

**Bachelor Projekt 2021/22**  
**Digital Health & Machine Learning**  
**Prof. Dr. Christoph Lippert**

**Human-in-the-loop Deep Learning auf  
handgeschriebenen Patientenakten**

Deep Learning ist heutzutage der Go-To-Approach für Anwendungen im Bereich Computer Vision. Damit diese Algorithmen erfolgreich sind, müssen Menschen oftmals eine hohe Anzahl an qualitativ hochwertigen Annotationen erstellen. Dies nimmt ohne geeignete Softwareunterstützung viel Zeit und Kosten in Anspruch. In diesem Projekt soll eine nutzerfreundliche und smarte Annotations- und Lernsoftware entwickelt werden, die den Menschen in den Mittelpunkt stellt, Zeit einspart und Feedback gibt. Hierbei konzentrieren wir uns auf einen medizinischen Anwendungsfall: der automatischen Digitalisierung von handgeschriebenen Patientenakten.

The image shows a large, dense handwritten table on a grid background. The table has numerous columns and rows, with many cells containing numbers and some containing text. Red ink is used for highlighting certain cells and for writing annotations. The text is somewhat difficult to read due to the handwriting and the density of the data. There are some larger annotations in red ink on the right side of the table, including the name 'Allergien der Patientin' and some other illegible text. The overall appearance is that of a detailed data log or a complex spreadsheet used in a research or development project.

## **Anwendung**

Hierfür reisen wir gedanklich ins Jahr 1950, als der deutsche Arzt Walter Kempner an der Duke University eine diätetische Behandlung für schweren (bösartigen) Bluthochdruck einführte, der damals eine durchschnittliche Überlebenszeit von <6 Monaten hatte. Seine Diät, die arm an Natrium, Proteinen und einfachen Kohlenhydraten ist, war äußerst erfolgreich, wofür er die höchste Auszeichnung der American Heart Association erhielt. Kempner veröffentlichte Fallberichte und untersuchte zwischen 1950 und 1985 über 18.000 Patienten an der Duke University. Sein Team führte detaillierte Untersuchungen mit sich täglich wiederholenden Beobachtungen über Monate durch, die auf "gelben Blättern" (siehe oben) festgehalten wurden. Da diese Daten aufgrund fehlender Digitalisierung bis heute nicht systematisch analysiert werden konnten, ist bis heute unklar, welche Komponenten dieser Diät primär verantwortlich für seine Ergebnisse waren.

## **Ziel**

Das Ziel des Projekts ist die Planung, Entwicklung und das Testen einer Software, die den Digitalisierungsprozess Walter Kempner's handgeschriebener Patientenakten optimiert. Spezifisch soll die Software ermöglichen, Patientendaten einzulesen, Bounding Boxes (siehe Bild) inkl. Text als Annotation abzuspeichern und mit den Annotationen eine Deep Learning Pipeline für die Optical Character Recognition (OCR) zu trainieren und zu evaluieren. Der Fokus liegt hierbei auf dem Anwender: Smarte Annotationswerkzeuge und Feedbacksysteme über die aktuelle Erkennungsrate sind Beispiele hierfür.

## **Projektpartner**

In diesem Projekt arbeitet ihr mit mehreren Partnern zusammen: Die Duke University, Scalable Minds und dem Visian Team. Die Duke University stellt die Daten und initiale Annotationen zur Verfügung und gibt in regelmäßigen Meetings Fachexpertise, Anforderungen und Feedback zur Anwendung. Scalable Minds ist ein aus dem HPI aus- gegründetes Startup, das Bildanalysetools für medizinische Anwendungen entwickelt. Zusammen mit dem Visian Team, das aus einem ehemaligen Bachelorprojekt zur Annotation und dem Training von Deep Learning Modellen auf MRT-Daten hervorgegangen ist, erhaltet ihr Unterstützung bei der Planung und Umsetzung nutzerfreundlicher Software. Der Lehrstuhl unterstützt euch zudem fachlich bei der Umsetzung der Deep Learning Pipeline.

## **Über dich**

In diesem Projekt werden Deep Learning, Software Engineering und Usability Engineering mit einer faszinierenden Health-Anwendung kombiniert. Du/Ihr solltet Interesse in einem oder mehreren dieser Felder mitbringen und entsprechende Kurse belegt haben.

## Über uns



**Prof. Dr. Christoph Lippert**  
Chair Digital Health - Machine Learning



**Prof. Dr. Pao-Hwa Lin**  
Duke University



**MSc. Norman Rzepka**  
Scalable Minds



**BSc. Paul Brachmann**  
Visian

**BSc. Jonas Kordt**  
Visian

**BSc. Clara Uktar**  
Visian

**BSc. Maximilian Lindner**  
Visian