

## KI Industrialisierung

**Abstrakt: Während praktisch alle Voraussetzungen erfüllt sind, verhindern die heutigen Methoden zur Datengewinnung nicht nur eine allgemeine Verbreitung, sondern hemmen auch die Ausschöpfung des vollen Potenzials von Künstlicher Intelligenz. Wir beschreiben einen neuartigen Ansatz für die Datenproblematik, den wir als Eckpfeiler für eine Anwendung von KI in der Breite betrachten.**

Als Komplettlösungsanbieter für kundenspezifische Deep-Learning-Lösungen bietet die EvoTegra GmbH das gesamte Spektrum von Beratung und Datenservices bis hin zur Prozessintegration an. Mit der Fähigkeit, kostengünstige Lösungen in großem Maßstab anzubieten, richten wir uns an kleine und mittlere Unternehmen, die keine Möglichkeit haben ihre eigene KI-Kompetenz aufzubauen.

### Der aktuelle Stand von Künstlicher Intelligenz

Die Vorteile von KI in der Objekterkennung liegen in der Kombination von Genauigkeit und Geschwindigkeit in Kombination mit der Flexibilität praktisch alles zu erkennen, was ein Mensch sehen kann.

Während KI mittlerweile eine höhere Erkennungsgenauigkeit erreichen kann als der Mensch, sind KI Lösungen nicht grundsätzlich verlustfrei übertragbar. Jeder Unterschied in der Anwendung im Vergleich zu den Trainingsszenarien, kann die Erkennungsrate negativ beeinflussen. Auf der anderen Seite erhöht das Hinzufügen weiterer Szenarien die Wahrscheinlichkeit falscher Erkennungen. Während dies für Verbraucherprodukte akzeptabel sein mag, ist diese Eigenschaft für industrielle Lösungen in der Regel nicht geeignet. Daher glauben wir nicht an klassische Produkte mit integrierter KI (eine KI für alle Anwendungsfälle), sondern an einzelne plattformbasierte Lösungen, die für ein spezielles Szenario optimiert sind.

Die heutige KI-Technologie ist mittlerweile für die Produktion geeignet. Neuronale Netzwerke können zuverlässig trainiert und ausgeführt werden, C++ und Netzwerkoptimierung ermöglichen die Prozessintegration. Eine breite Palette verfügbarer Hardware, von eingebetteten Systemen bis hin zu High-End-Rechenzentrumslösungen, ermöglicht den Einsatz von KI-Lösungen in nahezu jedem Szenario.

Der Einsatz von KI-Lösungen in der Nähe von Maschinen (Edge-Processing) ist für Industrie- und Bildverarbeitungslösungen besonders wichtig. Während die hohe Bandbreite von Kameradaten in der Cloud nicht effizient verarbeitet werden kann, verhindern auch Unternehmensrichtlinien unter Umständen eine cloud-basierte Verarbeitung sensibler Produktionsdaten.

## Das Datenproblem

Auch wenn alle anderen Voraussetzungen erfüllt sind, bleiben die Daten das Haupthindernis für eine breite Einführung von KI. Für ein allgemeines Erkennungsproblem mit 100 verschiedenen Klassen empfehlen wir je nach Szenario insgesamt zwischen 100.000 und 1.000.000 Datenpunkte.

Zur Lösung des Datenproblems gibt es derzeit zwei gängige Ansätze:

- 1.) Der Kunde stellt die Daten mittels einer speziellen Anwendung bereit
- 2.) Ein Dienstleister

Nach unserer Schätzung kann eine einzelne Person bis zu 500 Bilder pro Tag markieren. Die Probleme mit negativem Einfluss auf die Qualität sind dabei die monotone Arbeitsbelastung sowie die Subjektivität der einzelnen Individuen. Infolgedessen erfordert dieser Prozess eine zusätzliche Qualitätskontrolle um sicherzustellen, dass die Ergebnisse in einem KI-System verwendet werden können.

