

Maschinen sehen mehr, als wir glauben: Wie Computer Vision die Baustelle sichtbar macht.

Die bauma 2025 war die 34. Auflage der weltgrößten Fachmesse für Baumaschinen, Baustoffmaschinen, Bergbaumaschinen, Baufahrzeuge und Baugeräte. Sie fand vom 7. bis 13. April 2025 auf dem Messegelände in München statt. Die Messe setzte thematische Schwerpunkte in den Bereichen Klimaneutralität, alternative Antriebskonzepte, vernetztes und nachhaltiges Bauen sowie Digitalisierung und Innovationen im Mining-Bereich. Ich war vor Ort, habe verschiedene Ausstellungsbereiche besucht und mich insbesondere auf die Themen Digitalisierung und Künstliche Intelligenz konzentriert. Im Fokus stand dabei ein KI-Anwendungsfall, der mithilfe von Microsoft Azure AI die Ermittlung der Objektanzahl anhand einer Fotoaufnahme ermöglicht.



Markus Nussbaum

Die Bauherrenhilfe stellt allen Mitgliedern und Partnern einen „KI-Beauftragten“ zur Seite! Markus Nussbaum hat die HTL für Hochbau sowie Bauingenieurwesen studiert und bereits eine jahrelange Berufserfahrung hinter sich gebracht. Er ist zudem zertifizierter KI-Manager.

Kontakt:
m.nussbaum@bauherrenhilfe.org

Die bauma 2025 (7. bis 13. April 2025) ist vorbei und bot erneut spannende Einblicke in die Welt der Baumaschinen und Überraschung: Auch hier ist Künstliche Intelligenz (KI) allgegenwärtig. Von neu gedachten Muffenverbindungen für Bewehrungsseisen über optimierte Bohrköpfe bis hin zur KI-gestützten Datenerfassung zeigte sich deutlich: KI wird ein wesentlicher Bestandteil der Baustelle der Zukunft. Während meines Messebesuchs gab es für KI-Interessierte viel zu entdecken. Ein Beispiel sind die autonomen Baumaschinen von Liebherr, die kontinuierlich Leistungsdaten wie Energieverbrauch und Verschleiß erfassen, analysieren und optimieren. Ein weiteres Beispiel sind intelligente Überwachungssysteme, die mithilfe KI-gestützter Kameras Bewegungsmuster in Echtzeit analysieren. Die KI erkennt dabei Personen, Fahrzeuge und allgemeine Bewegungen in der Umgebung. Die moderne Baustelle, auf der ein Bauwerk schrittweise realisiert wird, bietet zahlreiche Dinge, die „gesehen“ werden können. Dazu zählt auch das entstehende Bauwerk selbst. Hier kann KI bereits während der Bauphase kontinuierlich den Fortschritt überwachen, den Ist-Zustand erfassen und diesen mit dem Soll, etwa den Projektplänen oder einem BIM-Modell, abgleichen.¹

Maschinen sehen

Nicht nur Menschen können auf der bauma etwas sehen, auch Maschinen sind mittlerweile in der Lage, ihre Umgebung zu

erfassen und menschliche Wahrnehmung zu simulieren. In vorherigen Beiträgen haben wir bereits einen Einblick in den Bereich der Computer Vision (CV) erhalten, ein Teilgebiet der Künstlichen Intelligenz, das Maschinen das „Sehen“ und anschließende „Verstehen“ ihrer Umgebung ermöglicht.

KI zählt Objekte auf der Baustelle

Fast schon versteckt befand sich auf dem Messestand der Doka Österreich GmbH (ein Anbieter von Schalungs- und Gerüstsystemen für den Betonbau) ein Bereich, der einer Computer-Vision-Anwendung gewidmet war. Im Gespräch mit einem Entwickler von Doka wurde mir ein Tool vorgestellt, das derzeit ausschließlich firmenintern eingesetzt wird, aber künftig auch als App für Bauunternehmen verfügbar sein soll. Dabei wird ein möglichst rechtwinklig aufgenommenes Foto, etwa von einem Bündel Bewehrungsstäbe oder einem Stapel Kanthölzer, analysiert. Das KI-System verarbeitet das Bild und gibt die Anzahl der erkennbaren Kanthölzer aus. In der Regel arbeiten verschiedene KI-Techniken zusammen, um eine solche Anwendung überhaupt zu ermöglichen. Erst das Zusammenspiel aller eingesetzten KI-Modelle und -Verfahren ergibt ein funktionierendes KI-System. Wenn ich also schreibe „die KI erkennt“ oder „die KI macht“, ist damit immer ein komplexes KI-System gemeint, das unterschiedliche Technologien integriert. Ein zentraler Bestandteil dieses Systems ist die

¹ Künstliche Intelligenz erleichtert Mitarbeitenden den Alltag: <https://bauma-messe.de/artikel/kuenstliche-intelligenz-erleichtert-mitarbeitenden-den-alltag-001844>, Zugriff 28.04.2025.

