

AI IN DER VIDEO ERKENNUNG

BEURTEILUNG
VIDEOMATERIAL
MIT ALGORITHMEN
FÜR MASCHINELLES
LERNEN



AI IN DER VIDEO ERKENNUNG

BEURTEILUNG VIDEOMATERIAL MIT ALGORITHMEN FÜR MASCHINELLES LERNEN

IN VIELEN UNTERNEHMEN FÄLLT IN INSPEKTION, QUALITÄTSSICHERUNG, ÜBERWACHUNG ODER PROZESSMANAGEMENT VIDEOMATERIAL AN. DIE BEURTEILUNG DIESER BILDER ERFORDERT MEIST ERHEBLICHEN PERSONELLEN AUFWAND. KÜNSTLICHE INTELLIGENZ (AI) KANN DIES ERLEICHTERN.

Ohne AI geht es fast schon gar nicht mehr, die Bandbreite ist riesig: von selbstfahrenden Autos und Apps, die beurteilen können, ob ein Mutterfleck hautkrebsgefährdet ist, bis zur automatischen Brunsterkennung bei Rindern mithilfe von Drohne-Bildern - alles ermöglicht durch künstliche Intelligenz, insbesondere maschinelles Lernen bis hin zum Deep-Learning. Hier beschreiben wir den Einsatz des maschinellen Lernens in der automatischen Erkennung und Klassifizierung von Videobildern. Für welche Situationen der Videoerkennung ist AI einsetzbar? Welche Vorteile bietet es? Wie sieht das Geschäftsszenario aus?

ANWENDUNGSSPEKTRUM DER VIDEOERKENNUNG

In diesem Frühjahr wurden in den Niederlanden viele Nester mit Kameras versehen, sodass Vogelfreunde beim Brüten zuschauen konnten. Nur passiert bei solchen Feeds oft stundenlang gar nichts. Da wäre es praktisch, wenn ein Algorithmus automatisch nur die besonderen Ereignisse erfasst und am Ende des Tages ein Highlights-Video zeigt? Eine ähnliche AI-Anwendung kann zur Auswertung der Bilder von Überwachungskameras benutzt werden.

Die Überwachungsmitarbeiter brauchen dann nicht mehr alle Kamera-Feeds zu verfolgen, weil ihnen nur Echtzeit-Bilder gezeigt werden, wo wirklich etwas passiert. Die Bilder werden dann immer noch von Menschen ausgewertet, nur brauchen sie keine Bilder mehr zu überwachen, auf denen gar nichts passiert.

Noch eine AI-Kategorie: die automatische (Vorrats-)Erfassung von Objekten, z. B. von Produkten auf einer Verpackungslinie. In den meisten Schlachthöfen werden unterschiedliche Teile nacheinander verpackt, weil die Linie die Fleischarten nicht selbst identifizieren kann. Der Mitarbeiter sortiert das Fleisch in Kisten, dadurch liegt das Fleisch zu lang und trocknet aus. Da wäre es doch praktisch, wenn die Kamera ein rundes Steak von einem Kotelett unterscheiden und automatisch die nächsten Schritte steuern kann? Dann könnte man von der seriellen zur parallelen Verarbeitung wechseln.

Weitere Beispiele für die Video-Verfolgung von Objekten: Zählen von Containern an einem Umschlagstandort, Verfolgung des Wachstums von landwirtschaftlichen Nutzpflanzen oder Unterscheidung zwischen Nutzpflanzen und Unkraut.

Eine Anwendung der Video-Erkennung ist die Inspektion. So findet die Deichschau nach wie vor durch tatsächliches Begehen statt. Bei plötzlichem Anstieg des Wasserstands sind die Inspektoren ständig im Einsatz, um alle kritischen Deiche zu inspizieren. Wäre es nicht viel sicherer, wenn eine Drohne einen Deich aus nächster Nähe filmen könnte und ein trainierter Algorithmus den Inspektoren die Risikostellen zeigt?

'MIT IN DER BILDERKENNUNG TRAINIERTEN AI-MODELLEN KANN VIDEOMATERIAL AUTOMATISCH AUSGEWERTET WERDEN'

AI IN DER VIDEOERKENNUNG: WIE FUNKTIONIERT DAS?

