

3D-Vision kombiniert mit Deep Learning-Software für die automatisierte Früchteerkennung

Basler hat zusammen mit der Firma Data Spree aus Berlin eine Deep Learning-Anwendung zur Erkennung und Klassifikation von Früchten entwickelt. Das Vision System basiert auf einer Basler blaze Time-of-Flight-Kamera und einer Deep Learning-Plattform von Data Spree.

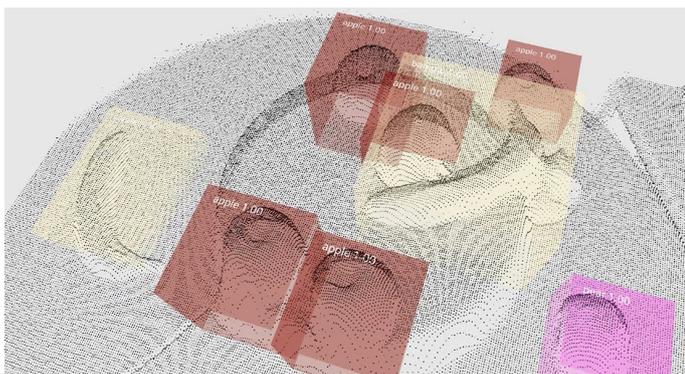
1. Überblick

Deep Learning-Ansätze bieten nicht nur durch eine erhöhte Genauigkeit Vorteile gegenüber klassischen Verfahren der Bildverarbeitung, sondern auch durch die reduzierte Entwicklungszeit von einsatzbereiten Systemen. Maschinen, die auf einem gut trainierten neuronalen Netz basieren, sind nicht nur robust gegenüber Varianten von zu erkennenden Gegenständen, sondern können auch kontinuierlich über den kompletten Lebenszyklus der Systeme verbessert werden und somit unter geänderten Randbedingungen gleichbleibende Ergebnisse erbringen.

Ein sehr gutes Beispiel hierfür ist die Sortierung und Verarbeitung von landwirtschaftlichen Erzeugnissen. Diese können sich in Form und Farbe stark voneinander unterscheiden, was klassische Bildverarbeitungsmethoden vor große Herausforderungen stellt.

2. Lösung

Basler hat in enger Zusammenarbeit mit dem Software-Anbieter Data Spree eine Vision-Lösung für die Detektion und Klassifikation von Früchten entwickelt, indem 3D-Daten als Rohmaterial für die Objekterkennung genutzt werden. Die Stärken der 3D-Time-of-Flight-Technologie werden mit einem sehr einfach anwendbaren Deep Learning-Algorithmus kombiniert und bietet so eine hochgenaue und robuste Echtzeitleistung für die Früchte-sortierung.



3D-Punktwolkenbild der Früchte

2.1 Hardware

Das Deep Learning-basierte Vision System besteht aus einem Standard PC und einer Basler blaze ToF-Kamera.

Die Basler blaze Kamera bietet hochauflösende 3D-Bilder mit annähernd millimetergenauer Präzision dank neuester Sony DepthSense™ IMX556PLR Sensortechnologie. Es handelt sich um eine 3D-Kamera, die über die Time-of-Flight (ToF) Methode nicht nur eine Graustufenaufnahme als Intensitätsbild erzeugt, sondern zusätzlich über Laufzeitmessungen von Lichtimpulsen im nahen Infrarotbereich (NIR), Distanzmessungen für jeden einzelnen Pixel vornimmt. Die resultierende Aufnahme liegt dann als 3D-Punktwolke vor und liefert zusätzliche Informationen über die abgebildete Szene. Im Vergleich zu 2D-RGB-Aufnahmen werden hier die Farbinformationen durch Forminformationen ersetzt, welches nicht nur Vorteile bei der gleichzeitigen Erkennung von roten und grünen Äpfeln hat, sondern zusätzliche Applikationen ermöglicht, wie z.B. das genaue Positionieren und Vermessen der erkannten Objekte.



Basler blaze Kamera

2.2 Software

Die von der Basler blaze Kamera aufgenommenen Bilder werden von der Data Spree Software „Deep Learning DS“ ausgewertet, eine Softwarelösung auf Basis tiefer neuronaler Netze (Deep Learning). Die Data Spree Lösung ist extrem anwenderfreundlich und erlaubt die einfache Entwicklung von Deep Learning-Modellen ohne Vorkenntnisse.

