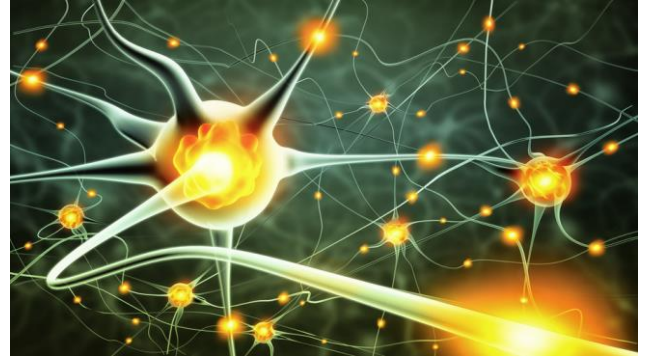


Opening the Blackbox – Neurowissenschaftliche Methoden für das Verständnis von künstlichen neuronalen Netzen

Wissenschaftliche Fragestellung

Die Forschung im Bereich der Neurowissenschaften beschäftigt sich unter anderem mit dem Verständnis über das Gehirn und dessen Funktionsprinzipien um große Fluten an Informationen verschiedenster Art, die der Mensch tagtäglich aufnimmt, zu verarbeiten. Aus verschiedensten logistischen und ethischen Gründen ist der wissenschaftlich saubere Umgang mit dem Gehirn als Untersuchungsgegenstand problematisch, da in der Regel kein großes Ensemble von Gehirnen für eine systematische Untersuchung und statistische Auswertung der Ergebnisse zur Verfügung stehen. Auf der anderen Seite bietet die Forschung über künstliche neuronale Netze (KNN) die Möglichkeit beliebige Untersuchungen und Manipulationen an den Netzen vorzunehmen, ohne irreversiblen Schaden zu verursachen. Die immer weiter ansteigende Größe und Komplexität von KNNs aus dem aktuellen Stand der Technik erfordert jedoch neue Methoden zum tieferen Verständnis über die Art und Weise, wie Wissen in diesen Netzen repräsentiert ist. Die vorliegende Arbeit beschäftigt sich mit der Verwendung von neurowissenschaftlichen Methoden zur Untersuchung von Wissensrepräsentationen in künstlichen neuronalen Netzen.



© adimas - Fotolia.com

Die Möglichkeit beliebige Untersuchungen und Manipulationen an den Netzen vorzunehmen, ohne irreversiblen Schaden zu verursachen. Die immer weiter ansteigende Größe und Komplexität von KNNs aus dem aktuellen Stand der Technik erfordert jedoch neue Methoden zum tieferen Verständnis über die Art und Weise, wie Wissen in diesen Netzen repräsentiert ist. Die vorliegende Arbeit beschäftigt sich mit der Verwendung von neurowissenschaftlichen Methoden zur Untersuchung von Wissensrepräsentationen in künstlichen neuronalen Netzen.

Wissenschaftliche Methodik

Zur wissenschaftlichen Methodik der Arbeit gehört eine einsteigende Literaturrecherche zur historischen Entwicklung der neurowissenschaftlichen Forschung im Bereich der visuellen Informationsverarbeitung. Begleitend sollen jüngste Erkenntnisse aus dem Forschungsbereich Computer Vision recherchiert werden und der neurowissenschaftlichen Forschung gegenübergestellt werden. Als Untersuchungsgegenstand dient ein künstliches neuronales Netz (VGG-16) zur Bildklassifizierung, welches mit neurowissenschaftlich inspirierten Methoden dahingehend untersucht werden soll, an welchen Stellen im Netz Wissen hinterlegt ist, und wie dieses repräsentiert ist. Eine Ablationsuntersuchung des Netzes, bei der systematisch Areale des Netzes entfernt werden, soll Aufschluss über die Struktur des Wissensrepräsentation geben.

Ziel und erwartete Ergebnisse

Das Ziel der Arbeit besteht in der Aufbereitung der neurowissenschaftlichen Forschungshistorie im Bereich der visuellen Informationsverarbeitung und einer Gegenüberstellung zur aktuellen Forschung aus dem Bereich Computer Vision. Darüber hinaus, soll eines der bekanntesten künstlichen neuronalen Netze zur Bildverarbeitung, das VGG-16, mittels einer Ablationsstudie untersucht werden. Die erwarteten Ergebnisse sollen Aufschluss über die Lokalisierung von hinterlegtem Wissen geben und die Art der Repräsentation dieses Wissen charakterisieren. Darüber hinaus werden die Ergebnisse im Rahmen der Abschlussarbeit dokumentiert und in einer Präsentation vorgestellt und verteidigt.

Art der Arbeit

Bachelor-/Master-/Projektarbeit

Ansprechpartner

Richard Meyes, M.Sc.

Tel.: 0241-80-91146

richard.meyes@ima-ifu.rwth-aachen.de