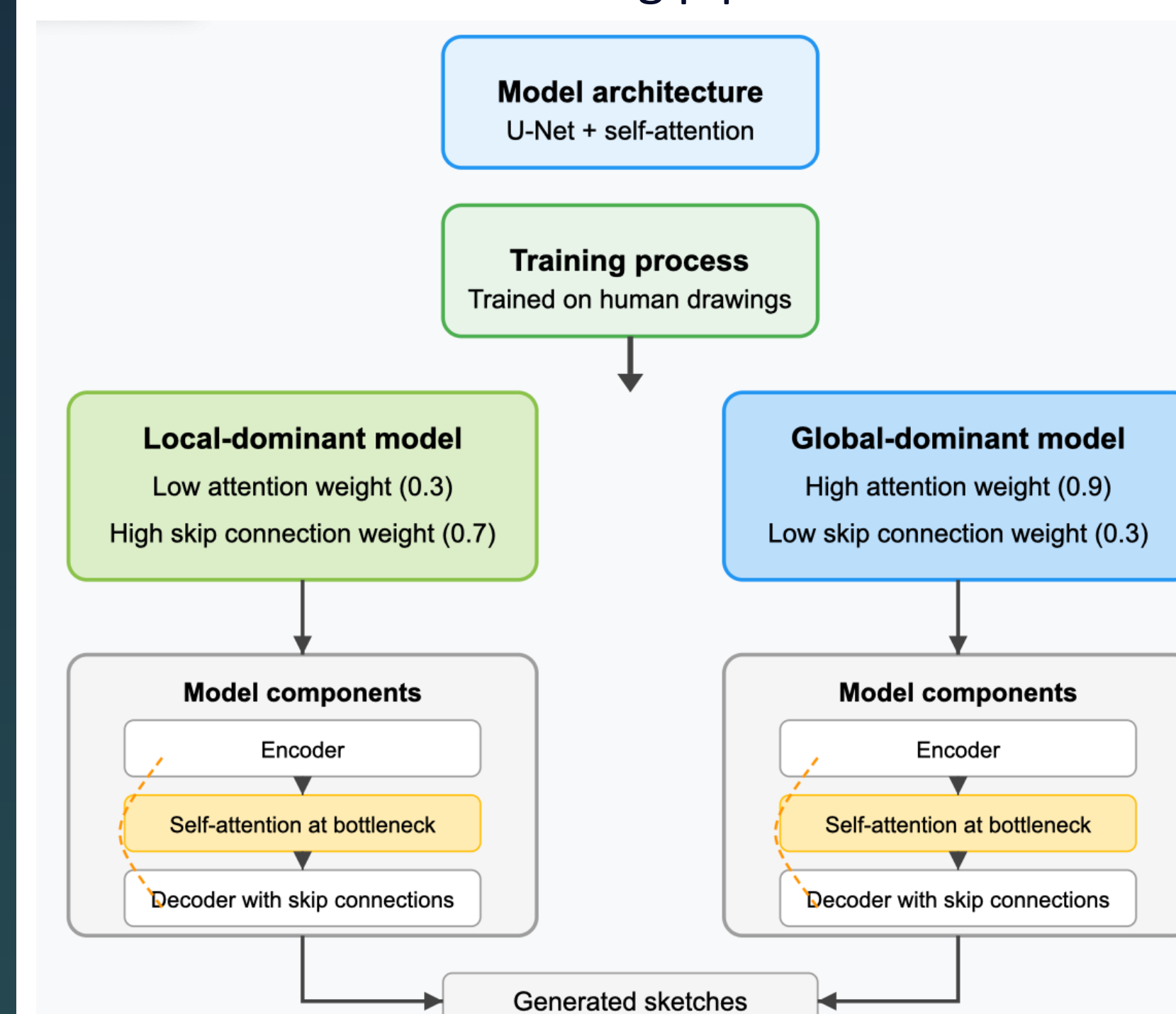
Generated examples



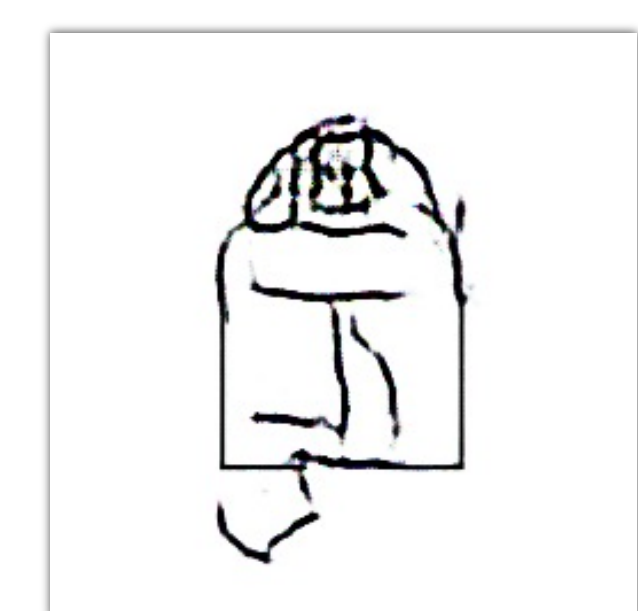
Custom Neural Network