Dienstleister benötigen eine Anlaufphase, in der die Datenanforderungen spezifiziert werden und die Schnittstellen des Anbieters implementiert werden müssen. Je nach Anbieter kann dieser Vorgang Tage bis Monate dauern. Aber auch während der produktiven Phase muss der Anbieter gegebenenfalls bei der Lösung von Fragen und Eckfällen unterstützt werden. Qualität, Durchsatz und Kosten variieren je nach Anbieter. Infolgedessen wird ein unabhängiger zusätzlicher Qualitätskontrollprozess empfohlen.

Die überwiegende Mehrheit der potenziellen Projekte von heute wird an diesem Punkt enden, da die impliziten Kosten und Risiken der Datenerfassung keinen validen Business Case ermöglichen.

Unabhängig von der Datenstrategie kann das zentrale Problem beider Ansätze mit dem Wort „Veränderung“ beschrieben werden. Die anfängliche Definition des Klassenmodells basiert normalerweise auf menschlicher Erfahrung, auch bekannt als Annahme. Die Wahrscheinlichkeit, dass ein Klassenmodell während des Projekts angepasst werden muss, ist daher hoch. Da diese Zusatzaufwände in der Regel nicht geplant sind, überschreiten KI-Projekte an dieser Stelle in der Regel das Budget.

Von Menschen bearbeitete Daten haben jedoch einen weiteren negativen Effekt auf Künstliche Intelligenz.

## Die bittere Lektion

Die menschliche Wahrnehmung wird unter anderem durch Erwartung, Ablenkung, Kontext, Schlafniveau, Stimmung oder genetische Disposition beeinflusst. Unsere Wahrnehmung ist daher sehr subjektiv.

Die heute in der Praxis verwendete "schwache künstliche Intelligenz" ist eine selbstoptimierende mathematische Annäherung an eine unbekannt komplexe Funktion. Da die künstliche Intelligenz von heute kein Bewusstsein dafür entwickelt was sie lernt, ist KI vollständig objektiv.

Im März 2019 veröffentlichte Rich Sutton, Forscher bei Google DeepMind und Professor für Informatik, den Artikel „The bitter lesson“. In dem Artikel beschreibt er, dass KI-Forscher oft menschliches Wissen als kurzfristige Lösung zur Verbesserung der Ergebnisse in ihre Systeme einbauen, was jedoch in der weiteren Entwicklung typischerweise zu einer Stagnation führt. Was er als die größte Lektion in 70 Jahren KI-Forschung bezeichnet ist, dass die Berechnung mit mehr Rechenleistung bei weitem die effektivste Lösung zur Verbesserung der KI ist.

Um das volle Potenzial der KI auszuschöpfen, darf menschliches Wissen daher nicht eng mit künstlicher Intelligenz verknüpft werden. Genau dies erfolgt jedoch in allen aktuellen Lösungen des Datenproblems. Das menschliche Wissen kommt in Form der Definition von Klassen und vom Menschen gekennzeichneten Daten. Ein objektives System mit subjektiven Daten zu vermischen bedeutet, objektiv inkonsistente Daten bereitzustellen. Dies hat negative Auswirkungen auf die Verallgemeinerung und verhindert daher, dass die heutigen Deep Learning-Lösungen ihr volles Potenzial nicht ausschöpfen können.

### **Subjektive versus objektive Wahrnehmung**

Um unsere Annahme zu bestätigen, haben wir unsere Methode auf den Tsinghua-Tencent 100K Datensatz zur Verkehrszeichenerkennung angewendet. Dabei haben wir ausschließlich die Bilder der Kategorie „Negative“ analysiert, die keine Verkehrszeichen enthalten sollten. Dabei ergab unsere Analyse, dass tatsächlich zahlreiche Bilder Verkehrszeichen enthielten, die für Menschen entweder schwer zu erkennen waren oder sich in ungewöhnlichen Stellen im Bild befanden. Dies lässt vermuten, dass diese Bilder von Menschen markiert wurden. Die ausgewählten Bilder enthalten ausschließlich Schilder in der Frontalansicht, die wir zur besseren Erkennbarkeit mit einem grünen Rahmen markiert.