Mit in der Bilderkennung trainierten AI-Modellen lassen sich Videobilder automatisch auswerten. Damit das Modell richtig trainiert werden kann, muss es ausreichend Bilddaten mit einem klaren Bezug zum AI-Einsatzziel geben. Soll eine Überwachungskamera Einbrecher von Mitarbeitern und Passanten unterscheiden, braucht sie ausreichend Bilder von allen drei Situationen. Nur hat man relativ wenige Bilder von Einbrechern, das macht es fast unmöglich, Mitarbeiter und Einbrecher voneinander zu unterscheiden. Man kann aber Bilder von Personen, die den Standort betreten, von reinen Passanten unterscheiden.

Außer genug Datenmaterial mit einem klaren Bezug zum Überwachungszweck muss der Lernprozess auch überwacht werden. Typisch für AI: das Modell lernt selbst von allen neuen Daten. Das geht nur, wenn die neuen Daten auch einen klaren Bezug zum Zweck haben, also der Algorithmus verstehen kann, was er lernen soll. Registriert die Kamera eindringende Kleintiere wie Katzen oder Kaninchen, kann sie lernen, diese als solche zu erkennen. Aber dieser Lernprozess muss überwacht werden, damit der Algorithmus nicht einen unter dem Stacheldraht hindurchkriechenden Mann für einen Hund hält.

AI kann nämlich etwas Entscheidendes nicht so gut, was uns als Menschen keine Schwierigkeiten bereitet: Kontext erkennen und entsprechend handeln. AI erkennt genau das, wofür sie spezifisch trainiert wurde - und sonst gar nichts. Ein Beispiel: Bei der Erkennung von Brustkrebs aus Mammographien sind diese Algorithmen bei der Diagnose der Krebs-Anfangsstadien besser und schneller als der Radiologe - können aber keine anderen Erkrankungen erkennen. Dadurch eignen sie sich sehr gut für die Reihenuntersuchung, aber wenn Frauen mit allgemeinen Beschwerden zum Arzt kommen, ist eine breitere Sicht erforderlich.

ANWENDUNGSBEISPIEL: STRAßENINSPEKTIONEN

Wir möchten das jetzt an einem praktischen Anwendungsbeispiel erläutern: Inspektion von Asphalt durch Bauunternehmen BAM. Zur Inspektion werden von einem Aufnahmeauto Kamerabilder aufgezeichnet, dann wird das gesamte Bildmaterial von Experten gesichtet. Wenn sie Schäden erkennen, markieren sie sie mit einem Rahmen und notieren einen Vermerk.

Zusammen mit ICT Group hat BAM untersucht, ob die Schadenserkenkung auch von einem selbstlernenden Algorithmus ausgeführt werden kann. Zum Einsatz kam ein Modell, das mit 2500 Bildern von acht verschiedenen Asphalt-Schadenstypen trainiert wurde. So lernte das Modell, Schäden zu erkennen und Schadenstypen zu unterscheiden.

Im Lernprozess entdeckte das Team, dass einige Schäden zu breit eingerahmt wurden, was das Lernen für den Algorithmus erschwerte. So erkannte der Algorithmus Tierkörper auf dem Asphalt nicht richtig, weil im Vermerk auch größere saubere Asphaltflächen gezeigt wurden. Durch eine Verkleinerung der Rahmengröße braucht der Algorithmus jetzt weniger „Störinformationen“ zu verarbeiten und werden die Prognosen genauer. Darüber hinaus sind für Menschen schwer erkennbare Schäden (etwa Kornausbrüche) das auch für den Algorithmus. Logisch, weil auch Bilder von nicht als Schaden eingerahmten Kornausbrüchen in das Trainingsmaterial aufgenommen wurden. Manchmal ist der Unterschied zwischen einer einwandfreien und einer beschädigten Straße wirklich klein. Für solche Schadenstypen wird ein größerer Trainingsatz in den Algorithmus aufgenommen. Schlussendlich soll die menschliche Experten-Auswertung durch eine Qualifikation von AI-erkannten Schäden ersetzt werden.

