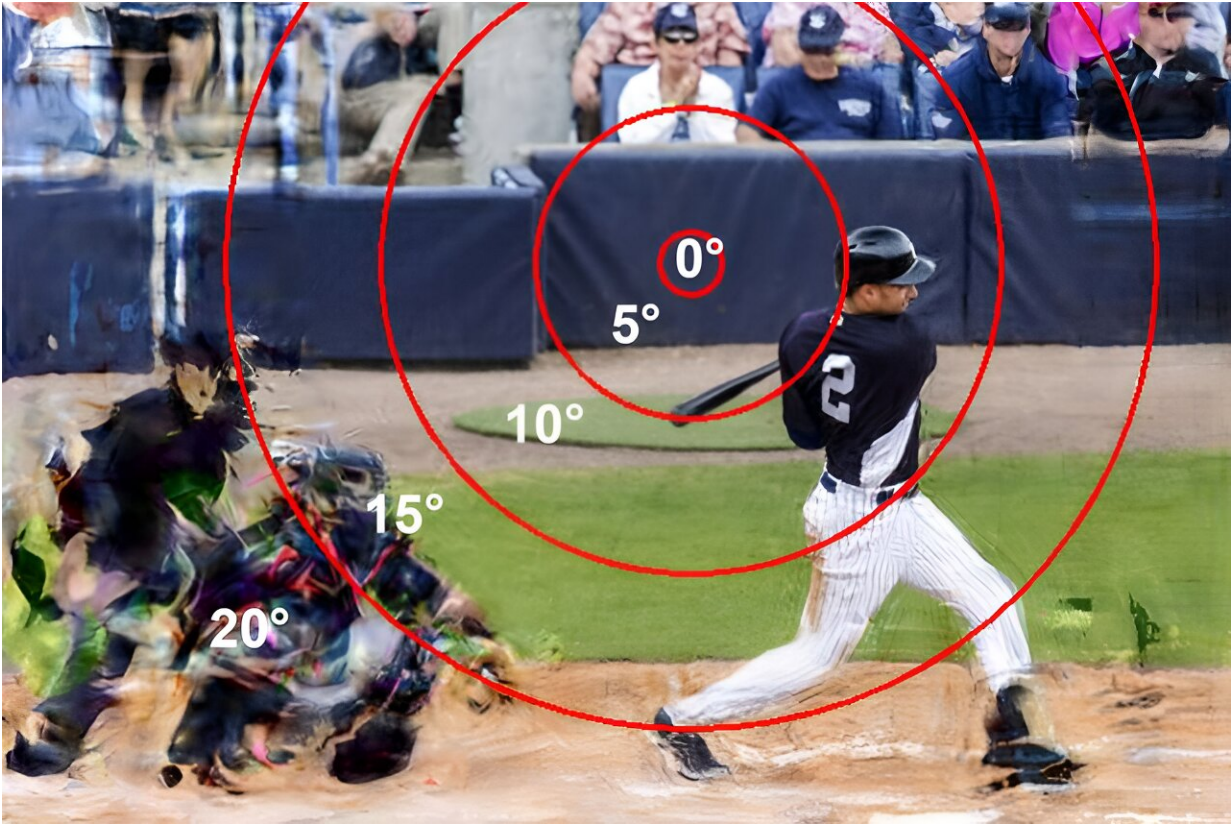


# Researchers enhance peripheral vision in AI models

March 8 2024, by Adam Zewe



The researchers used a modified technique to transform images so they appear more textual in certain areas to represent the loss of detail that occurs when a human looks further into the periphery. They used these transformed images to build their dataset. This image shows the transformation. Credit: Massachusetts Institute of Technology

Peripheral vision enables humans to see shapes that aren't directly in our line of sight, albeit with less detail. This ability expands our field of vision and can be helpful in many situations, such as detecting a vehicle approaching our car from the side.

Unlike humans, AI does not have peripheral vision. Equipping computer vision models with this ability could help them detect approaching hazards more effectively or predict whether a human driver would notice an oncoming object.

Taking a step in this direction, MIT researchers developed an image dataset that allows them to simulate peripheral vision in machine learning models. They found that training models with this dataset improved the models' ability to detect objects in the visual periphery, although the models still performed worse than humans.

Their [results](#) also revealed that, unlike with humans, neither the size of objects nor the amount of visual clutter in a scene had a strong impact on the AI's performance.

"There is something fundamental going on here. We tested so many different models, and even when we train them, they get a little bit better but they are not quite like humans. So, the question is: What is missing in these models?" says Vasha DuTell, a postdoc and co-author of a paper detailing this study.

Answering that question may help researchers build machine learning models that can see the world more like humans do. In addition to improving driver safety, such models could be used to develop displays that are easier for people to view.

