

Bildverarbeitung von Fluidbewegungen im Pharmabereich mittels KI

Viele Medikamente werden in flüssiger (fluid) Form transportiert. Beim Transport gibt es Schwingungen, die sich auf die flüssigen Medikamente auswirken und diese in Bewegung versetzen können. Diese Instabilitäten tragen das Risiko, dass die in der Flüssigkeit enthaltenen Wirkstoffe (oftmals Proteinverbindungen) zerstört werden. Roche Pharma (Schweiz) AG versucht diesen Zerstörungseffekt in Experimenten zu untersuchen und die besten Transportbedingungen zu finden. Die Experimente werden mit Hochgeschwindigkeitskameras aufgenommen und danach manuell analysiert. In dieser Arbeit wird ein Ansatz untersucht, die Analyse der aufgenommenen Videos zu automatisieren. Hierfür wird ein Convolutional Neural Network, kurz CNN, mit selbst annotierten Daten trainiert. Die Daten werden im Laboratorium der ZHAW auf einem hauseigenen Shaker aufgenommen und danach manuell mit dem MATLAB Video Labeler annotiert. Um die Daten für das Netzwerk brauchbar zu machen, werden diese zuerst durch eine eigene Pipeline aufbereitet. Um Overfitting aufgrund eines kleinen und unausgeglichene Datensets zu überwinden, werden zudem Data Augmentation und Balancierung der Daten durchgeführt. Die Effekte dieser Verfahren auf das trainierte Modell und den Trainingsverlauf werden in definierten Experimenten untersucht. Ebenfalls werden mehrere Modelle mit verschiedenen Hyperparametern und Architekturen trainiert, um die besten Einstellungen für das vorliegende Klassifizierungsproblem zu finden. Es wird festgestellt, dass sich Data Augmentation und die Balancierung positiv auf das Ergebnis auswirken. Von den untersuchten Hyperparametern werden die besten Einstellungen für Batchsize, Kernel-Grösse, Anzahl Kernels, sowie Anzahl Dense Layer und Neuronen identifiziert. Mit dem besten trainierten Modell konnte ein FBeta-Score von 96 % bei der Erkennung der Instabilitäten erzielt werden. Jedoch ist das Modell auf den verwendeten Shaker sowie die vorhandenen Lichtverhältnisse spezifiziert. Ein Wechsel in der Experimentumgebung kann die Ergebnisse drastisch verschlechtern und zur Folge müssten neue Trainingsdaten gesammelt werden. Durch ein grösseres und vielfältigeres Dataset kann die Genauigkeit und die Verwendbarkeit des Modells weiter gesteigert werden.

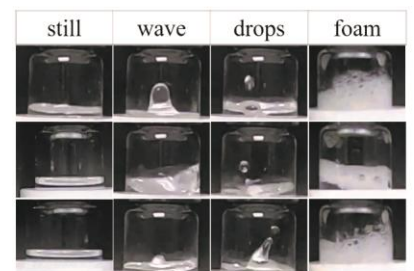


Diplomierende
Sven Aebersold
William Wong

Dozent
Gernot Kurt Boiger



Shaker im ZHAW ICP Lab



Beispielbilder der zu klassifizierenden Fluidzustände