Für den Aufbau sind folgende Arbeitsschritte erforderlich:

1. Datenakquise

Jedes neuronale Netz benötigt eine Fülle an Beispielbildern für die spätere Klassifikation. So werden zunächst ca. 500 Bilder von Früchten (Bananen, Äpfeln und Birnen) mit der Basler blaze Kamera aufgenommen. Die Kamerasoftware erstellt sodann 2-Kanal-Bilddaten aus der Graustufen-Intensitätsaufnahme und dem Tiefenbild, welches für jeden Pixel den Abstand zum Objekt in Millimetern enthält. Diese Bilddaten werden in die Deep Learning DS Plattform geladen.

2. Annotation

Anschließend werden die Bilddaten mit Metadaten angereichert. Dazu werden manuell Boxen um die Früchte gezeichnet und die entsprechende Kategorie (Apfel, Birne etc.) wird zugeordnet. Damit wird festgelegt, was dem neuronalen Netz im folgenden Schritt „beigebracht“ wird. Dieser zunächst mühselige Prozess wird bereits nach circa 100 manuell annotierten Bildern beschleunigt, indem ein initiales Deep Learning-Modell bereits Vorschläge für die folgenden Aufnahmen generiert, die anschließend nur noch korrigiert werden müssen.

3. Training

Sind alle 500 Bilder annotiert, wird über wenige Mausklicks ein weiteres Modell erstellt und das Deep Neural Network (DNN) wird automatisiert trainiert. Während des Trainings wird zyklisch die Erkennungsgenauigkeit anhand eines Testdatensatzes evaluiert.

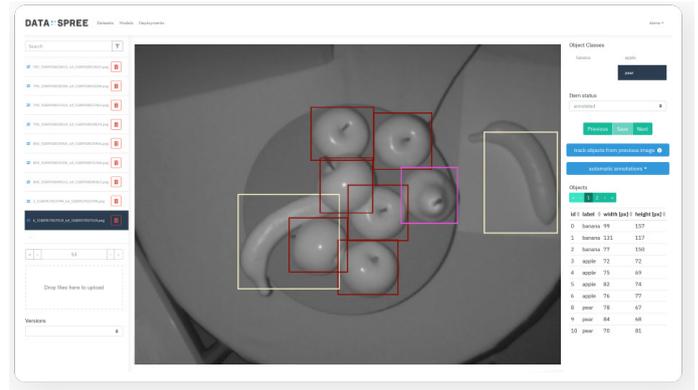
4. Bereitstellung (Deployment) und Anwendung (Inferenz) des trainierten Netzes auf der Zielhardware

Sobald das Training abgeschlossen ist, wird das fertig trainierte Modell auf die Zielhardware geladen, wo es über die Data Spree Ausführungssoftware Inference DS direkt ausgeführt werden kann. Die Basler blaze ToF-Kamera inklusive entsprechender Vorverarbeitung ist bereits vollständig integriert, sodass die Deep Learning-Applikation direkt gestartet werden kann.

2.3 Vorteile der Lösung

Die Lösung bietet verschiedene Vorteile:

- Genauere und zuverlässigere Erkennung und Klassifikation von Objekten durch Einbindung der räumlichen Information mittels einer 3D-Kamera (Time-of-Flight) in das Anlernen der neuronalen Netze



Annotation: Es wird festgelegt, was dem neuronalen Netz im folgenden Schritt „beigebracht“ wird.

- Geringere Komplexität der Anwendung, da eine komplementäre Sensorik in einer Vielzahl von Applikationen nicht mehr notwendig ist
- Dank der einfachen und komfortablen Software-Lösung von Data Spree können Deep Learning-Anwendungen auch ohne Programmier- und Deep Learning-Vorkenntnisse erstellt werden.
- Genaue Messergebnisse auch bei wenig Licht, bei Tageslicht und ohne Kontrast
- Einsatz von industriebewährter und langlebiger Kamera-Hardware mit IP67 Schutzklasse
- Einfache Hardwareinstallation durch integrierte Beleuchtung und kalibrierte Optik

3. Fazit

Mit der dargestellten Beispielanwendung zur Erkennung von Früchten beweist Basler in Kooperation mit Data Spree, dass die Kombination einer robusten 3D Kamera mit Deep Learning-Modellen es erlaubt, komplexe Aufgaben zeit- und kosteneffizient zu lösen. Die zusätzliche Tiefeninformation der Time-of-Flight-Kamera blaze von Basler kann bei einer Vielzahl von Anwendungen helfen, wenn Volumen, Formen oder die Position von Objekten analysiert werden sollen.

Basler ist der richtige Partner für die Entwicklung von intelligenten 3D-Vision Lösungen, die robuste Hardware mit anwenderfreundlicher Software vereinen.