Rotating Snakes Illusion Reveals Limitations of Visual-Motion Models in Explaining Human Vision

Isabella E. Rosario* (ier2108@columbia.edu)
Department of Neuroscience;
The Mortimer B. Zuckerman Mind Brain Behavior Institute;
Columbia University, New York, NY, USA

Fan L. Cheng* (fc2803@columbia.edu)
Department of Neuroscience;
The Mortimer B. Zuckerman Mind Brain Behavior Institute;
Columbia University, New York, NY, USA

Zitang Sun (sun.zitang.c09@kyoto-u.jp)
Cognitive Informatics Lab;
Department of Intelligence Science and Technology;
Kyoto University, Sakyo, Kyoto, JP

Shin'ya Nishida (nishida.shinya.2x@kyoto.u.ac.jp)
Cognitive Informatics Lab;
Department of Intelligence Science and Technology;
Kyoto University, Sakyo, Kyoto, JP
Human Information Science Laboratory;
NTT Communication Science Laboratories;
Nippon Telegraph and Telephone Corporation, 3-1, Morinosato Wakamiya Atsugi-shi, Kanagawa, JP

Nikolaus Kriegeskorte (nk2765@columbia.edu)
Departments of Psychology and Neuroscience;
The Mortimer B. Zuckerman Mind Brain Behavior Institute;
Columbia University, New York, NY, USA

*These authors contributed equally to this work.

Abstract

Deep Neural Network (DNN) models provide a computational framework that enables rigorous understanding of vision. Recent DNN-based motion models have successfully replicated illusions like reverse-phi and barber pole, suggesting possible shared computational principles with human motion processing. However, findings have been mixed on whether DNN models can replicate the "Rotating Snakes" illusion—static patterns that induce motion perception in humans. We tested representative optical flow estimation models on both grayscale and color versions of Rotating Snakes, including those featuring recurrent architectures and different training approaches. None of the models predicted optical flows matching the continuous rotational motion humans perceive, either when presented with consecutive static images or under simulation conditions believed to trigger the illusion, such as saccadic eye movements and stimulus onset. Only the motion energy sensor and self-attention based Dual model estimated partial rotation in expected regions, matching or opposing predicted directions—an effect absent in controls. Our results highlight the gap between current DNN-based motion models and human vision. Future models tested in experimental loops should incorporate mechanisms accounting for possible explanations of the Rotating Snakes illusion, such as pupil dilation, eye movements, and contrast-dependent processing latency, as well as color- and contrast-sensitive adaptation functionality.

Keywords: Motion perception; deep neural networks; optical flow; visual illusions

Introduction

The "Rotating Snakes" illusion (Kitaoka & Ashida, 2003) consists of static patterns with concentric micropatterns of asymmetric luminance that create illusory rotational motion, particularly in peripheral vision (Murakami et al., 2006). A unified account for this phenomenon remains unestablished, with several competing explanations proposed in the literature: contrast-dependent differences in neural response timing in V1 and MT interpreted as motion by first-order detectors (Conway et al., 2005; Fermüller et al., 2010; Bach & Atala-Gérard, 2020); erroneous estimation of local motion signals and fixational eye movement effects (Murakami et al., 2006; Fermüller et al., 2010); non-linear saturation of motion detectors and saccades upon pattern onset (Backus & Oruç, 2005); reflexive pupil dilation responding to transitory luminance changes (Mather & Cavanagh, 2025).

DNNs have emerged as a powerful tool for modeling vision (Kriegeskorte, 2015). Researchers found that PredNet could reproduce the Rotating Snakes illusion, suggesting predictive coding as a putative mechanism (Watanabe et al., 2018). Reproduction of this effect lacks robustness across network instances and consistent alignment with psychophysical and electrophysiological findings (Kirubeswaran & Storrs, 2023). Recent DNN-based motion models have successfully replicated several classic motion illusions, suggesting potential shared computational principles with human motion processing (Solari et al., 2015; Z. Sun et al., 2023, 2025). Here, we investigate whether representative DNN-based motion models are capable of reproducing the Rotating Snakes illusion.

