



Neue Geschäftsmodelle dank Advanced Analytics

Einblicke in die Verfahren und die Skalierbarkeit

„Big Data ist in aller Munde ...“, so fangen viele Vorträge an, in denen es um Big Data geht. Im Zusammenhang mit Analysemöglichkeiten wird oft anhand von Beispielen gezeigt, wie aus unterschiedlichen Datenarten, von strukturiert über polystrukturiert bis unstrukturiert, Erkenntnisse gewonnen werden.

Dabei sollen die beiden berühmten „Vs“ des Big Data, Volume und Velocity, mittels der Skalierbarkeit der Daten bezwingbar werden. Wie Bild 1 zeigt, begleitet uns diese Einsicht laut einer Google-Analyse bereits seit 2010.

Ein Großteil des Vs „Volume“ besteht heute allerdings aus unstrukturierten Daten wie Bildern und Videos. Um diese Daten konsequent mit zu berücksichtigen, bedienen sich Experten Methoden der deskriptiven Analytik, die aus der Business-Intelligence-Welt bekannt ist und sich nach klassischen Fragestellungen reaktiv mit vergangenen Ereignissen befasst. Diese Methoden werden durch den proaktiven Zweig der Datenanalytik erweitert und münden damit in einem neuen Vorgehen: Advanced Analytics. Der Umfang von Advanced Analytics reicht von Bilderkennung und Bildklassifizierung über die semantische Auswertung von Texten bis zur prädiktiven Modellierung von Zusammenhängen. Somit fungiert Advanced Analytics als ein großer Treiber für neue Geschäftsmodelle.

Wir fokussieren uns in diesem Artikel ganz auf die Bildanalytik und versuchen die Domänen der Bilderkennung sowie der zugrundeliegenden Technologien klar abzugrenzen und auf die Skalierbarkeit einzugehen.

Domänen der Bildanalytik

Eine maschinelle Bildanalyse lässt sich in viele Unterbereiche aufteilen, vom trivialen Sortieren nach Farbstimmungen bis hin zu komplexen Themen, die heutzutage unter dem Begriff „Computer Vision“ zusammengefasst werden. Auf drei repräsentative Bereiche gehen wir in diesem Zusammenhang näher ein und stellen anschließend eine Verbindung zu Machine Learning und Deep Learning her:

- Selektion ähnlicher Bilder
- Objekterkennung
- Objektklassifizierung

Fangen wir mit einer eher simplen Problemstellung in der Bildanalytik an, nämlich dem Selektieren von ähnlichen Bildern in einer Datenbank. Hierzu

werden üblicherweise Hashing- oder Histogrammverfahren verwendet: Das heißt, ein Programm erstellt pro Bild ein Histogramm (etwa LBPH – Local Binary Pattern Histogramm) und clustert anschließend die sortierten Vektoren entsprechend ihres Abstands zueinander. Aus den Clustern wird dann wiederum ein Index für die spätere Suchfunktion gebildet.

LIRE (Lucene Image Retrieval), ein Plugin für Lucene, funktioniert zum Beispiel auf diese Weise und ermöglicht so eine Suche nach ähnlich aussehenden Bildern. Dabei geht LIRE nach einem fest definierten Regelsatz vor. Das System unterscheidet also nicht zwischen den sichtbaren Objekten im Bild sondern abstrahiert seine Farbstimmung. [LIRE 2017]

Objekterkennung

In der Bildanalytik geht es aber auch deutlich komplexer. Ein Paradebeispiel ist die Objekterkennung. So benötigt ein Algorithmus für die Objekterkennung eine genaue Vorgabe. Er muss wissen, nach welchen Alleinstellungs-

