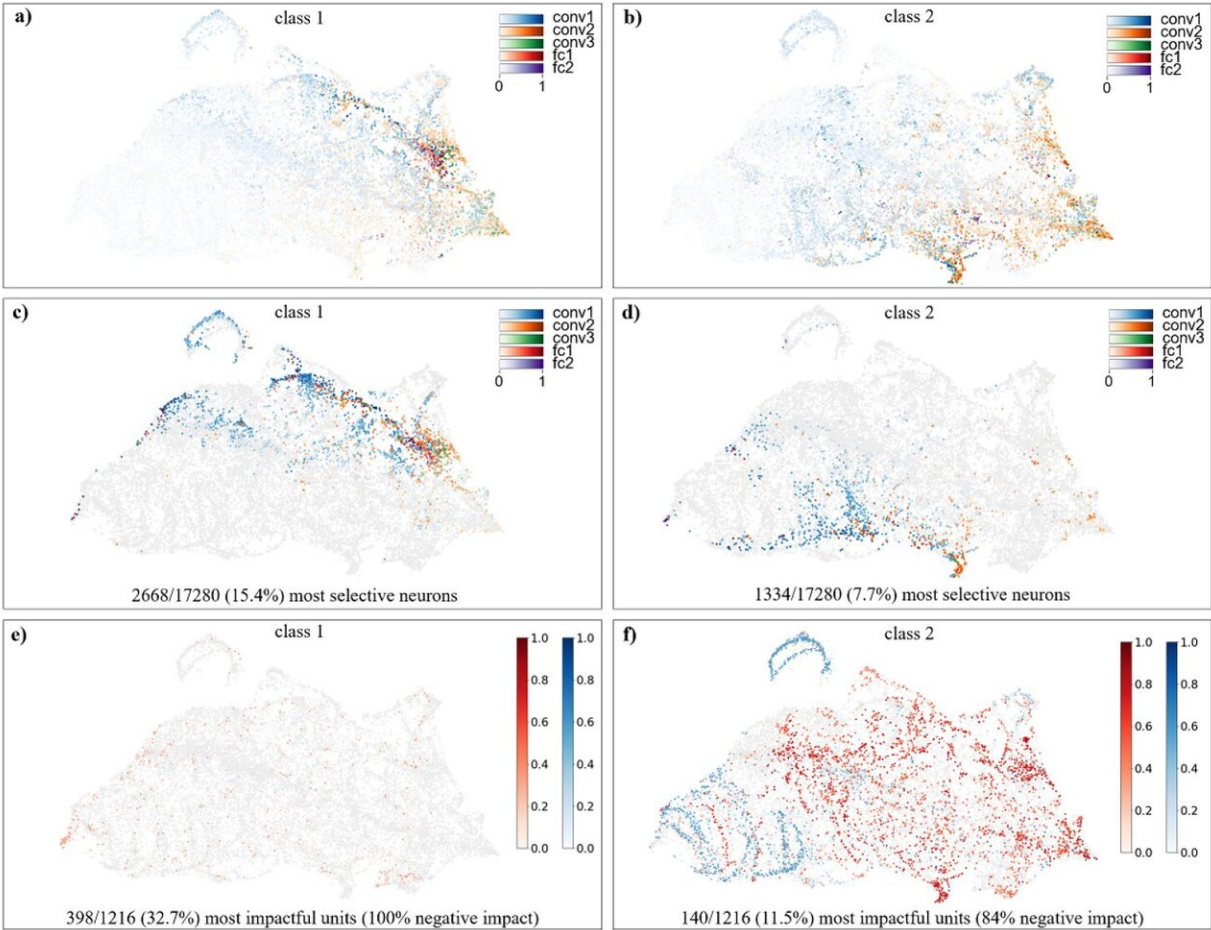


Investigating the role of different neurons in artificial neural networks

April 23 2020, by Ingrid Fadelli



Credit: Meyes et al.