Controlled Gestalt attention

Processing pipeline

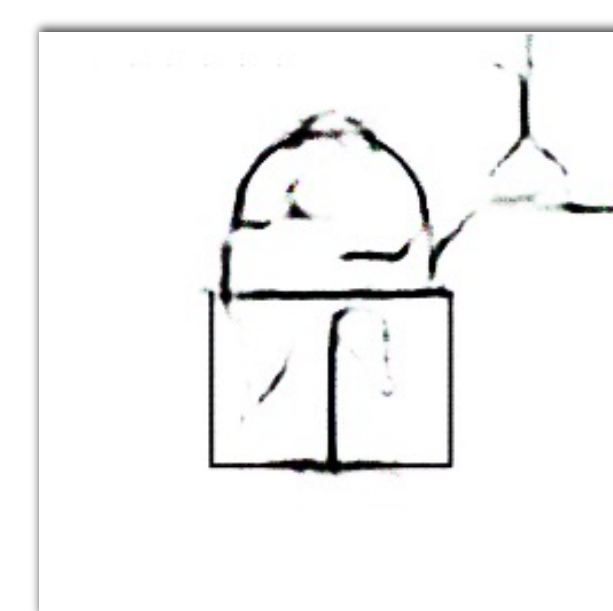


Example

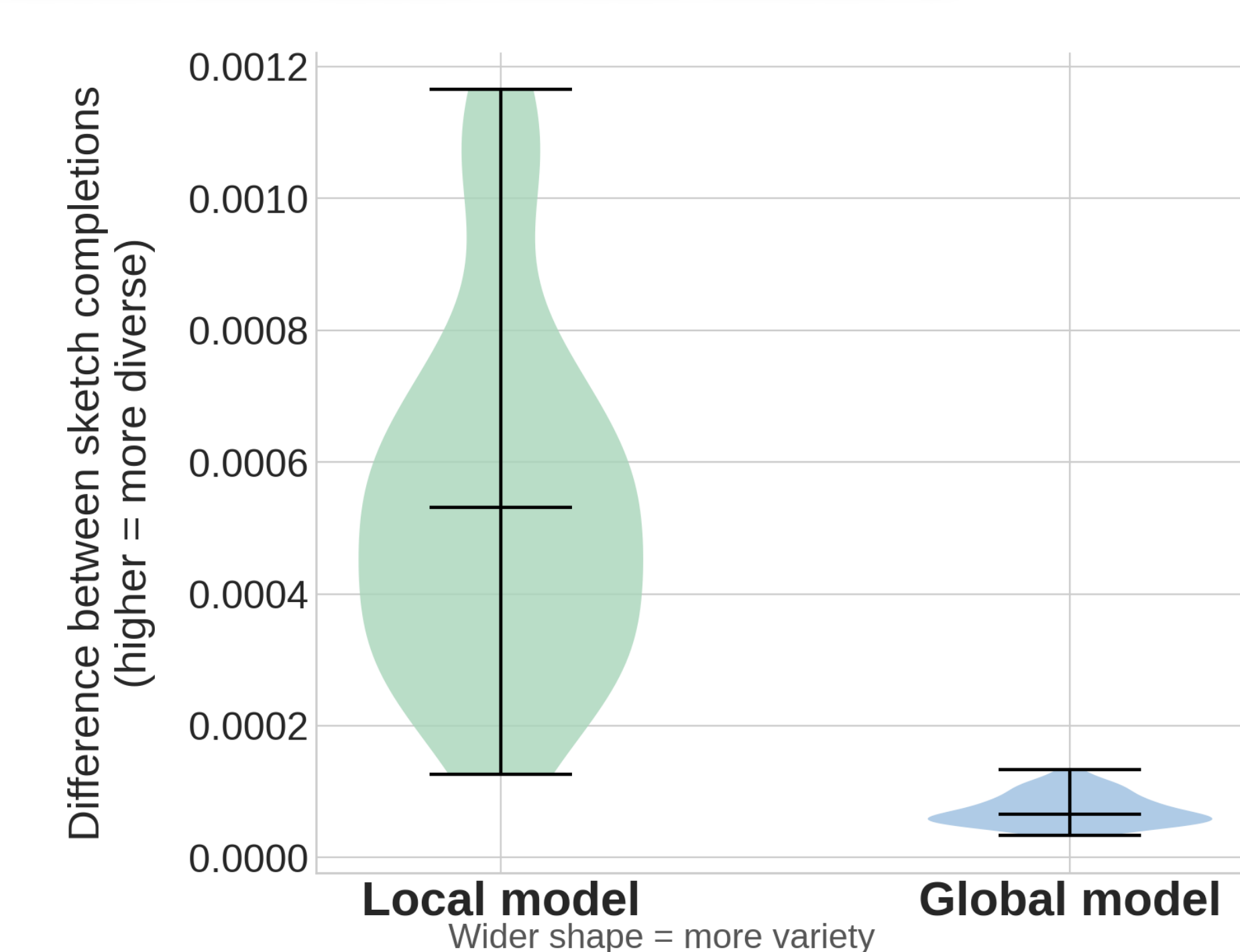