Zur fernbedienten Auslesung der installierten Basis gibt es zwei Methoden: Am einfachsten ist es, wenn Ihre eigene auf dem Modell installierte Software so strukturiert ist, dass falsch-negative Meldungen verhindert werden, d. h. es werden keine Bilder mit Schäden verpasst. Darum enthält die erste Version vergleichsweise mehr falsch-positive Meldungen: Bilder, an denen der Algorithmus zweifelte und die von menschlichen Experten in der Nachbewertung als unbeschädigt eingestuft wurden. Dieser Input floss dann in die Feinjustierung des Algorithmus ein, damit dieser den Grad der Asphalt Schäden besser vorhersagen kann. Der erste von BAM verwendete Algorithmus konnte 80 % der Bilder als unbeschädigt einstufen, d. h. die Inspektoren brauchten nur die restlichen 20 % der Bilder auszuwerten. Die eingesparte Zeit wurde zum weiteren Training des Algorithmus und Überwachung des Lernprozesses genutzt.

Längerfristig wird sich dieses Verhältnis zu 99 % automatischer Erkennung und 1 % menschlicher Auswertung entwickeln. Auch auf lange Sicht wird menschliches Eingreifen erforderlich bleiben, weil es immer untypische Bilder geben kann, die der Algorithmus noch nie gesehen hat und nicht erklären kann. Weil solche Ausnahmen sehr selten sind, hat der Algorithmus zu wenig Trainingsmaterial, um diese Situationen selbst zu erkennen. Darum wird diese AI-Anwendung eine Mensch-Maschinen-Kombination bleiben.

EIN DATENSATZ, VERSCHIEDENE TRAININGSZIELE

Manchmal kann derselbe Datensatz zu verschiedenen Zwecken verwendet werden. Im BAM-Beispiel sind die Asphaltbilder angepasst von den Aufnahmen einer Horus-Autokamera, ähnlich wie das von Google Streetview verwendete Gerät. Die Originalbilder enthalten also noch viel mehr Informationen. Beispielsweise können sie auch Ort und Zustand von Verkehrszeichen zeigen. Mit diesen Informationen könnte man deren Wartung planen. Außerdem kann beurteilt werden, wie übersichtlich Verkehrssituationen sind: ist ihr Standort sinnvoll, sind die Zeichen deutlich sichtbar? Gerade für selbstfahrende Fahrzeuge sind übersichtliche Situationen noch viel wichtiger. Mithilfe von AI lässt sich dies viel schneller auswerten.

Das Beispiel der Verkehrszeichen zeigt auch, dass für jede Anwendung anderes Trainingsmaterial benötigt wird. In jedem Land gibt es nur eine bestimmte Anzahl von Verkehrszeichen, die der Algorithmus nur einmal zu lernen braucht. Trotzdem kann die AI z. B. eine Person mit einem auf dem T-Shirt abgedruckten Verkehrszeichen nicht richtig erkennen, darum bleibt menschliches Eingreifen zur Beurteilung aller Situationen unverzichtbar. Das Trainingsmaterial kann aber viel kleiner sein als bei einer Anwendung zur Erkennung von Asphalt Schäden.

VORTEILE DES EINSATZES DER AI

DER EINSATZ VON AI BIETET ZAHLREICHE VORTEILE:

1. HÖHERE EFFIZIENZ

Menschen brauchen sich nicht mehr sämtliche Kamerabilder anzuschauen, sondern nur Bilder mit möglichen Abweichungen. So bewältigen sie mehr Arbeit in weniger Zeit.

2. INTERESSANTERE ARBEIT

Die Arbeit der Experten, die die Bilder auswerten, wird interessanter, weil sie nur die komplizierten Fälle zu beurteilen brauchen.

3. HÖHERE QUALITÄT

Anders als Menschen werden Algorithmen nicht müde oder unkonzentriert. Der Output eines AI-Modells ist immer vorhersehbar. Seine Qualität hängt davon ab, wie gut der Algorithmus trainiert wurde. Zu Anfang schneidet ein Algorithmus suboptimal ab. Die Herausforderung besteht darin, das Modell so zu trainieren, dass die Zahl der falsch-negativen Meldungen auf 0 gesenkt wird.