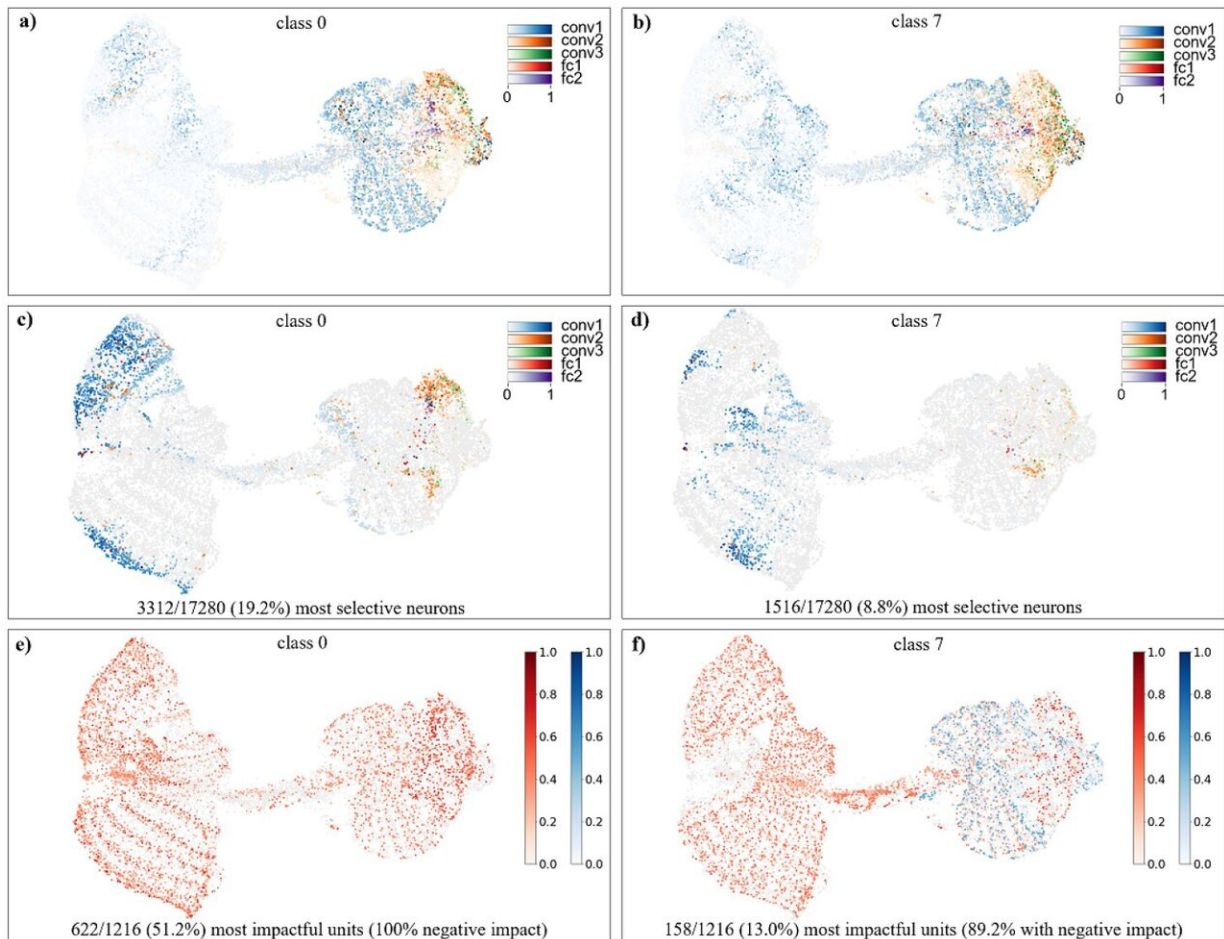
Over the past decade or so, researchers worldwide have been developing

increasingly advanced artificial neural networks (ANNs), computational methods designed to replicate biological mechanisms and functions of the human brain. While some of these networks have achieved remarkable results in a variety of tasks, the decision-making processes underlying their predictions are not always clear.

Researchers at the Institute of Technologies and Management of Digital Transformation and the Institute of Information Management in Mechanical Engineering in Germany have recently carried out a study aimed at better understanding the mechanisms underlying the decision-making processes of ANNs. Their findings, outlined in a paper [pre-published on arXiv](#), shed some light on the roles of both individual [neurons](#) and distinct neuron populations within these networks.

"Our recent study was based on [our previous work](#), in which we characterized the importance of single neurons within [artificial neural networks](#) for their learned tasks," Richard Meyes, one of the researchers who carried out the study, told TechXplore. "In our new study, we extended our investigation to larger groups of neurons with the objective of identifying functional neuron populations in which the contained units act jointly to solve the network's trained task."

To investigate the role of different neurons in ANNs, Meyes and his colleagues drew inspiration from techniques that are commonly employed in neuroscience studies. Their ultimate goal was to characterize the representations that a network acquired over time by observing how it behaved when presented with different stimuli, while also activating or deactivating some of its neurons to better understand their role in processing these stimuli.



Credit: Meyes et al.

"Simply put, we asked whether the network exhibits specific patterns of activity in response to distinct images and investigated which neurons in the network contribute to these activity patterns," Meyes explained. "We also used network ablations to knock out single neurons and groups of neurons from the network to determine what role these neurons play in creating these activity patterns."

Their neuroscience-based approach allowed Meyes and his colleagues to collect a number of interesting observations. Most notably, the

researchers identified neuron populations that had different functions. They also found that some groups of neurons acted together when responding to specific types of stimuli or that deactivating (i.e., ablating) certain neurons had a similar effect on the ANN's overall performance.

"Our most intriguing finding was that there are distinct subsets of neurons within the network that are important for different parts of the processed image data," Meyes said. "This suggests that various aspects of knowledge learned about the training data is locally represented in specific parts of the network."

The recent study carried out by Meyes and his colleagues offers valuable insight that could have implications both for the development of ANNs and for future neuroscience research. In addition to unveiling the role of different neurons in ANNs, their findings imply that specific knowledge stored in a network could be isolated and extracted into a new network, without having to train the new network on the same knowledge.

"In our future research, we plan to investigate whether it is possible to isolate specific parts of the learned knowledge in a network and transfer it into other networks," Meyes said. "This would facilitate a new way of building neural networks in a modular fashion, combining small circuits of different networks that contain specific knowledge into a larger [network](#)."

More information: Under the hood of neural networks: characterizing learned representations by functional neuron populations and network ablations. arXiv:2004.01254 [cs.NE]. arxiv.org/abs/2004.01254

Ablation studies in artificial neural networks. arXiv:1901.08644 [cs.NE]. arxiv.org/abs/1901.08644

© 2020 Science X Network

Citation: Investigating the role of different neurons in artificial neural networks (2020, April 23)
retrieved 25 April 2024 from

<https://techxplore.com/news/2020-04-role-neurons-artificial-neural-networks.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.