## Die gelernte Lektion

Als Lösungsanbieter sind wir auf kostengünstige, qualitativ hochwertige Daten in großen Mengen und kurzer Zeit angewiesen. Aufgrund unserer Erfahrungen mit von Menschen beschrifteten Daten haben wir diese Schlussfolgerungen in den letzten Jahren in unseren Datenprozessen umgesetzt. Statt die Daten direkt von Menschen beschriften zu lassen, werden die Daten in einem von Menschen überwachten Prozess hochautomatisiert beschriftet.

Der Aufwand zum Beschriften von Bildern hängt allgemein von der Gesamtzahl der Objekte und der Anzahl der Klassen ab. Bevor jedoch Daten automatisch markiert werden können, muss die automatische Verarbeitung erst eingerichtet werden. Abhängig von der Projektphase und den Anforderungen werden beim hochautomatisierten Markieren unterschiedliche Techniken aus den Bereichen (nicht) überwacht maschinelles Lernen aber auch klassischen Methoden verwendet. Der Aufwand bei der anschließenden Datenverarbeitung ist generell anfangs höher und nimmt mit der Zeit ab.

In einem Referenzszenario von:

- 1.000.000 Gesamtbilder
- 50% der Bilder enthalten ein Objekt
- 100 verschiedene Klassen

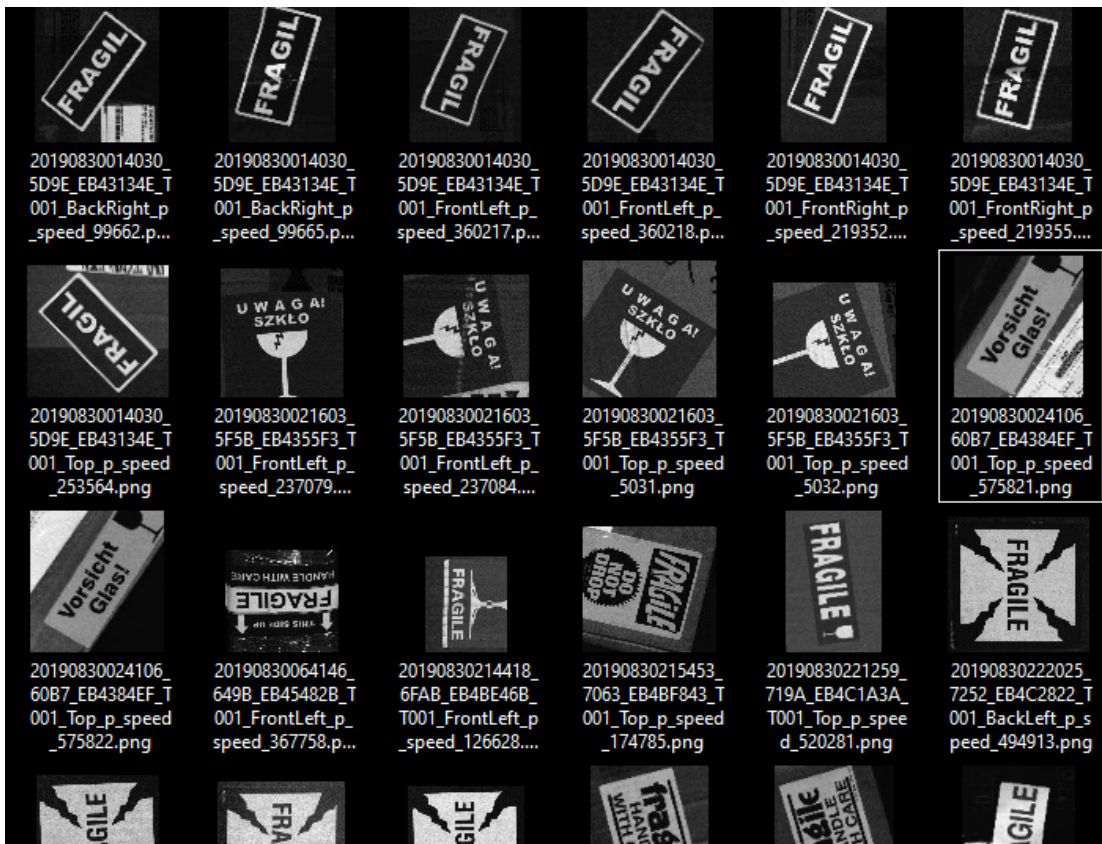
kann eine einzelne Person die automatische Markierung von *durchschnittlich* 10.000 bis 20.000 Bildern pro Tag überwachen. Im Vergleich zur Bearbeitung durch den Menschen bedeutet dies eine Produktivitätssteigerung von bis zu 4000%. Normalerweise werden dabei 100.000 Bilder in einem Durchlauf verarbeitet.

