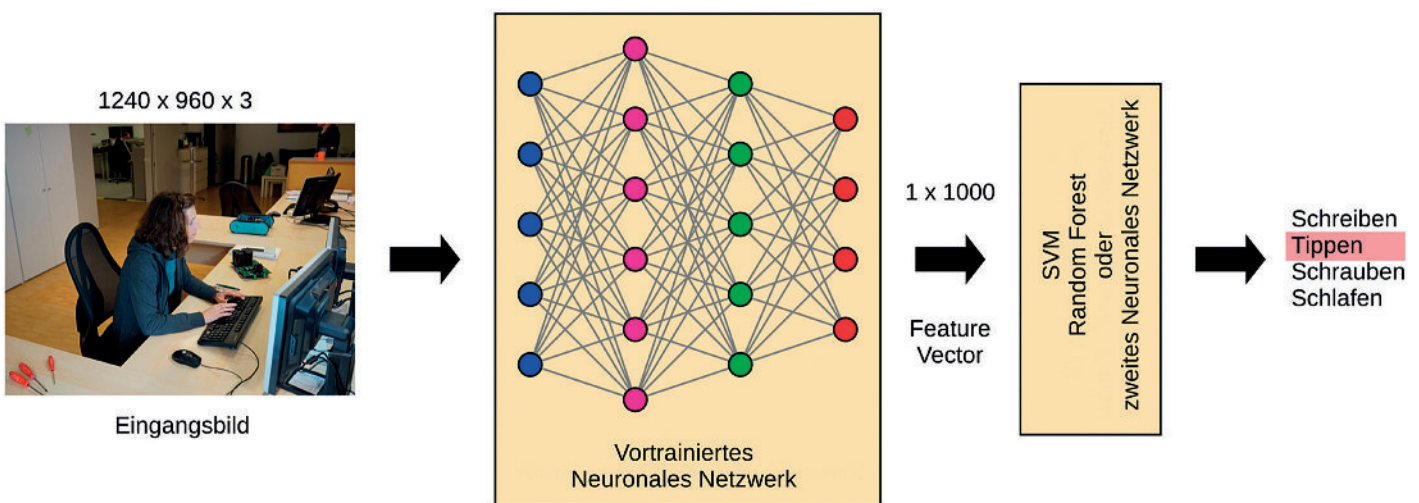


## Industrielles Machine Learning

# Was bisher schwer denkbar war ...

Die Rechenleistung von Mikroprozessoren und von industriellen PCs wächst beständig. Dadurch werden industrielle Lösungen nicht nur schöner und schneller, sondern es eröffnen sich zunehmend auch ganz neue Möglichkeiten und Strategien. In aller Munde ist heute das Maschinelle Lernen. Der folgende Selbstversuch zeigt Möglichkeiten auf.

» Jonas Schmid und Dr. Jürg M. Stettbacher



Stettbacher

Layout des geschilderten Projektbeispiels.

Zeitungen und Zeitschriften sind voll mit Gerede über die sogenannte Künstliche Intelligenz (KI) – oder englisch: Artificial Intelligence (AI). Eine wunderbare goldene Zukunft oder apokalyptische Abgründe werden uns prophezeit. Zur ersten Kategorie zählen die immer wieder gepriesenen intelligenten Kühlschränke, die dereinst aufgebrauchte (oder abgelaufene) Joghurts und Salate automatisch nachbestellen werden; zur zweiten Kategorie rechnen wir Siri, Alexa & Co., die heute schon im Auftrag der krakenhaften Torwächter des Internets – Google, Amazon, Apple, Microsoft – die intimsten Geheimnisse aushorchen. Nebst diesen fragwürdigen und teilweise beängstigenden Entwicklungen übersieht man leicht, dass die sogenannte Künstliche Intelligenz, oder besser:

das Machine Learning, längst auch in der Industrie angekommen ist. Hier eröffnen sich interessante, neuartige Möglichkeiten für das Smart Manufacturing, die Qualitätskontrolle, prädiktive Verfahren, Mustererkennung etc., die mit konventionellen Methoden – wenn überhaupt – nur schwer realisierbar wären.

