Technische Hochschule
Brandenburg
University of
Applied Sciences
Fachbereich
Informatik und Medien

Instanzsegementierung auf Edge AI-Plattformen: Integration von Basler Hochgeschwindigkeitskameras und Yolo-Seg

Lasse Broer

Bachelorarbeit • Studiengang Medieninformatik • Fachbereich Informatik und Medien • 26.03.2025

Aufgabenstellung

Ziel dieser Arbeit ist die Demonstration der Leistungsfähigkeit aktueller Instanzsegmentierungsverfahren auf der Edge-AI-Plattform NVIDIA Jetson AGX Orin. Im Fokus steht das einstufige, auf Convolutional Neural Networks (CNN) basierende Modell YOLO-Seg in Kombination mit Hochgeschwindigkeitskameras der Firma Basler. Es wurde untersucht, wie sich die eingesetzten Komponenten optimal konfigurieren lassen, um eine robuste und zuverlässige Instanzsegmentierung auch unter Bedingungen hoher Bildraten sicherzustellen.

Konzept

Zur Erreichung des formulierten Ziels wurden zunächst alle relevanten Soft- und Hardwarekomponenten installiert und umfassend evaluiert. Aufbauend auf den gewonnenen Erkenntnissen wurde ein Demonstrator konzipiert und realisiert, der die praktische Anwendbarkeit sowie die Effektivität verschiedener Konfigurationsmöglichkeiten unter realistischen Einsatzbedingungen demonstriert.

Evaluation der Leistungsfähigkeit von YOLO-Seg

Für die systematische Bewertung von YOLO-Seg auf dem Jetson AGX Orin wurde ein dediziertes Testframework entwickelt. Dieses ermöglicht die parametrisierte Ausführung des Modells und erfasst währenddessen relevante Leistungskennzahlen wie GPU-Auslastung, Stromverbrauch und Inferenzzeiten. In 10 Experimenten wurde untersucht, wie sich unterschiedliche Modellgrößen (Nano, Small) und Präzisionsmodi (FP32, FP16, INT8) auf die Leistungsfähigkeit, Genauigkeit und den Ressourcenverbrauch des Modells auswirken.

Evaluation der Basler-Kamera

Neben der Inbetriebnahme der Kamera wurde die maximal erreichbare Bildrate unter Variation zentraler Parameter wie Pixel-Format und Region-of-Interest (ROI) ermittelt.

Demonstrator

Zur praxisnahen Veranschaulichung wurde ein einfacher Versuchsaufbau realisiert: Eine Rampe mit einer Neigung von 45° diente als Testumgebung, wobei die Kamera in entsprechendem Winkel oberhalb der Rampe installiert wurde. (Abb. 1) Die softwareseitige Umsetzung erfolgte unter Verwendung von NVIDIA DeepStream – einem leistungsfähigen Toolkit zur Entwicklung und Ausführung von Stream-Processing-Pipelines für KI-basierte Anwendungen.

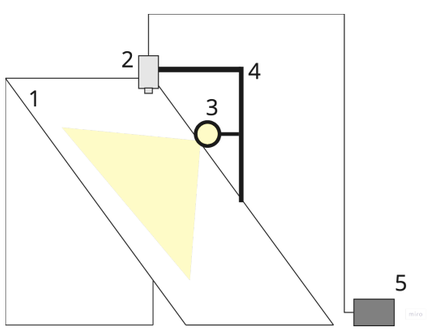

Abb. 1: Versuchsaufbau des Demonstrators mit Rampe (1), Kamera (2), Licht (4), Stativ (4) und Jetson AGX Orin (5)

Test des Demonstrators

Basierend auf den Erkenntnissen der Einzelkomponenten-Evaluation wurden zentrale Parameter identifiziert, die maßgeblichen Einfluss auf die Erkennungsgenauigkeit und Verarbeitungsgeschwindigkeit des Gesamtsystems haben:

- Präzision: FP-16, FP-32
- Modellvariante: Nano, Small
- Bildauflösung: UXGA (1440 × 900), WUXGA (1920 × 1200)
- Farbformat: RGB, YUV

Zur Bestimmung der optimalen Kombination dieser Parameter wurden sämtliche Permutationen in einem standardisierten Testablauf erprobt. Insgesamt wurden 16 Experimente durchgeführt. Während der Tests wurde die Systemlatenz erfasst. Zudem wurden Erkennungsgenauigkeit und Maskenqualität subjektiv anhand einer Skala von 1 (sehr schlecht) bis 5 (sehr gut) bewertet.

Als Testobjekte dienten zwei handelsübliche Haushaltsscheren, die in geöffnetem Zustand die Rampe hinunterrutschten und dabei von der Kamera erfasst wurden. (Abb. 2)

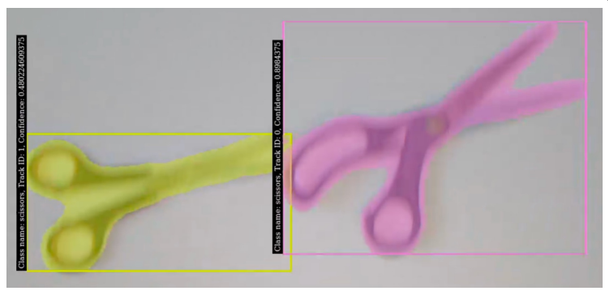

Abb. 2: Auschnitt eines Testdurchlaufs

Ergebnisse

Die Testergebnisse bestätigten insgesamt die zuverlässige Funktionalität der implementierten DeepStream-Pipeline: Die Objektinstanzen wurden konsistent erkannt, präzise nachverfolgt und die Segmentierungsmasken korrekt dargestellt.

Je nach gewählter Konfiguration zeigten sich jedoch feine Qualitäts- und deutliche Geschwindigkeitsunterschiede, die differenzierte Rückschlüsse auf die Wirksamkeit der eingesetzten Parameter und deren Einfluss auf das Gesamtsystem ermöglichten.

Besonders hervorzuheben ist, dass in der Hälfte der durchgeführten Experimente die Segmentierung selbst bei sehr hohen Bildraten von bis zu 198 Hz erfolgreich und mit subjektiv guter Qualität in Echtzeit ausgeführt werden konnte.

Fazit

In dieser Arbeit wurde gezeigt, dass sich die gewählten Komponenten grundsätzlich für den Einsatz in Hochgeschwindigkeitsanwendungen eignen. Allerdings erfordert die Konfiguration der Komponenten eine sorgfältige Abstimmung auf das jeweilige Anwendungsszenario sowie auf die spezifischen Anforderungen an Modellgenauigkeit und Verarbeitungsgeschwindigkeit.

Betreuung Dipl.-Inform. Ingo Boersch • Technische Hochschule Brandenburg • Prof. Dr. Emanuel Kitzelmann • Technische Hochschule Brandenburg