WEB-TIPP:
www.optitz-consulting.com

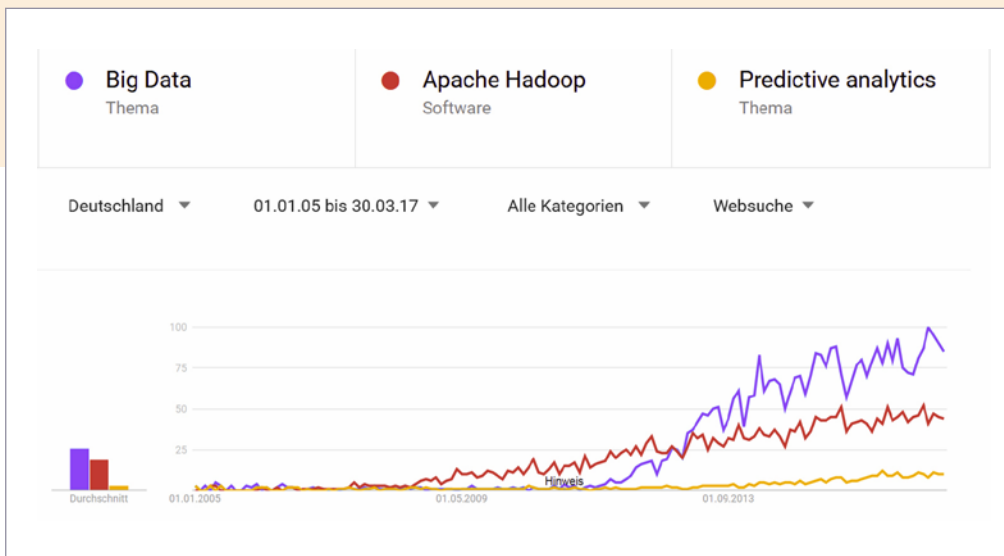
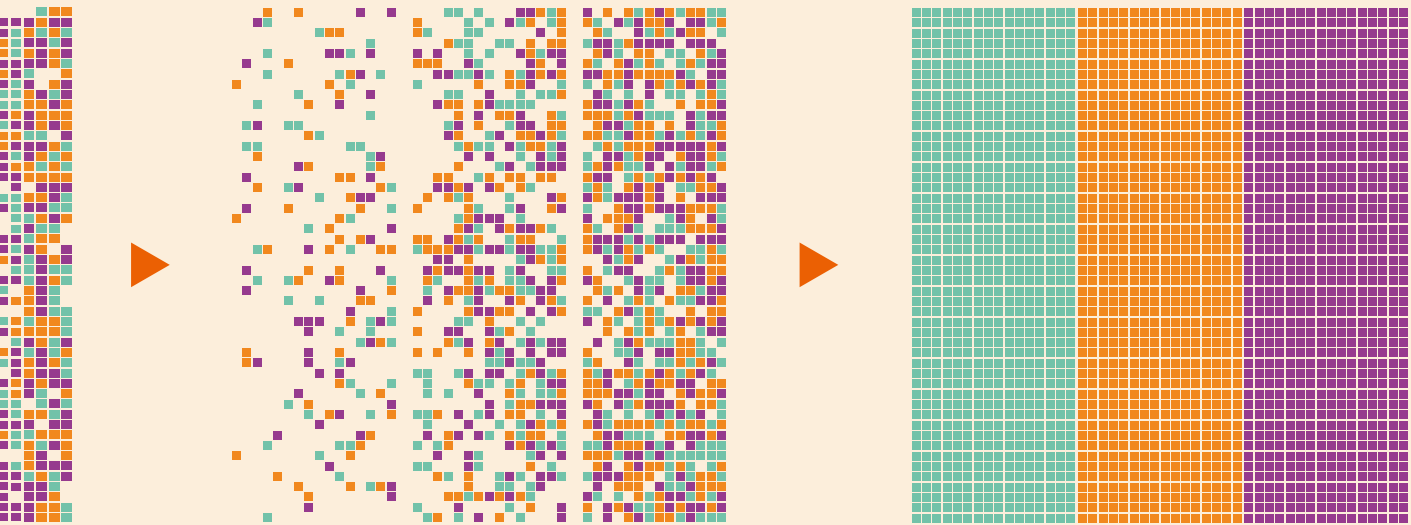


Bild 1: Google-Auswertung: Big Data wird als Begriff in Deutschland ab ca. 2010 deutlich wahrnehmbar .

merkmalen (Features) in einem Bild gesucht werden soll. Das bedeutet, dem Algorithmus werden sogenannte Basis-Features übergeben – also eine Art Schablone, mit deren Hilfe das Bild mit dem gesuchten Objekt abgeglichen wird. So werden beispielsweise für eine Gesichtserkennung mehrere Features definiert, die ein menschliches Gesicht beschreiben. Meist sind das direkte Informationen zu Hell-Dunkel-Unterschieden (HAAR-Cascades) oder aber Histogramme, die Hell-Dunkel-Unterschiede beschreiben. [Haar 2017] Findet der Algorithmus in einem Bild eine gewisse Anzahl an Übereinstimmungen

mit den Basis-Features, so wird das Bild als „TRUE“ klassifiziert.