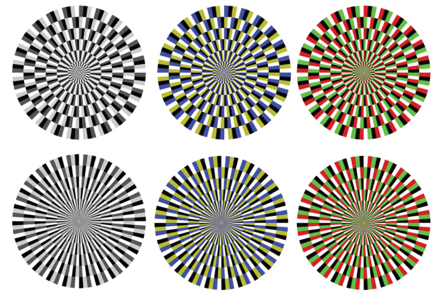

Figure 1: Images used for model inputs. Rotating Snakes illusion stimuli (top row) induce perception of counter-clockwise motion in human observers, while control stimuli (bottom row) with modified color/luminance ordering of micropatterns induce no motion perception. From left to right: Grayscale, blue-yellow, and red-green versions.

Table 1: Summary of motion estimation models. M-S: multi-scale methods; R-D: recurrent decoding methods; SV: supervised learning; USV: unsupervised learning; Bio.: Biocomputing methods.

Model	Param.	M-S	R-D	SV	USV	Bio.
FFV1MT	N.A	•				•
PWC-Net	8.75 M	•		•		
RAFT	5.26 M	•	•	•		
LiteFlowNet2	6.42 M	•		•		
DorsalNet	55,296			•		•
ME-Attention	14.7 M	•	•	•		•
Dual	25.6 M	•	•	•		•

Results

We generated modified versions of the Rotating Snakes illusion (Kitaoka & Ashida, 2003) consisting of concentric repeated cycles of four stepwise luminance levels, with RGB values matched to that of Uesaki et al. (2024), along with corresponding controls (Fig. 1). We tested models' motion estimation on illusion and control stimuli using three presentation conditions: static (consecutive identical images), microsaccade (rightward-downward stimulus translation), and stimulus-onset (blank to stimulus). The microsaccade condition simulated the experiment in Otero-Millan et al. (2012) with approximately 0.2° visual angle saccade while viewing an 8°-diameter disk. To verify models could detect real motion, we generated stimuli by physically rotating control images. We evaluated alignment between model predictions and human

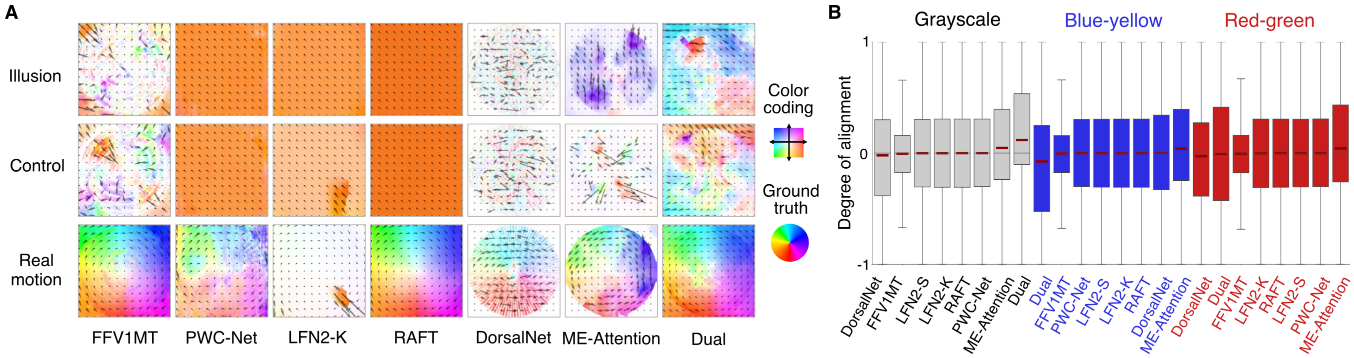

Figure 2: Test results under microsaccade simulation. (**A**) Optical flow estimates from models for the grayscale illusion stimuli. Color saturation indicates flow speed. Inset shows ground truth flow from physically counterclockwise rotating stimuli. (**B**) Degree of alignment with human perception across visual space. Box plots show 25th-75th percentile range with means (dark red lines). Values closer to 1 indicate counterclockwise predictions, while values closer to -1 represent clockwise predictions.

perception using $1 - e/90$, where e is the angular error between predicted and ground truth optical flow vectors that mimic human judgment (counterclockwise rotation).

