

# Classify steps from surgical procedures using simulation data

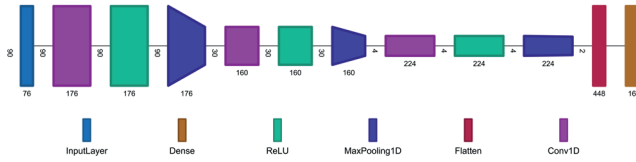


Abb. 1: CNN-Model mit 3 Convolutional + Max Pooling Layern

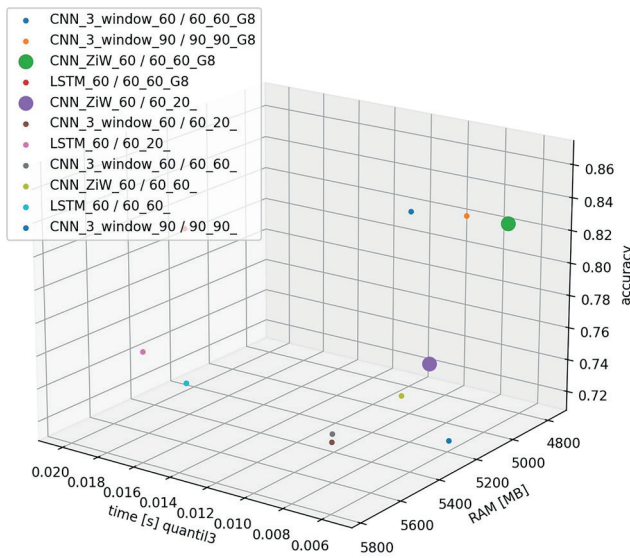


Abb. 2: Evaluation Zeit, RAM, Genauigkeit

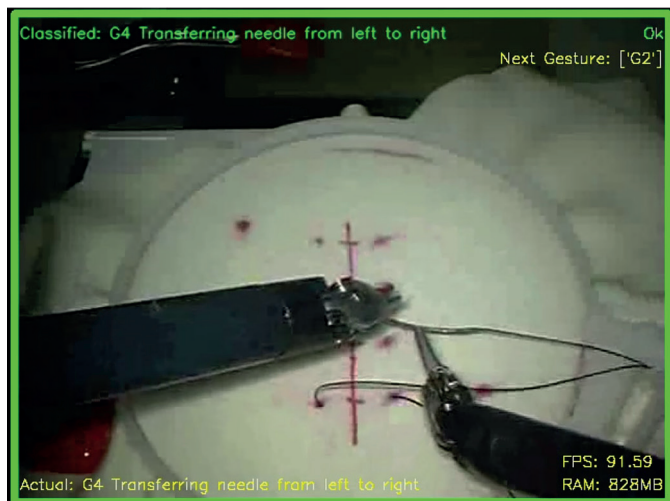


Abb. 3: Ausschnitt aus Gestenklassifizierung-Sequenz mit Actionevaluator

## Problemstellung

Die Firma VirtaMed AG erstellt Simulatoren, mit welchen chirurgische Aufgaben in einer Risikofreien Umgebung trainiert werden können. Für diese Simulatoren soll ein Gestenklassifizierer implementiert werden, welcher die ausgeführten Gesten automatisch erkennt. Zudem soll anhand der erkannten Gesten ein Feedback erstellt werden können, mit welchem dem Benutzer mitgeteilt wird, ob die Aufgabe korrekt ausgeführt wird.

## Lösungskonzept

Zu Beginn des Projektes wurde eine Recherche über Machine Learning Classifiers im Allgemeinen und Gestenerkennung in einem chirurgischen Umfeld durchgeführt. Es wurde entschieden, dass der Klassifizierer auf Convolutional Neural Networks basieren soll. Da keine aktuellen Simulator Daten vorlagen, wird das öffentliche JIGSAWS Dataset verwendet. Die erkannten Gesten werden an den Actionevaluator übergeben, welcher die Gesten auswertet.

## Realisierung

Die verschiedenen Teilsysteme (Modelle und Actionevaluator) wurden mit Python implementiert.

Zuerst mussten verschiedene Klassifizierer-Architekturen miteinander verglichen werden, um das optimale Modell zu evaluieren. Das Modell, welches am besten abgeschnitten hatte, ist in Abbildung 1 zu sehen. Anschliessend wurden zwei verschiedene Ansätze für den Actionevaluator implementiert.

## Ergebnisse

Die verschiedenen Modelle wurden miteinander verglichen, basierend auf der Genauigkeit, der Geschwindigkeit und benötigtem RAM. Der Vergleich ist in Abbildung 2 dargestellt, das beste Modell erzielte eine Genauigkeit von 86%. Des Weiteren wurden zwei Actionevaluator implementiert. Der eine zwingt den Benutzer nach einer bestimmten Abfolge vorzugehen, der andere sucht basierend auf verschiedenen Sequenzen die nächste erwartete Geste. Mit der cv2 Library konnten die Videos des JIGSWAS Dataset mit den Ergebnissen dargestellt werden (Abbildung 3).



**Diplomand**  
Frei Simon

**Dozent**  
Prof. A. Rumsch

**Themengebiet**  
Mechatronik/Automation/Robotik

**Projektpartner**  
VirtaMed AG