Objekterkennung auf dem jeweiligen Bild, die durch ein spezielles Computer-Vision-Modell erfolgt, beispielsweise das Modell „YOLO“ (You Only Look Once). Dabei kommen sogenannte Convolutional Neural Networks (CNNs) zum Einsatz, eine Form neuronaler Netzwerke, die speziell für die Analyse von Bilddaten entwickelt wurden.



Abbildung 1: Ein Bild von einem Stapel Kanthölzern. Was erkennt hier eine „KI“?

Abbildung 1 zeigt einen Stapel gelagerter Kanthölzer. In der Objekterkennung wird erfasst, wo sich die Objekte im Bild befinden und was sie sind, in diesem Fall unsere Kanthölzer.

Das Prinzip dahinter

Das Computer-Vision-Modell „YOLO“ erhält das gesamte Bild und beginnt unmittelbar damit, Merkmale zu extrahieren, die darauf hinweisen, dass es sich mit gewisser Wahrscheinlichkeit um einzelne Kanthölzer handelt. Jedes erkannte Kantholz wird dabei mit einer rechteckigen Umrahmung versehen, einer sogenannten Bounding Box, und als „Kantholz“ klassifiziert. Der Vorgang ähnelt dem menschlichen Sehen: Alle Objekte in unserer Umgebung weisen charakteristische Anordnungen von Kanten und Ecken auf, die in ihrer Gesamtheit eine Fläche und letztlich ein Objekt formen, in diesem Fall die typische, eher quadratische Form eines Kantholzes. Wenn wir auf den Querschnitt eines unbekannten Objekts blicken,

klassifizieren wir es intuitiv: „Ah, das ist ein Kantholz!“ Damit diese Einordnung funktioniert, muss zuvor definiert werden, dass ein solches Holzstück, basierend auf seiner Form, als Kantholz gilt. Ein Brett oder eine Bohle hingegen weisen einen eher rechteckigen Querschnitt auf, bei dem eine Kante deutlich länger ist als die andere. Wir haben also festgelegt (zum Beispiel in einer Norm), dass solche Formen als Bretter oder Bohlen bezeichnet werden (wobei Bohlen dicker sind als Bretter). Kanthölzer, Bohlen und Bretter sind alle eine Art von Schnittholz, die sich in ihren jeweiligen geometrischen Merkmalen unterscheiden. Abbildung 2 stellt das Prinzip der Objekterkennung schematisch dar.

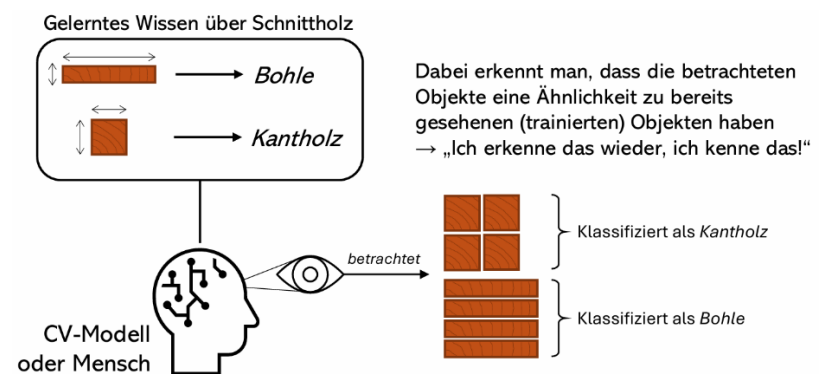


Abbildung 2: Schema der Objekterkennung und Klassifizierung durch gelerntes Wissen analog zum Menschen

Menschen und Maschinen erkennen unterschiedliche Merkmale, indem sie lernen, diese zu unterscheiden. Bei Maschinen geschieht das durch maschinelles Lernen, beispielsweise durch überwachtes Lernen, also das gezielte Training einer Künstlichen Intelligenz. Vereinfacht ausgedrückt muss der Maschine im Vorfeld eine große Anzahl von Bildern gezeigt werden, auf denen Kanthölzer und Bohlen eindeutig gekennzeichnet (gelabelt) sind, damit sie lernt, zwischen diesen beiden Arten von Schnittholz zu unterscheiden. Wenn das Computer-Vision-Modell ein Objekt korrekt klassifiziert und mithilfe einer Bounding Box lokalisiert (also markiert), muss im Grunde nur noch die Anzahl dieser Bounding Boxes erfasst werden, um die Stückzahl der erkannten Objekte zu bestimmen.

