

## Zusammenfassung

Jüngste Fortschritte in den neurobildgebenden Verfahren haben neue Einblicke in die neuronalen Grundlagen psychiatrischer Störungen ermöglicht. Die Anwendung von maschinellem Lernen zur Übersetzung dieser Erkenntnisse in klinisch nützliche Werkzeuge für Diagnose und Behandlung stößt jedoch auf erhebliche Barrieren. Unter anderem begrenzen kleine Stichprobengrößen in neurobildgebenden Studien komplexe Deep-Learning-Modellierungsansätze.

In dieser Arbeit wurde systematisch die Vorhersagegenauigkeit von maschinellem Lernen bei verschiedenen Stichprobengrößen und Bildgebungsmodalitäten untersucht. Die Ergebnisse deuten darauf hin, dass wahrscheinlich erhebliche Leistungsreserven bestehen, aber um bedeutende Verbesserungen zu erzielen, könnten unpraktisch große Datensätze benötigt werden, die um Größenordnungen über das derzeit Machbare hinausgehen. Selbst bei sehr großen Datensätzen bleibt der klinische Nutzen ungewiss, da die projizierten Vorhersagegenauigkeiten besorgniserregend niedrig bleiben.

Bei gängigen Vorhersageaufgaben wie Alter und Geschlecht konkurrieren einfache lineare Modelle auch bei großen Stichprobengrößen mit dem Deep Learning, was auf Grenzen bei der Ausnutzung nichtlinearer Beziehungen aus Gehirnbildern hinweist.

Die Einbeziehung multimodaler Bildgebung kann die Genauigkeit vergleichbar mit der Verdoppelung der Stichprobengröße verbessern und unterstreicht den Wert komplementärer Informationen über die Modalitäten hinweg. Die optimalen Bildgebungsmodalitäten hängen jedoch von der Stichprobengröße ab, was die Bedeutung multimodaler Analysen unterstreicht.

Insgesamt erfordert die robuste Bewertung vielversprechender Techniken des maschinellen Lernens in der Psychiatrie eine sorgfältige Berücksichtigung von Stichprobengröße, Bildgebungsmodalität und Aufgabenkomplexität. Obwohl Aussichten auf inkrementellen Fortschritt bestehen, verbleiben erhebliche Hindernisse auf dem Weg zu maschinellem Lernen mit wirklichem klinischem Nutzen in der neurobildgebenden Psychiatrie.

## Abstract

Recent advances in neuroimaging techniques have enabled new insights into the neural bases of psychiatric disorders. However, applying machine learning to translate these findings into clinically useful tools for diagnosis and treatment in psychiatry faces significant constraints. Among other challenges, small sample sizes in neuroimaging studies limit complex machine learning approaches.

Here, we systematically evaluated neuroimaging machine learning prediction performance across a range of sample sizes and imaging modalities. Results suggest substantial performance reserves likely exist, but realizing major gains may require impractically large datasets orders of magnitude beyond what is currently feasible. Even with massive datasets, clinical usefulness remains uncertain as projected accuracies stay concerningly low.

On common prediction tasks like age and sex, simple linear models rival deep learning even at large sample sizes, suggesting limits in exploitability of nonlinear relationships from brain images.

Incorporating multimodal imaging can improve accuracy comparable to doubling sample size, highlighting the value of complementary information across modalities. However, the optimal imaging modalities depended on sample size, underscoring the importance of multimodal analyses.

Overall, robustly evaluating promising machine learning techniques in neuroimaging psychiatry requires carefully considering sample size, imaging modality, and task complexity. While prospects remain for incremental gains, substantial barriers persist in achieving machine learning solutions with actual clinical utility.