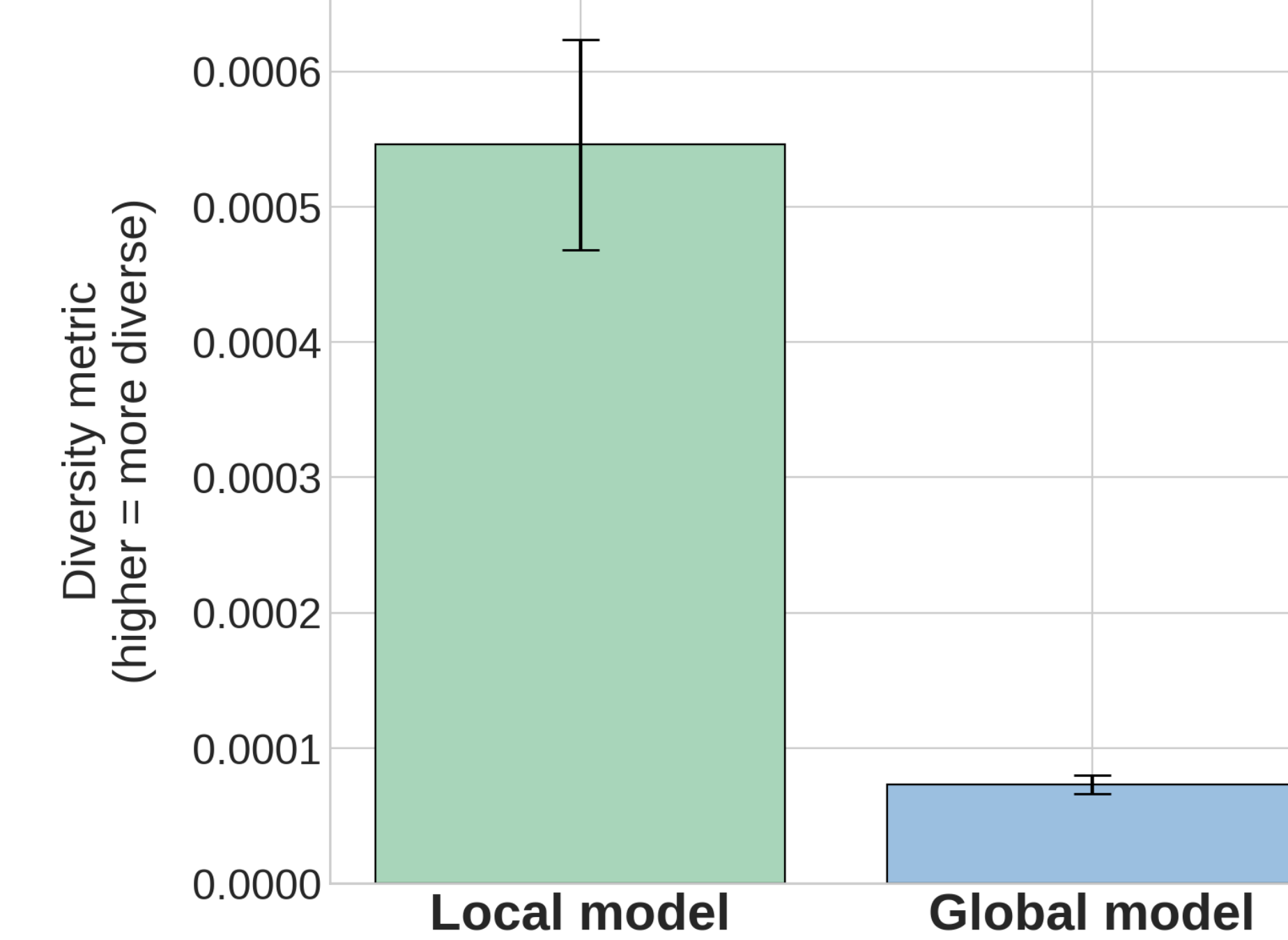


How do generated sketches vary by local-attention and global-attention models?