Models showed correct predictions of optical flow for the real motion condition, except for **LiteFlowNet2** (Hui et al., 2020) variants which predicted little to no optical flow except for a region of rightward motion (Fig. 2A). **FFV1MT** (Solari et al., 2015), a motion energy model inspired by the V1-MT processing pathway, predicted optical flow in the direction of the stimulus shift. Despite being previously shown to capture reverse-phi motion phenomena, **ME-Attention** (Z. Sun et al., 2023) generated vertical flow estimates and did not produce responses in the direction of the stimulus shift. This model is also built on trainable spatiotemporal filters and implements a self-attention motion integration mechanism. Spatial pyramid based models (**PWC-Net** (D. Sun et al., 2018) and **LiteFlowNet2**) and a representative state-of-the-art DNN-based optical estimation model (**RAFT** (Teed & Deng, 2020)) exhibit biases toward structural motion. Training datasets and fine-tuning methods also did not impact the generated optical flow predictions of **LiteFlowNet2**, as versions trained and fine-tuned either on KITTI (Geiger et al., 2013) or Sintel (Mayer et al., 2016) datasets performed similarly. In contrast, **DorsalNet** (Mineault et al., 2021) a 3DResnet trained with the objective of predicting dorsal stream neural recordings and containing units tuned for complex optical flow patterns such as rotation and expansion, generated opposing flow predictions and greater flow magnitudes at the edges of the stimulus micropatterns. **Dual** (Z. Sun et al., 2025), a V1-MT pathway inspired model capable of first- and second-order motion perception with self-attention, predicted regions of counterclockwise rotation of the grayscale illusion as well as a region of motion estimation in the direction of the stimulus shift.

For static and stimulus-onset presentation conditions, although optical flow maps were qualitatively different from the microsaccade condition, all models failed to reproduce continuous motion that aligns with human perception, with only **Dual** being capable of predicting partially counterclockwise motion.

Quantification analysis shows only **Dual** exhibiting closer alignment with counterclockwise motion in the grayscale illusion condition (Fig. 2B). Although rotation motions were also predicted by **Dual** for color versions, the direction was opposite (clockwise) for the blue-yellow version, while both directions existed in the optical flow map for the red-green version.

These findings indicate that current DNN models struggle with ambiguous luminance-defined local motion signals and lack the ability to globally integrate them, limiting their compatibility with stimuli used in human vision research.

Conclusions

We show that visual-motion models fail to predict optical flow in the direction of illusory motion perceived by humans. Non-biocomputing methods track only microsaccade direction due to their focus on absolute image correspondence between frames and vulnerability to changes in pixel intensity. In comparison, biocomputing methods captured motion flow beyond the microsaccade without showing rotational patterns, except for **Dual**, the only tested model capable of second-order motion perception. DNN performance depends heavily on architecture and training data, with models often overfitting to natural color datasets and generalizing poorly to artificial and grayscale stimuli. Future model development should account for eye movements and implement mechanisms for naturalistic brain-like spatiotemporal dynamics in response to luminance changes responsible for first-order motion perception.

References

Bach, M., & Atala-Gérard, L. (2020). The rotating snakes illusion is a straightforward consequence of nonlinearity in arrays of standard motion detectors. , *11*(5), 2041669520958025.

Backus, B. T., & Oruç, I. (2005). Illusory motion from change over time in the response to contrast and luminance. *Journal of Vision*, *5*(11), 10.

Conway, B. R., Kitaoka, A., Yazdanbakhsh, A., Pack, C. C., & Livingstone, M. S. (2005). Neural basis for a powerful static motion illusion. *Journal of Neuroscience*, *25*(23), 5651–5656.

