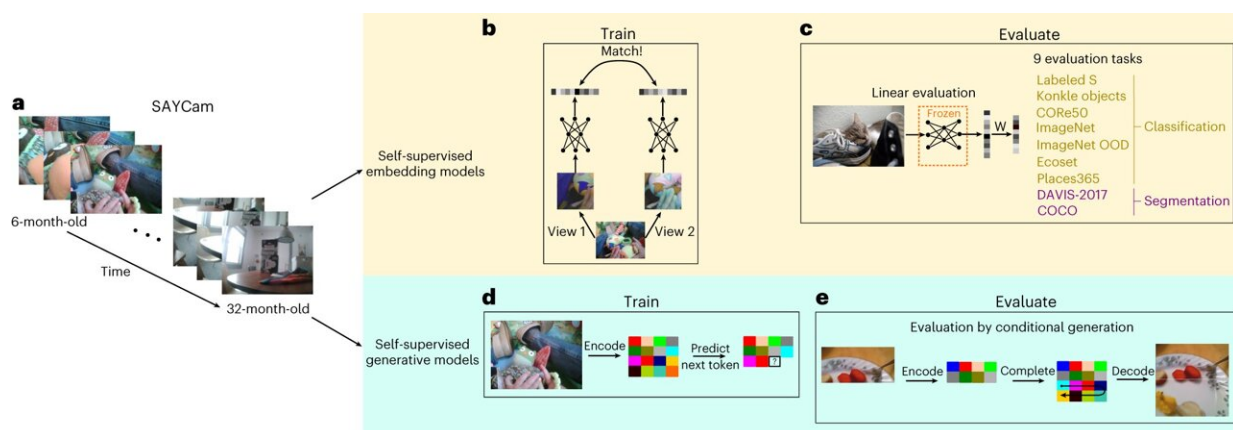


Training artificial neural networks to process images from a child's perspective

March 20 2024, by Ingrid Fadelli



Schematic overview of the experiments. Credit: *Nature Machine Intelligence* (2024). DOI: 10.1038/s42256-024-00802-0

Psychology studies have demonstrated that by the age of 4–5, young children have developed intricate visual models of the world around them. These internal visual models allow them to outperform advanced computer vision techniques on various object recognition tasks.

Researchers at New York University recently set out to explore the possibility of training [artificial neural networks](#) on these models without domain-specific inductive biases. Their [paper](#), published in *Nature Machine Intelligence*, ultimately addresses one of the oldest philosophical questions, namely the "nature vs. nurture" dilemma.

The nature vs. nurture dilemma disputes whether humans possess innate inductive biases influencing how they perceive objects, people and the world around them overall, or whether they are initially a "blank slate," developing biases as a result of their experiences. Some of the hypothesized innate biases are related to the ability to categorize and label objects.

The team at New York University set out to investigate this dilemma from a modern standpoint. To do this, they trained state-of-the-art self-supervised deep neural networks on a large dataset containing videos taken from young children's perspective using headcams (cameras attached to a hat or helmet).

"Young children develop sophisticated internal models of the world based on their visual experience," A. Emin Orhan and Brenden M. Lake wrote in their paper. "Can such models be learned from a child's visual experience without strong inductive biases? To investigate this, we train state-of-the-art neural networks on a realistic proxy of a child's [visual experience](#) without any explicit supervision or domain-specific inductive biases."

Orhan and Lake trained two types of deep learning techniques, namely embedding and generative models, on approximately 200 hours of headcam video footage collected from a single child over a two-year period. After pre-training more than 70 of these models, they tested their performance on a series of computer vision and object recognition tasks, comparing it with other state-of-the-art computer vision models.

"On average, the best embedding models perform at a respectable 70% of a high-performance ImageNet-trained [model](#), despite substantial differences in training data," Orhan and Lake wrote. "They also learn broad semantic categories and object localization capabilities without explicit supervision, but they are less object-centric than models trained

on all of ImageNet.

"Generative models trained with the same data successfully extrapolate simple properties of partially masked objects, like their rough outline, texture, color or orientation, but struggle with finer object details."

To validate their findings, the researchers carried out further experiments involving two other [young children](#). Their results were consistent with those gathered during their first experiment, suggesting that higher-level [visual representations](#) can be learned from a child's unique visual experiences without integrating strong inductive biases.

The findings of this recent work by Orhan and Lake could serve as an inspiration for psychologists and neuroscientists, informing further studies exploring the nature vs. nurture dilemma using computational tools. Overall, the team suggests that object categorization biases depend on the unique characteristics of the human visual system, which result in different images from those typically used to train deep learning models.

"We hope that our work will inspire new collaborations between [machine learning](#) and developmental psychology, as the impact of modern deep learning on [developmental psychology](#) has been relatively limited thus far," Orhan and Lake conclude in their paper.

"Future algorithmic advances, combined with richer and larger developmental datasets, can be evaluated through the same approach, further enriching our understanding of what can be learned from a child's experience with minimal inductive biases."

More information: A. Emin Orhan et al, Learning high-level visual representations from a child's perspective without strong inductive

biases, *Nature Machine Intelligence* (2024). [DOI: 10.1038/s42256-024-00802-0](https://doi.org/10.1038/s42256-024-00802-0)

© 2024 Science X Network

Citation: Training artificial neural networks to process images from a child's perspective (2024, March 20) retrieved 27 April 2024 from <https://techxplore.com/news/2024-03-artificial-neural-networks-images-child.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.