Darüber hinaus können mit Hilfe der hochautomatisierten Prozesse Änderungen der Klassendefinition beliebig oft umgesetzt werden, wodurch das Projektrisiko erheblich reduziert wird. Gleichzeitig erhält man Daten auf einem Qualitätsniveau, das nicht nur erheblich besser ist, sondern auch während des gesamten Projekts ständig überwacht wird.

Im Ergebnis verwenden wir jetzt ein objektives System mit objektiven Daten, die von der menschlichen Intuition gesteuert werden. Unsere Beobachtungen zeigen, dass die Verallgemeinerung dadurch ein Niveau erreicht welches es der KI ermöglicht die Fehler in den menschlichen Annahmen zu finden. Dies ermöglicht der KI Bilder nicht mehr nur durch ihren Inhalt, sondern nach ihrer Bedeutung zu erkennen.

Menschliche Intuition und Ethik ergänzen sich perfekt mit der Produktivität und Objektivität der KI. Da die KI falsche oder unvollständige menschliche Annahmen erkennt, ist das Lernen beidseitig. Und da die menschliche Intuition mittelfristig nicht durch künstliche Intelligenz ersetzt werden kann, empfehlen wir generell den Einsatz von KI als Assistenzsystem.

## Der Rosetta Stein der KI



Dies ist ein Beispiel für unseren automatisierten Prozess, bei dem die KI „Fragile“ Aufkleber erkennt. Aufgrund der fehlenden Standardisierung existieren nahezu unbegrenzte Variationen dieses Aufklebers in verschiedenen Sprachen und Symbolen. Trotz der Tatsache, dass viele dieser Etiketten nicht in den Trainingsdaten enthalten sind, zeigen alle obigen Bilder einen Aufkleber mit der *Bedeutung* „Zerbrechlich“ und werden von uns daher als richtig bewertet. Die Wahrscheinlichkeit, dass eine Gruppe von Personen die Bilder auf ähnliche Weise konsistent bearbeiten würde, ist im Grunde nicht gegeben.

Der hier beschriebene Effekt ist keine Einzel- sondern eine allgemeine Beobachtung. Daher muss davon ausgegangen werden, dass sich die Klassendefinition während eines Projekts sogar mehrfach ändert. Ein hochautomatisierter Prozess ist daher für diese Methode des evolutionären Lernens zwingend erforderlich.

## Ergebnis

Um KI zu verstehen, muss man die Welt mit den Augen einer KI sehen. Objektive Systeme, die mit objektiven Daten verwendet werden, nutzen das verborgene Potenzial und steigern die Produktivität von KI-unterstützten Mitarbeitern um bis zu 4000%. Die Methode des evolutionären Lernen macht Daten für höchst spezifische Anwendungsfälle in kurzer Zeit und in großen Mengen zu geringen Kosten verfügbar, und ebnet der KI den Weg zu einer breiten Anwendung zum Vorteil des Menschen.

Das ist keine Vision. Das ist unser alltägliches Geschäft