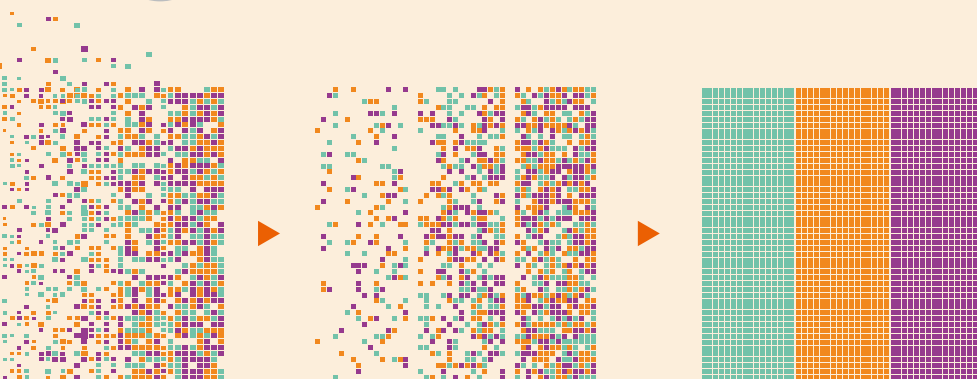
Zur Klassifizierung der gefundenen Features werden meist Support Vector Machines (SVM) eingesetzt. OpenCV (Open Computervision) ist eine bekannte Bibliothek, die viele solcher Ob-

jekterkennungsverfahren zusammenfasst. Es gibt noch weitere Ableger auf dem Open-Source-Markt wie das Open Biometrics. Damit kann beispielsweise eine Gesichtserkennung ohne großen Aufwand implementiert werden und darüber hinaus auch eine Gesichtsid-

Big Data im Fokus auf it-daily.net

- o Wie gelingt die Migration bestehender Systeme auf Big Data
- o Warum Demoskopien falsch liegen

Zu jeder Ausgabe neue Fokus Themen im Blick.



		Beschreibe mir was auf dem Bild zu sehen ist
	Zeige mir alle Gesichter auf dem Bild	Output
Zeige mir alle farblich ähnlichen Bilder	Output	Abgleich mit den „erlerten“ Features
Output	Abgleich mit den „erlerten“ Features	Verdichtung zu komplexen Features
Ein „hart verdrahtetes“ Programm	Vorgegebener Feature-Extraktionsalgorithmus für fest definierte Features	Einfache Features extrahieren
Input	Input	Input
Regelbasierte Systeme	Machine Learning	Deep Learning

Bild 2: Regelbasierte Systeme vs. Machine Learning vs. Deep Learning (nach [Boney 2017]).

tifikation. Hierbei muss jedoch bedacht werden, dass sich der Feature-Raum mit der Zunahme an zu vergleichenden Gesichtern proportional vergrößert, was sich negativ auf der Rechenzeit auswirkt.

Objektklassifizierung

Der anspruchsvollste Bereich in der Bildanalytik ist die generische Objektklassifizierung. Diese benötigt einen generischen Algorithmus, der robust genug ist, alle Merkmale eines Objekts wiederzuerkennen und zwar unabhängig vom Bildhintergrund, und das für eine beliebige Anzahl an Objekten.

Bild 2 zeigt Unterschiede zwischen regelbasierten Systemen, Machine Learning und Deep Learning. Die Domänen können folgendermaßen mit Bildanalytik in Zusammenhang gebracht werden:

- Die Berechnung von Hash-Werten, die ein Bild beschreiben oder die Farbstimmung in einem Hash oder

einem Histogramm kodieren, bewegt sich im Bereich der regelbasierten Systeme.

- Jedes Verfahren, das einen exklusiven Algorithmus zur Feature-Extraktion benutzt, um eine bestimmte Aufgabenstellung zu lösen, kann dem Bereich Machine Learning zugeordnet werden. Dazu gehört zum Beispiel das Erkennen von bestimmten Objekten (nur Autos, nur Fußgänger, nur Gesichter usw.).
- Beim Deep Learning kennt das System nicht die repräsentativen Features eines Objekts, sondern lernt nach und nach, diese Features in einem beliebigen Objekt selbst zu extrahieren. Mit der Zeit entwickeln solche Systeme ein robustes Modell und können damit beliebige Objekte selbstständig erkennen. Die Praxis zeigt, dass mit Hilfe von gefalteten neuronalen Netzwerken, die dem Bereich Deep Learning zugeordnet werden,

auch präzise Gesichtserkennungen und -identifizierungen realisiert werden können. [digitaltrends 2016]

