

ZUSAMMENFASSUNG

Moderne Machine-Learning-Techniken sind mit einer Vielzahl von Herausforderungen konfrontiert. Künstliche neuronale Netze sind seit langem dafür bekannt, schwer interpretierbar zu sein und werden stellenweise kritisch als Black-Box-Modelle bezeichnet. Dank des stetigen Anstiegs von Datenmengen und dem einhergehenden Wachstum von neuronalen Netzwerkarchitekturen, hat sich dieses Problem zunehmend intensiviert. Besonders offensichtlich ist diese Problematik im Bereich Computer Vision und bei der Arbeit mit Videodatensätzen. Die hier behandelten Daten sind besonders umfangreich und enthalten große Mengen von Noise und nur ein Bruchteil der Daten enthalten relevante Features.

Dies macht das Training von Machine-Learning-Applikationen im Bereich Computer Vision nicht nur Kosten- und Ressourcen-intensiv, es erschwert auch die Analyse der Modelle. Unerwünschtes Verhalten kann nur schwer untersucht und festgestellt werden. Dies ist problematisch, da neuronale Netze einen ungewünschten Fokus auf unbeabsichtigte Sekundär-Features entwickeln können. Beispielsweise könnten sie fälschlicherweise lernen Schienen, Straßen oder Wolken zu erkennen, obwohl sie Züge, Fahrzeuge und Flugzeuge klassifizieren sollten.

Sogenannte Aufmerksamkeit (engl. attention) stellt einen vielversprechenden Lösungsansatz für diese Probleme dar. Wie künstliche neuronale Netze sind diese Mechanismen auch von der Natur inspiriert und versuchen nachzuahmen, wie Menschen ihre Umgebung wahrnehmen. So wie Menschen möglicherweise in einem Bild ihre Aufmerksamkeit auf ein bestimmtes Objekt lenken, so sollen Aufmerksamkeitsmechanismen ein neuronales Netz dazu bewegen, sich während des Trainings auf möglichst relevante Features zu konzentrieren.

Diese Masterarbeit untersucht, wie verschiedene Implementierungen von visuellen Aufmerksamkeitsmechanismen die Ergebnisse bei der Erkennung von menschlichen Handlungen in Videodaten verbessern können. Weiterhin werden die Vorteile von Aufmerksamkeit bei der Interpretation des Verhaltens von neuronalen Netzen untersucht.