4. SCHNELLERE ENTSCHEIDUNGEN

Weil Mitarbeiter mehr Arbeit in kürzerer Zeit bewältigen können, können sie den Entscheidungsprozess für Kamerabilder, bei denen ein Eingreifen erforderlich ist, beschleunigen. Besonders in Notfällen ist das ein großer Vorteil.

5. KONTINUITÄT

Ein AI-Modul wird nicht krank, hat keinen Urlaub und arbeitet rund um die Uhr.

6. SKALIERBAR

Ein AI-Modell lässt sich leicht auf andere virtuelle Maschinen duplizieren, was die Verarbeitungsgeschwindigkeit erhöht. So wird die gleichzeitige Verarbeitung von riesigen Bildmaterial-Mengen möglich.

7. KOSTENGÜNSTIGER

Alle oben beschriebenen Vorteile - besonders die gesteigerte Effizienz - ermöglichen erhebliche Kostenvorteile.

8. NEUE GESCHÄFTSMODELLE

Durch niedrigere Kosten bei höherer Qualität werden Chancen für neue Geschäftsmodelle geschaffen. Beispielsweise können die Inspektionen öfter durchgeführt werden, sodass Fehler schneller behoben werden. Außerdem wird eine viel frühere Schadenserkennung und exakte Verfolgung des Schadensprozesses möglich, was ganz neue Möglichkeiten für die Wartungsplanung eröffnet.

9. MENSCHLICHE KAPAZITÄT

Die menschliche Kapazität ist nicht mehr der begrenzende Faktor.

ZAHLEICHE GESCHÄFTSSZENARIEN IN GREIFBARER NÄHE

In der Bilderkennung gibt es viele mögliche AI-Anwendungen, davon werden aber noch sehr wenig in die Praxis umgesetzt. Man glaubt noch zu oft, diese Technologie stünde noch in den Kinderschuhen. Das Gegenteil ist der Fall. Schon in den 1980er Jahren wurden neurale Netzwerke eingesetzt. Die ersten Bilderkennungsanwendungen wurden in den 1990ern entwickelt. In unserem Jahrtausend kommt die Bilderkennung z. B. in der Polizeiarbeit breit zum Einsatz (Nummernschilderkennung/automatische Gesichtserkennung). Auch in der Medizin wird AI intensiv genutzt, z. B. bei der Lokalisierung von Krebszellen in einem MRI. Die Technologie ist ausgereift, dadurch sind die Kosten von selbstlernenden Modellen gesunken.

Die tatsächlichen Kosten werden von der Qualität der Trainingsdaten bestimmt. Je größer der Datensatz und je genauer die Vermerke, desto schneller kann der Algorithmus trainiert werden. Nur wenn das Trainingsmaterial schlecht aufbereitet ist, muss erst dessen Qualität verbessert werden. Wie schon gesagt, lernt ein Algorithmus viel schneller, wenn er sofort versteht, was er in einem Bild oder Film erkennen soll.

Wollen Sie herausfinden, wie Ihr Geschäftsszenario aussehen könnte, sprechen Sie uns doch einfach mal an. Wir können die Qualität Ihrer Daten beurteilen und Ihnen eine Vorstellung von den Kosten geben, die zur Verbesserung erforderlich sind.

FÜR WEITERE INFORMATIONEN STEHEN IHNEN GERN ZUR VERFÜGUNG:

BAS LAMME

BUSINESS DEVELOPMENT MANAGER

+31 (0)6 25 73 33 27

bas.lamme@orangenxt.com

FRANK THOMSON

DATA SCIENTIST

+31 (0)6 28 60 09 89

frank.thomson@orangenxt.com



TURBO-BOOST YOUR DIGITAL TRANSFORMATION



ORANGENXT

PEOPLE • DEVICES • DATA

John F. Kennedylaan 2 • 5612 AB Eindhoven • Postfach: Kopenhagen 9 • 2993 LL Barendrecht
+31 (0)88 908 2000 • info@orangenxt.com

WWW.ORANGENXT.COM

an **ICT|GROUP** company