Die Frage der Skalierbarkeit

Vielleicht denken Sie, dass all diese Bereiche mittlerweile zum Alltag gehören, auch auf einem Smartphone realisiert werden können und das Thema Skalierbarkeit deshalb nicht die erste Priorität haben kann. Anhand eines Beispiels möchten wir erklären, wann und warum Skalierbarkeit für ein solches Vorhaben dennoch wichtig wird:

Besonders im Bereich der Objekterkennung, wenn gefaltete neuronale Netzwerke (CNN) Anwendung finden, spielen Skalierbarkeit und Speichereffizienz eine sehr wichtige Rolle. Denn je nach Anzahl der zu trainierenden Objekte sowie je nach Anzahl an verfügbaren Bildern, die ein Objekt in mehreren Situationen und vor mehreren Hintergründen zeigen, ergibt sich die Anzahl an notwendigen Iterationen. Beispielsweise erfordert ein kompletter Durchlauf eines CNN mit einem Bild – je nach Architektur des CNN, je nach Anzahl der angewendeten Filter und je nach Bildgröße – Speicherplatz im dreistelligen MB-Bereich (z. B. etwa 140 MB an „Footprint“ pro Bild/Iteration bei vgg-NET als CNN mit einer Bildauflösung von 224x224 Pixel). Bei einer Anzahl von beispielsweise 300.000 Iterationen fällt diese Größe deutlich ins Gewicht. [CS231n 2017]

Skalierbarkeitsszenarien

- Eine Möglichkeit der Skalierung wäre eine vertikale Skalierung über die GPU. Dabei werden die Algorithmen über CUDA oder OpenCL auf die GPU portiert und dort ausgeführt. Dies bringt einen Performance-Zuwachs um einen Faktor, der im zweibis dreistelligen Bereich liegen kann.
- Eine weitere Möglichkeit, die Berechnungen zu beschleunigen, kann über eine horizontale Skalierung in einem Cluster realisiert werden. Dies ist etwas komplizierter, da eine Parallelisierung über CPU-Threads auf mehreren Cluster-Nodes nicht ohne weiteres möglich ist. Hierzu muss die Applikation dies entweder selbst be-

herrschen, oder sie muss an die bekannten Skalierungs-Frameworks adaptiert werden. Yahoo folgte diesem Ansatz und erweiterte das bekannte Deep-Learning-Framework Caffe so, dass es nun mit Apache Spark skaliert. Allerdings beschränkt sich Caffe auf gefaltete neuronale Netzwerke und die damit verbundenen Use Cases. [Yahoo 2017]

- Alternativ zu Spark kann auch mit Apache Storm eine horizontale Skalierbarkeit erreicht werden. So wird beispielsweise die Gesichtserkennung mittels OpenCV auf mehrere Bolts verteilt, wobei jedem dieser Bolts entsprechend starke Hardware zugewiesen wird. So skaliert jeder dieser Bolts vertikal und erreicht die horizontale Skalierung gleichzeitig mit einer richtig gruppierten (Grouping-)Storm-Topologie. Dabei sollte im Hinterkopf behalten werden, dass ein Bild nicht parallel von mehreren Bolts analysiert werden kann. So könnte für Echtzeitanalyse von Kameradaten der Ansatz des nächsten Punktes interessant sein. [StormCV 2017]
- Mit einem Bestandteil von Spatial and Graph Analytics von Oracle kann Bildanalytik massiv im Cluster betrieben werden. Oracle bietet mit Oracle Multimedia Analytics ein Framework für Bildanalysen mit OpenCV oder anderen Bibliotheken. Das Framework bringt eine auf MapReduce basierende Technologie mit, um Videos beispielsweise in ihre Einzelbilder zu zerlegen und darauf eine anschließende Gesichtserkennung mit OpenCV laufen zu lassen. Skalierung dieser Art eignet sich für Szenarien, die keine Echtzeitanalyse erfordern. [Oracle 2017]
- Noch eine Möglichkeit, die Berechnungen zu beschleunigen ist, diese in die Cloud auszulagern. Dabei sprechen wir nicht von der Migration eines der oben beschriebenen On-Premises-Ansätze in die Cloud, sondern über die Nutzung einer Cloud-API für Bildanalysen. Google, IBM und Microsoft bieten Computer Vision APIs an, mit deren Hilfe Bilder klassifiziert und in natürlicher Sprache beschrieben werden können. Dabei muss man klar unterscheiden zwischen einem generalistischen Ansatz, den die API-Betreiber in ihren Produkten abbilden und ei-



