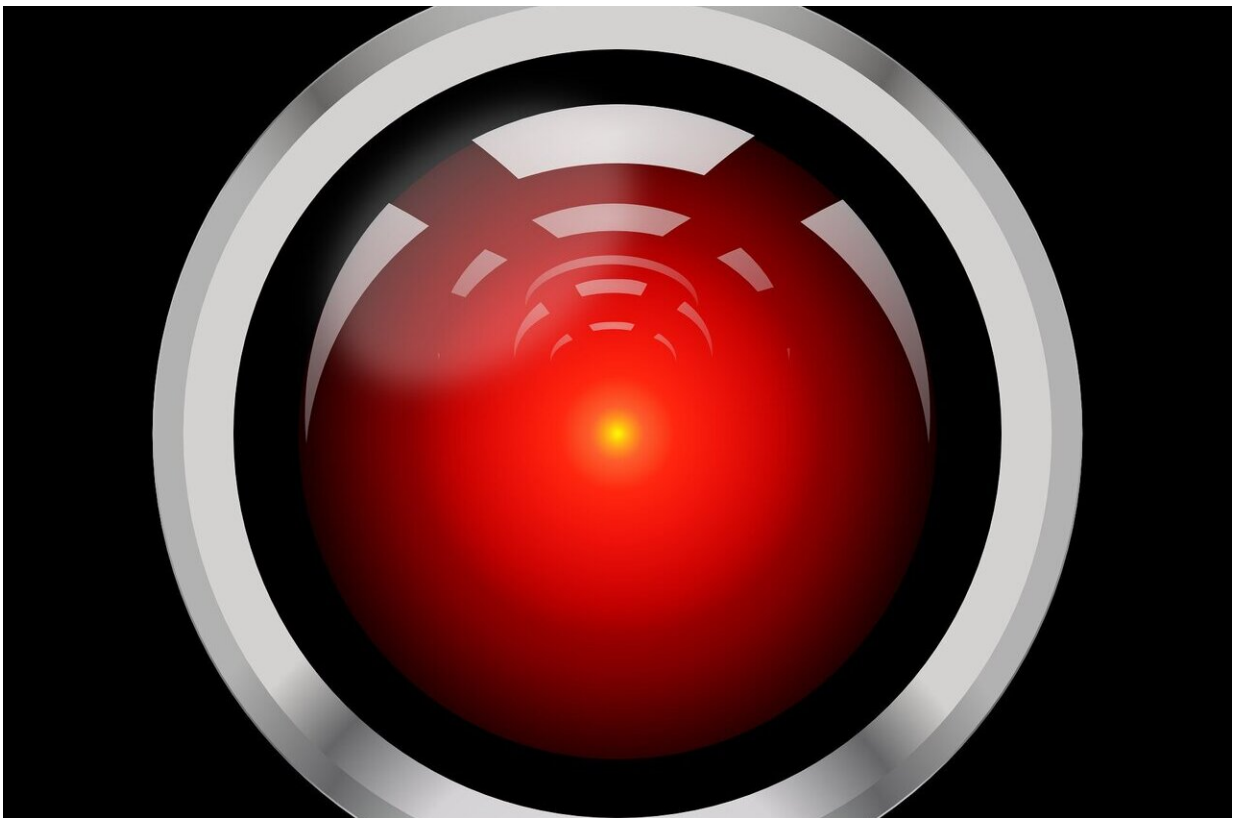


Computer scientist researches interpretable machine learning, develops AI to explain its discoveries

November 4 2020, by Marcus Wolf



Credit: Pixabay/CC0 Public Domain

Artificial intelligence helps scientists make discoveries, but not everyone can understand how it reaches its conclusions. One UMaine computer

scientist is developing deep neural networks that explain their findings in ways users can comprehend, applying his work to biology, medicine and other fields.

Interpretable machine learning, or AI that creates explanations for the findings it reaches, defines the focus of Chaofan Chen's research. The assistant professor of computer science says interpretable machine learning also allows AI to make comparisons among images and predictions from data, and at the same time, elaborate on its reasoning.

Scientists can use interpretable machine learning for a variety of applications, from identifying birds in images for wildlife surveys to analyzing mammograms.

"I want to enhance the transparency for [deep learning](#), and I want a deep neural network to explain why something is the way it thinks it is," Chen says. "What a lot of people have been starting to realize is that a deep neural network is like a black box, and people need to start figuring out ways to open the black box."

Chen began developing interpretable machine learning techniques while studying at Duke University, where he earned his Ph.D in computer science in May.



Credit: University of Maine

Before joining UMaine, Chen and research colleagues at Duke developed machine learning architecture known as a prototypical part network (ProtoPNet) to pinpoint and categorize birds in photos, then explain its findings. The ProtoPNet, which the team completed last year, would explain why the bird it identified was a bird and why it embodies a particular type of bird.

Researchers trained the ProtoPNet to determine what kind of bird is in a photo. The AI, for example, would learn a set of prototypical features that characterize each bird species, and compare different parts of a bird image with these prototypical features from a variety of bird species. For example, the ProtoPNet would compare what it thought was the head of a bird in the image to prototypical bird heads from a variety of bird classes. Using similarities to prototypical features of a [bird species](#), the

ProtoPNet can explain why the image was a particular kind of bird, Chen says.

The team shared its findings in a paper presented during the 33rd Conference on Neural Information Processing Systems last year in Vancouver, Canada.

"It's a very visual way of gaging the whole reasoning process ... that 'this bird is a clay colored sparrow because it contains parts that are prototypical of a clay colored sparrow,'" Chen says. "Bird recognition is a popular benchmark for fine-grained image classification, so I thought that it would be a good showcase for our technique."

The UMaine computer scientist has begun another AI study with colleagues and students from Duke University exploring how they can apply ProtoPNet to review mammograms for signs of breast cancer.

The ProtoPNet, however, struggles to focus on the crucial portions of the mammogram for pinpointing signs of breast cancer as it lacks the training instilled in doctors, Chen says. The team will train the network to evaluate mammograms like a medical professional and learn and identify crucial patterns in the imagery.

Chen's partners for the project, all from Duke University, include Ph.D. students Alina Jade Barnett and Yinhao Ren, undergraduate student Chaofan Tao, professor of computer science Cynthia Rudin, professor and Vice Chair for Research and Radiology Joseph Lo, and postdoctoral radiology researcher Fides Regina Schwartz.

"This has real impact," Chen says. "I certainly love seeing my work make a positive contribution to society."

Chen's research coincides with the UMaine AI initiative, an effort to

transform the state into a world-class hub for [artificial intelligence](#) research and education, and develop AI-based solutions that enhance social and economic wellbeing.

"It's satisfying for me to see not only the ability (for AI) to predict something and predict something well, but to emulate human thinking," he says.

More information: This Looks Like That: Deep Learning for Interpretable Image Recognition: papers.nips.cc/paper/9095-this-looks-like-that-deep-learning-for-interpretable-image-recognition.pdf

Provided by University of Maine

Citation: Computer scientist researches interpretable machine learning, develops AI to explain its discoveries (2020, November 4) retrieved 9 April 2024 from <https://techxplore.com/news/2020-11-scientist-machine-ai-discoveries.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.