Example

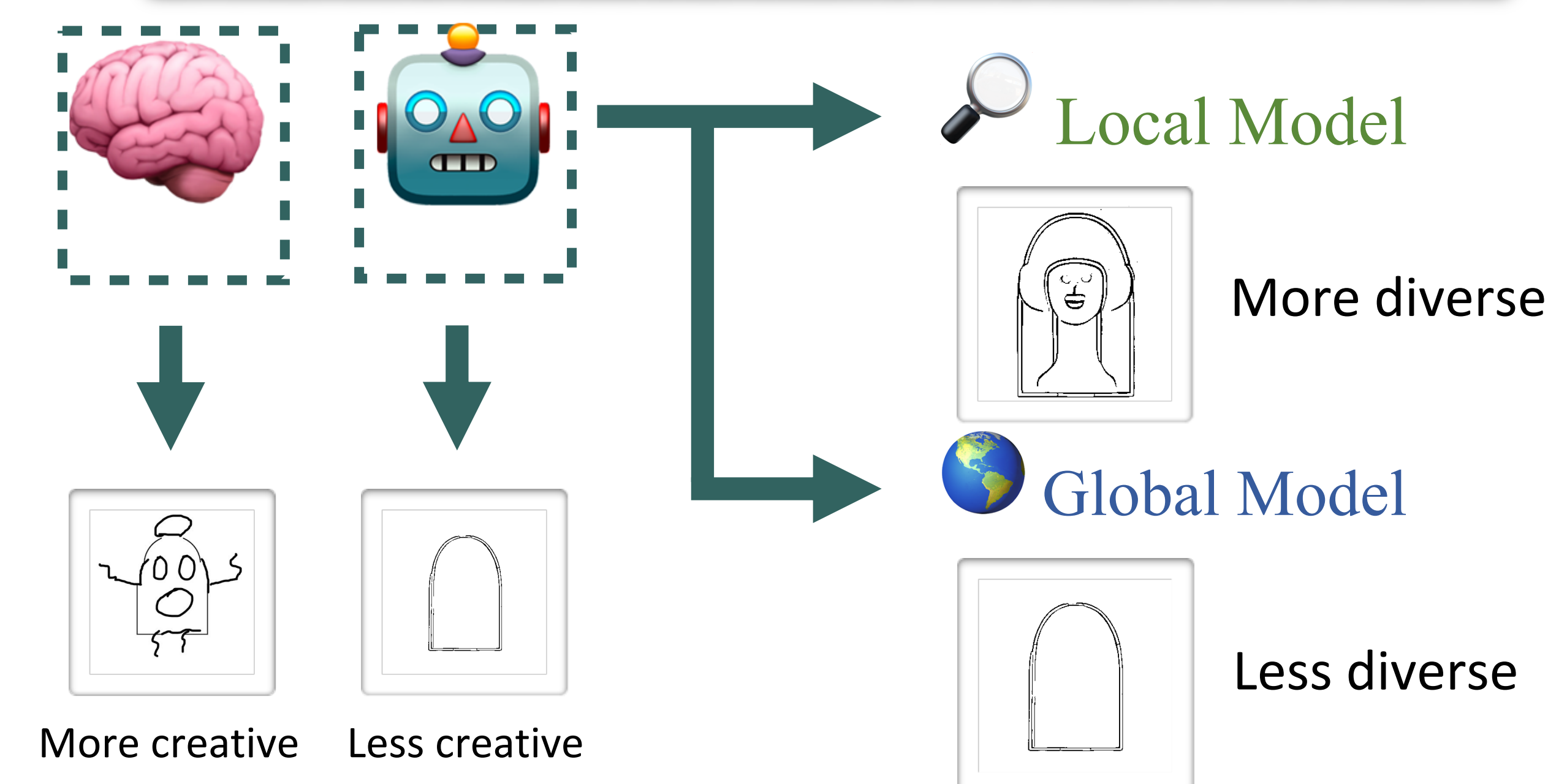


Local vs. Global self-attention



Sketches generated by the local-attention model showed higher diversity (variability) and pixel-wise distances (L1)

Conclusions



References

- Barbot (2018). *Front Psychol.*
- Patterson et al. (2023). *Behavioral Res Methods.*
- Podell et al. (2023). *arXiv.*
- Li et al. (2022). *arXiv.*

Contact

yaxin.liu@georgetown.edu

@yliu668