Fermüller, C., Ji, H., & Kitaoka, A. (2010). Illusory motion due to causal time filtering. *Vision Research*, *50*(3), 315–329.

Geiger, A., Lenz, P., Stiller, C., & Urtasun, R. (2013). Vision meets robotics: The KITTI dataset. *The International Journal of Robotics Research*, *32*(11), 1231–1237.

Hui, T.-W., Tang, X., & Loy, C. C. (2020). *A Lightweight Optical Flow CNN - Revisiting Data Fidelity and Regularization.* arXiv. (arXiv:1903.07414 [cs]) doi: 10.48550/arXiv.1903.07414

Kirubeswaran, O. R., & Storrs, K. R. (2023). Inconsistent illusory motion in predictive coding deep neural networks. *Vision Research*, *206*, 108195.

Kitaoka, A., & Ashida, H. (2003). Phenomenal characteristics of the peripheral drift illusion. *Vision*, *15*(4), 261–262.

Kriegeskorte, N. (2015). Deep Neural Networks: A New Framework for Modeling Biological Vision and Brain Information Processing. *Annual Review of Vision Science*, *1*(1), 417–446.

Mather, G., & Cavanagh, P. (2025). Pupil dilation underlies the peripheral drift illusion. *Journal of Vision*, *25*(2), 13.

Mayer, N., Ilg, E., Hausser, P., Fischer, P., Cremers, D., Dosovitskiy, A., & Brox, T. (2016). A Large Dataset to Train Convolutional Networks for Disparity, Optical Flow, and Scene Flow Estimation. In (pp. 4040–4048).

Mineault, P. J., Bakhtiari, S., Richards, B. A., & Pack, C. C. (2021). Your head is there to move you around: Goal-driven models of the primate dorsal pathway. *bioRxiv*. doi: 10.1101/2021.07.09.451701

Murakami, I., Kitaoka, A., & Ashida, H. (2006). A positive correlation between fixation instability and the strength of illusory motion in a static display. *Vision Research*, *46*(15), 2421–2431.

Otero-Millan, J., Macknik, S. L., & Martinez-Conde, S. (2012). Microsaccades and blinks trigger illusory rotation in the "rotating snakes" illusion. *Journal of Neuroscience*, *32*(17), 6043–6051.

Solari, F., Chessa, M., Medathati, N. K., & Kornprobst, P. (2015). What can we expect from a V1-MT feedforward architecture for optical flow estimation? *Signal Processing: Image Communication*, *39*, 342–354.

Sun, D., Yang, X., Liu, M.-Y., & Kautz, J. (2018). *PWC-Net: CNNs for Optical Flow Using Pyramid, Warping, and Cost Volume.* arXiv. (arXiv:1709.02371 [cs]) doi: 10.48550/arXiv.1709.02371

Sun, Z., Chen, Y.-J., Yang, Y.-H., Li, Y., & Nishida, S. (2025). Machine learning modeling for multi-order human visual motion processing. *arXiv preprint arXiv:2501.12810*.

Sun, Z., Chen, Y.-J., Yang, Y.-H., & Nishida, S. (2023). Modeling human visual motion processing with trainable motion energy sensing and a self-attention network. *Advances in Neural Information Processing Systems*, *36*, 24335–24348.

Teed, Z., & Deng, J. (2020). RAFT: Recurrent All-Pairs Field Transforms for Optical Flow. In *Computer Vision – ECCV 2020* (pp. 402–419). doi: 10.1007/978-3-030-58536-5_24

Uesaki, M., Biswas, A., Ashida, H., & Maus, G. (2024). Blue-yellow combination enhances perceived motion in rotating snakes illusion. , *15*(2), 20416695241242346.

Watanabe, E., Kitaoka, A., Sakamoto, K., Yasugi, M., & Tanaka, K. (2018). Illusory Motion Reproduced by Deep Neural Networks Trained for Prediction. *Frontiers in Psychology*, *9*, 345.