Plus, a deeper understanding of peripheral vision in AI models could help researchers better predict human behavior, adds lead author Anne

Harrington MEng '23.

"Modeling peripheral vision, if we can really capture the essence of what is represented in the periphery, can help us understand the features in a visual scene that make our eyes move to collect more information," she explains.

Their co-authors include Mark Hamilton, an [electrical engineering](#) and computer science graduate student; Ayush Tewari, a postdoc; Simon Stent, research manager at the Toyota Research Institute; and senior authors William T. Freeman, the Thomas and Gerd Perkins Professor of Electrical Engineering and Computer Science and a member of the Computer Science and Artificial Intelligence Laboratory (CSAIL); and Ruth Rosenholtz, principal research scientist in the Department of Brain and Cognitive Sciences and a member of CSAIL. The research will be presented at the International Conference on Learning Representations ([ICLR 2024](#)).

"Any time you have a human interacting with a machine—a car, a robot, a user interface—it is hugely important to understand what the person can see. Peripheral vision plays a critical role in that understanding," Rosenholtz says.

## **Simulating peripheral vision**

Extend your arm in front of you and put your thumb up—the small area around your thumbnail is seen by your fovea, the small depression in the middle of your retina that provides the sharpest vision. Everything else you can see is in your visual periphery. Your visual cortex represents a scene with less detail and reliability as it moves farther from that sharp point of focus.

Many existing approaches to model peripheral vision in AI represent this

deteriorating detail by blurring the edges of images, but the information loss that occurs in the optic nerve and visual cortex is far more complex.

For a more accurate approach, the MIT researchers started with a technique used to model peripheral vision in humans. Known as the texture tiling model, this method transforms images to represent a human's visual information loss.

They modified this model so it could transform images similarly, but in a more flexible way that doesn't require knowing in advance where the person or AI will point their eyes.

"That let us faithfully model [peripheral vision](#) the same way it is being done in human vision research," says Harrington.

The researchers used this modified technique to generate a huge dataset of transformed images that appear more textural in certain areas, to represent the loss of detail that occurs when a human looks further into the periphery.

Then they used the dataset to train several [computer vision models](#) and compared their performance with that of humans on an object detection task.

"We had to be very clever in how we set up the experiment so we could also test it in the machine learning models. We didn't want to have to retrain the models on a toy task that they weren't meant to be doing," she says.

## **Peculiar performance**

Humans and models were shown pairs of transformed images which were identical, except that one image had a target object located in the

periphery. Then, each participant was asked to pick the image with the target object.

"One thing that really surprised us was how good people were at detecting objects in their periphery. We went through at least 10 different sets of images that were just too easy. We kept needing to use smaller and smaller objects," Harrington adds.

The researchers found that training models from scratch with their dataset led to the greatest performance boosts, improving their ability to detect and recognize objects. Fine-tuning a model with their dataset, a process that involves tweaking a pretrained model so it can perform a new task, resulted in smaller performance gains.

But in every case, the machines weren't as good as humans, and they were especially bad at detecting objects in the far periphery. Their performance also didn't follow the same patterns as humans.

"That might suggest that the models aren't using context in the same way as humans are to do these detection tasks. The strategy of the models might be different," Harrington says.

The researchers plan to continue exploring these differences, with a goal of finding a model that can predict human performance in the visual periphery. This could enable AI systems that alert drivers to hazards they might not see, for instance. They also hope to inspire other researchers to conduct additional computer vision studies with their publicly available dataset.

"This work is important because it contributes to our understanding that human vision in the periphery should not be considered just impoverished vision due to limits in the number of photoreceptors we have, but rather, a representation that is optimized for us to perform

tasks of real-world consequence," says Justin Gardner, an associate professor in the Department of Psychology at Stanford University who was not involved with this work.

"Moreover, the work shows that neural network models, despite their advancement in recent years, are unable to match human performance in this regard, which should lead to more AI research to learn from the neuroscience of human vision. This future research will be aided significantly by the database of images provided by the authors to mimic peripheral human vision."

**More information:** COCO-Periph: Bridging the Gap Between Human and Machine Perception in the Periphery.

[openreview.net/pdf?id=MiRPBbQNHv](https://openreview.net/pdf?id=MiRPBbQNHv)

*This story is republished courtesy of MIT News ([web.mit.edu/newsoffice/](http://web.mit.edu/newsoffice/)), a popular site that covers news about MIT research, innovation and teaching.*

Provided by Massachusetts Institute of Technology

Citation: Researchers enhance peripheral vision in AI models (2024, March 8) retrieved 27 April 2024 from <https://techxplore.com/news/2024-03-peripheral-vision-ai.html>

<p>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.</p>
--