### Für komplexe Aufgaben

Das Maschinelle Lernen wird für komplexe Aufgaben eingesetzt, häufig für die optische Erkennung, Klassifikation und die Bestimmung von Lage und Orientierung von Objekten jeder Art. Nebst technischen Komponenten zählen dazu auch Personen, Gesichter, Gesten, Handschriften etc. In autonomen Fahrzeugen beispielsweise werden damit Hin-

dernisse, andere Fahrzeuge, Passagen, Markierungen etc. erkannt und lokalisiert.

### Versuchsaufbau

Für die Lösung von derartigen Aufgaben braucht es heute weder ein Team von Wissenschaftlern noch ein leistungsstarkes Rechenzentrum. Anhand eines einfachen Beispiels soll das illustriert werden. Das automatische Erkennen von Farbstiften, Schrauben oder Zahnrädern in einem Kamerabild wäre aber zu einfach. Darum eine etwas kniffligere Aufgabenstellung: Eine beliebige Person sitzt an einem Schreibtisch mit PC. Dabei geht sie einer von vier möglichen Tätigkeiten nach, nämlich: 1) Schreiben auf Papier, 2) Tippen auf der Tastatur, 3) mit Werkzeug hantieren

und 4) eingeknickt sein. Ein künstliches Neuronales Netzwerk soll anhand eines digitalen Fotos die jeweilige Tätigkeit aus den vier erkennen. Erschwerend kommt dazu, dass jede Person individuell gemusterte Kleidung trägt und dass sich im Hintergrund der betreffenden Bilder weitere Personen befinden können.

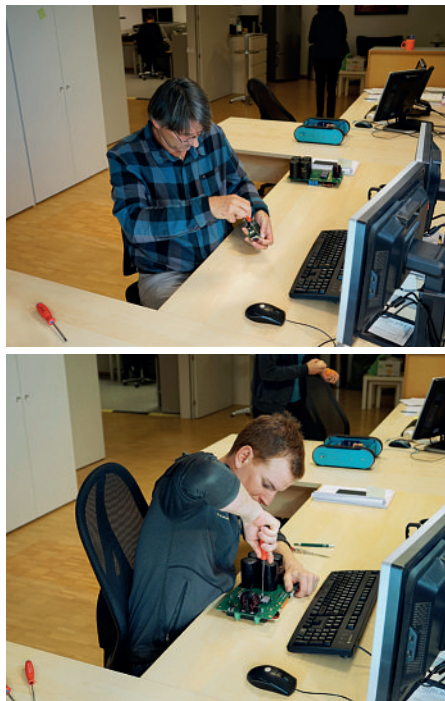
### Millionen von Trainingsbildern

In der jüngeren Vergangenheit wurden derartige Klassifizierungs-Aufgaben mit Hilfe von Deep Learning erfolgreich gelöst. Dabei wird ein grosses Neuronales Netz eingesetzt. Da dieses Netzwerk in der Regel viele hundert Millionen Parameter enthält, sind für sein Training tausende oder gar Millionen von Trainingsbildern notwendig. Zu jedem von diesen Bildern muss angegeben werden, welche Kategorie (im vorliegenden Beispiel: welche der vier Tätigkeiten) darauf abgebildet ist. In den meisten praktischen Fällen ist die Beschaffung und Klassifizierung einer solchen Menge von Bildern schlicht zu aufwändig und zu teuer.

### Komplexität reduziert

Um dies zu umgehen, wird für die Aufgabe hier ein vortrainiertes Netz verwendet. Ein derartiges Netzwerk ist bereits darauf trainiert, zum Beispiel 1000 verschiedene Features von Bildern zu separieren. Man nutzt das Neuronale Netz also nur als sogenannten Feature Extractor. Anstelle des vollständigen Bildes werden in der Folge für die Klassifikation nur diese 1000 Werte benutzt. Die Komplexität der Problemstellung reduziert sich damit auf einen überschaubaren Bruchteil.

Die Verarbeitung des Feature Vektor übernimmt zum Beispiel eine Support Vector Machine, ein Random Forest Klassifizierer oder ein zweites, kleineres Neuronales Netzwerk. Für das Training dieser zweiten Stufe reichen einige hundert bis einige tau-