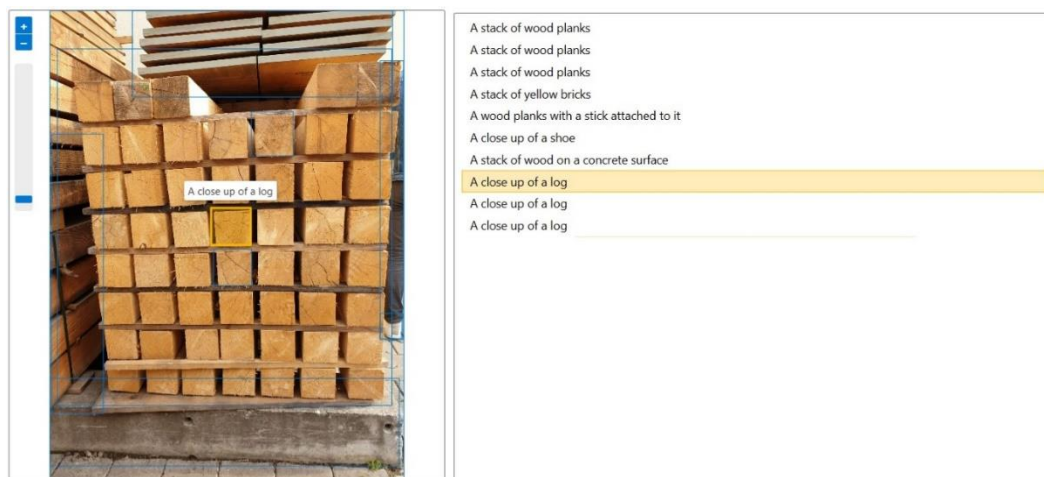


Abbildung 3: Bild von Kanthölzern und Objekterkennung mittels Vision Studio in Microsoft Azure AI

Wie in Abbildung 3 zu sehen ist, habe ich ein eigenes Bild in das Objekterkennungstool von Azure AI eingefügt, um einen Eindruck davon zu gewinnen, wie die KI, konkret das zugrunde liegende Computer-Vision-Modell, Objekte im Bild klassifiziert. Tatsächlich kann das Modell in diesem Fall zwischen Brettern (im Bild als „planks“ bezeichnet) und Kanthölzern („logs“) unterscheiden. In diesem einfachen Beispiel ist die Objekterkennung jedoch nicht in der Lage, alle einzelnen Kanthölzer zuverlässig zu erkennen, was auch nicht das Ziel dieses Tests war. Die zuvor erwähnten Bounding Boxes sind im Bild teilweise sichtbar. Es stellt sich dabei die Frage, wie zuverlässig eine KI-gestützte Objekterkennung und -zählung ist, wenn die Objekte nicht in geordneter Struktur vorliegen, etwa bei einem unstrukturierten Haufen von Stäben. Doka setzt bei der Entwicklung ihrer App-Anwendung auf Microsoft Azure AI. Dabei handelt es sich um einen cloudbasierten Dienst von Microsoft, der es Unternehmen und Entwicklern ermöglicht, Künstliche Intelligenz in ihre Anwendungen und Prozesse zu integrieren. Allerdings ist die Nutzung von Azure AI mit teilweise erheblichen Kosten verbunden. Die Bezahlung erfolgt für die tatsächliche Nutzung von externen „Serverressourcen“ von Microsoft.

Zwischen Make und Buy

Im Gespräch mit dem Entwickler von Doka kamen immer wieder Vertreter von Baufirmen und Handwerksbetrieben vorbei, um sich die App-Anwendung selbst anzusehen. Der Bedarf auf Seiten der Bauunternehmen scheint vorhanden zu sein. Besonders das

Baustellenpersonal zeigte großes Interesse daran, mithilfe einer KI-Anwendung die Stückzahlen von Objekten zu ermitteln. Mit Azure AI bezieht man grundlegende KI-Dienstleistungen von einem externen Anbieter und kann darauf aufbauend individuelle Lösungen entwickeln. Die cloudbasierte Plattform ermöglicht eine deutlich schnellere Implementierung, ohne dass tiefgehende KI-Expertise oder eigene Hardware erforderlich ist, beides würde sonst mit hohen Investitionskosten einhergehen. Azure AI stellt für jeden notwendigen Schritt vorgefertigte Services mit benutzerfreundlicher Oberfläche zur Verfügung, sodass alle benötigten Entwickler-Tools zentral von einem Anbieter bereitgestellt werden. Natürlich macht man sich dadurch in gewissem Maße vom externen Anbieter, in diesem Fall Microsoft, abhängig.

Fazit

Letztlich hängt die Entscheidung zwischen „Make or Buy“ stark davon ab, wie spezifisch die eigene KI-Anwendung sein soll und ob es bereits ein erprobtes, am Markt verfügbares Tool gibt, das den Anforderungen genügt und keine Neuentwicklung erfordert. In den meisten Fällen wird die Lösung irgendwo zwischen Eigenentwicklung und Fremdbezug von KI-Kompetenz und Infrastruktur liegen. Azure AI positioniert sich in der Mitte: Es ermöglicht, bereits vortrainierte Modelle mit eigenen Bilddaten nachzutrainieren, um sie an individuelle Anwendungsfälle anzupassen.