„Besonders im Bereich der Objekterkennung, wenn gefaltete neuronale Netzwerke (CNN) Anwendung finden, spielen Skalierbarkeit und Speichereffizienz eine sehr wichtige Rolle. Denn je nach Anzahl der zu trainierenden Objekte sowie je nach Anzahl an verfügbaren Bildern, die ein Objekt in mehreren Situationen und vor mehreren Hintergründen zeigen, ergibt sich die Anzahl an notwendigen Iterationen.“

Dimitri Gross,
Managing Consultant Solutions
Opitz Consulting Deutschland GmbH

nem Customized-Ansatz, der aktuell nur mit einem relativ hohen Entwicklungsaufwand realisiert werden kann.

Beispiel

Hier ein Beispiel, um den Unterschied deutlich zu machen: Geht es darum, die Objekte auf den Bildern zu erkennen und als strukturierten Text oder Beschreibung zurückzuerhalten, so eignen sich die Cloud-APIs ganz gut dafür. Auch wenn es darum geht, die Gesichter auf den Bildern zu finden und daraus gegebenenfalls noch Geschlecht und Alter abzuschätzen, kommen die Cloud-APIs in Frage, da man hier bei Hardware, Entwicklung und Betrieb ge-

genüber einer On-Premises-Lösung erhebliche Kosten sparen kann.

Geht es in einem Einsatzszenario allerdings darum, die Bilderdaten von Überwachungskameras zu analysieren und zum Beispiel auf den Bildern nach einer gewissen Person zu suchen, dann wird entweder eine Standardsoftware benötigt oder – falls die Lösung viele Individualanforderungen abdecken muss – eine eigenentwickelte Lösung, die je nach Use Case auch auf horizontale Skalierbarkeit ausgelegt ist.

Fazit

Bildanalytik zeigt eine starke Tendenz in Richtung Prozessautomatisierung. So gibt es bereits komplett neue Geschäftsbereiche im Bereich der Bildanalyse, die auf den Technologiesträngen Machine Learning und Deep Learning geschaffen wurden. Wie sich der Trend bei den Infrastrukturen weiter entwickelt, wird die Zukunft zeigen. Dass die Cloud-First-Strategie auch in diesem Bereich Einzug halten wird, lassen viele erfolgreiche Deep Learning Startups bereits jetzt erahnen.

DIMITRI GROSS

Quellen

- [digitaltrends 2017]
This scrappy Russian startup is beating Google and Facebook at facial recognition, <http://www.digitaltrends.com/cool-tech/ntechlab-facial-recognition-algorithm/> (abgerufen am 15.04.2017)
- [StormCV 2017]
Apache Storm + OpenCV = large scale distributed image and video analysis, <https://github.com/sensorstorm/StormCV>, (abgerufen am 15.04.2017)
- [Oracle 2017]
Big Data Spatial and Graph User's Guide and Reference, http://docs.oracle.com/cd/E65728_01/doc.43/e67958/GUID-4B15F058-BCE7-4A3C-A6B8-163DB2D4368B.htm#BDSPA-GUID-E9E5EAE8-0DB7-4BCA-B641-0A15A4E10771 (abgerufen am 15.04.2017)
- [Yahoo 2017]
Caffe on Spark – Github Project, <https://github.com/yahoo/CaffeOnSpark> (abgerufen am 15.04.2017)
- [CS231n 2017]
CS231n Convolutional Neural Networks for Visual Recognition, <http://cs231n.github.io/convolutional-networks/>, (abgerufen am 15.04.2017)
- [Boney 2017]
Theoretical Motivations for Deep Learning, <http://rinuboney.github.io/2015/10/18/theoretical-motivations-deep-learning.html>, (abgerufen am 15.04.2017)
- [Haar 2017]
Face Detection using Haar Cascades, http://docs.opencv.org/trunk/d7/d8b/tutorial_py_face_detection.html, (abgerufen am 15.04.2017)
- [LIRE 2017]
LIRE: Lucene Image Retrieval, <http://www.lire-project.net/>, (abgerufen am 15.04.2017)