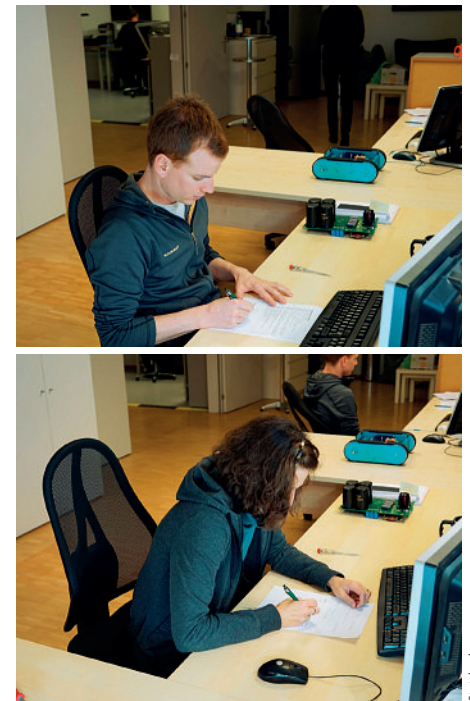


Bilder der verschiedenen Tätigkeiten für das Training des Klassifikators, links: Kategorie Schrauben, rechts: Kategorie Schreiben.

send Trainingsbilder. Für diese Aufgabe wurde ein Kamera-Setup aufgebaut, und ein kleines Hilfsprogramm vorbereitet, um jedem Bild – neben der Klasse der Tätigkeit – noch einige weitere Informationen zuzuordnen, zum Beispiel den Namen der Person im Bild. Mit diesem Aufbau wurden innerhalb von einem Tag knapp 900 Bilder mit neun verschiedenen Personen bei den vier Tätigkeiten aufgenommen, im Durchschnitt also 25 Bilder pro Person und Tätigkeit. Damit wurde der Klassifizierer trainiert.

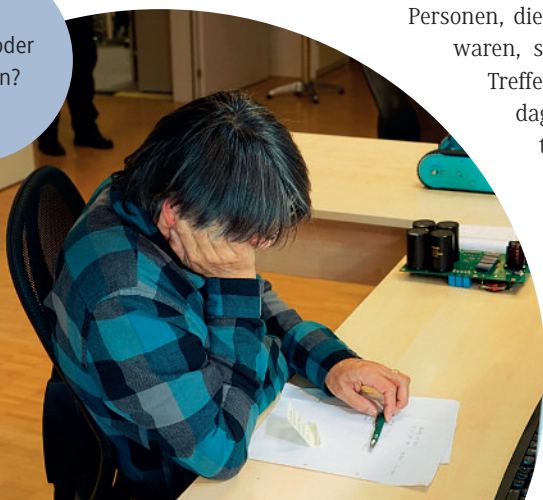
### Resultate

Bei der Verifikation des Systems dürfen diese Bilder selbstverständlich nicht nochmals verwendet werden. Stattdessen wurden neue Bilder aufgenommen. Enthalten diese Test-Bilder Personen, die nicht im Trainingsset dabei waren, so erreicht das System eine Trefferrate von 80 Prozent. Wird dagegen mit einer Person getestet, die in den Trainingsbildern bereits vorkam, so erreicht der Klassifizierer eine Zuverlässigkeit von rund 95 Prozent. Damit folgt: Trotz dem bescheidenen Umfang des Trainingsmaterials erreicht



der Klassifizierer bereits recht gute Trefferraten. Durch eine Erweiterung des Trainingssets auf zirka 4000 Bilder, entweder durch Augmentation – das heisst, durch Verändern der vorhandenen Bilder – oder mit zusätzlichen Aufnahmen der selben Personen und/oder mit weiteren Personen, kann die Trefferrate leicht bis gegen 100 Prozent gesteigert werden. Zu beachten ist aber, dass selbst die sogenannte Human Performance in der Regel keine 100 Prozent erreicht, denn es gibt immer Bilder, die sich nicht eindeutig zuordnen lassen. <<

Schlafen oder Schreiben?



### Firmenportrait

Stettbacher Signal Processing AG ist seit 25 Jahren F+E-Dienstleister für anspruchsvolle Projekte in industriellen Bereichen wie elektronische Mess-, Steuer-, Regelungs-, Antriebs-, Roboter- und Kommunikationstechnik. Typische Anwendungsgebiete sind Analytik, Qualitätssicherung, Medizin, Pharma, Verteidigung und Training, Autonome Systeme, etc. Zudem verfügt die Firma über eine eigene Produktion.

### Infoservice

Stettbacher Signal Processing AG  
Neugutstrasse 54, 8600 Dübendorf  
Tel. 043 299 57 23  
dsp@stettbacher.ch, www.stettbacher.ch