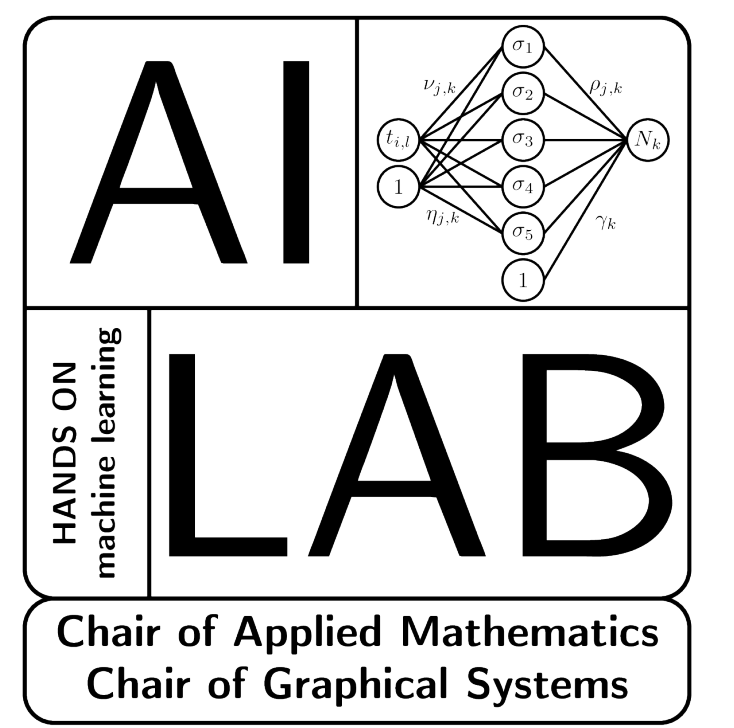


YOLO: ECHTZEIT-OBJEKTERKENNUNG IN EINER LERNFABRIK

Dr. Toni Schneidereit

Lehrlabor für Künstliche Intelligenz

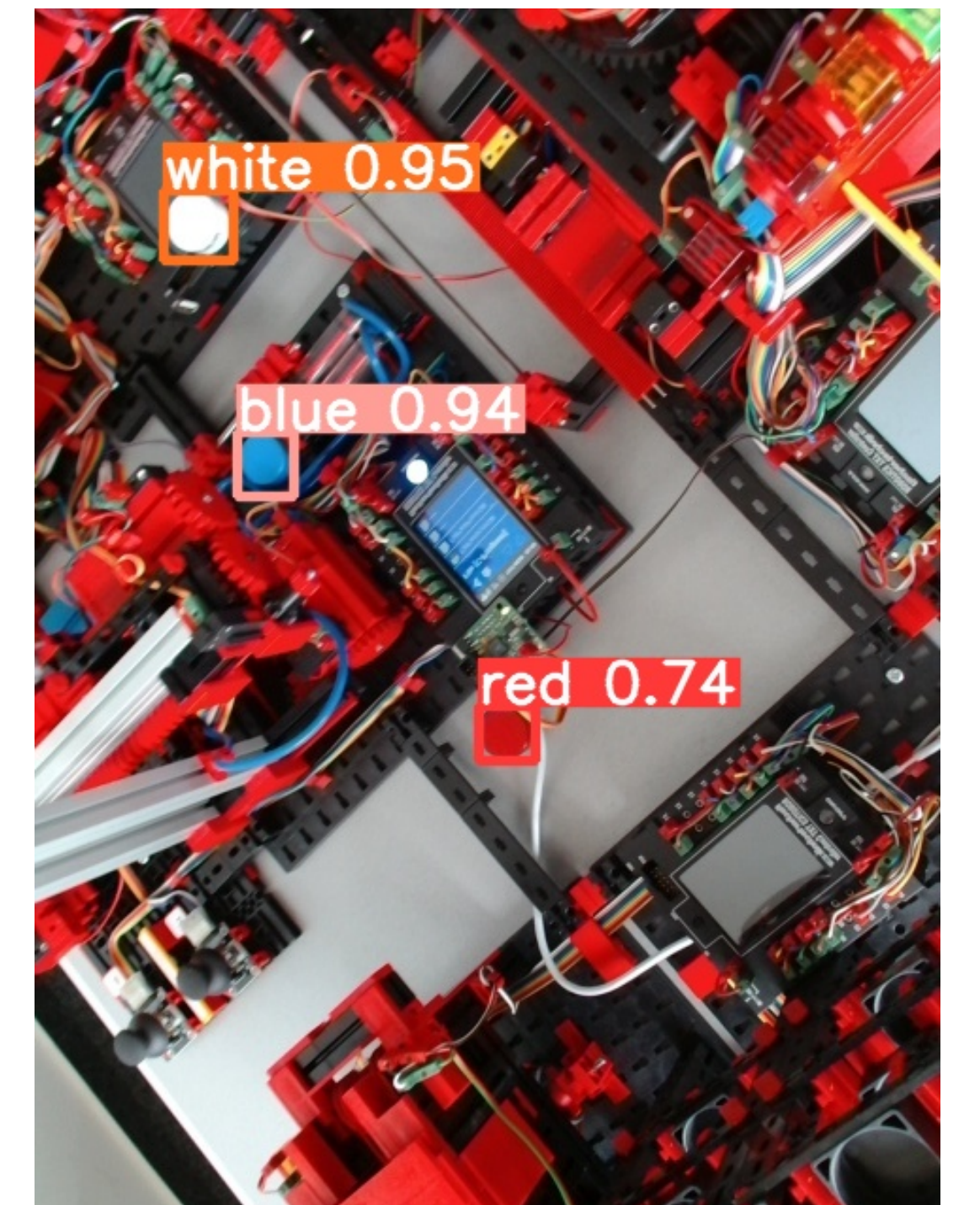
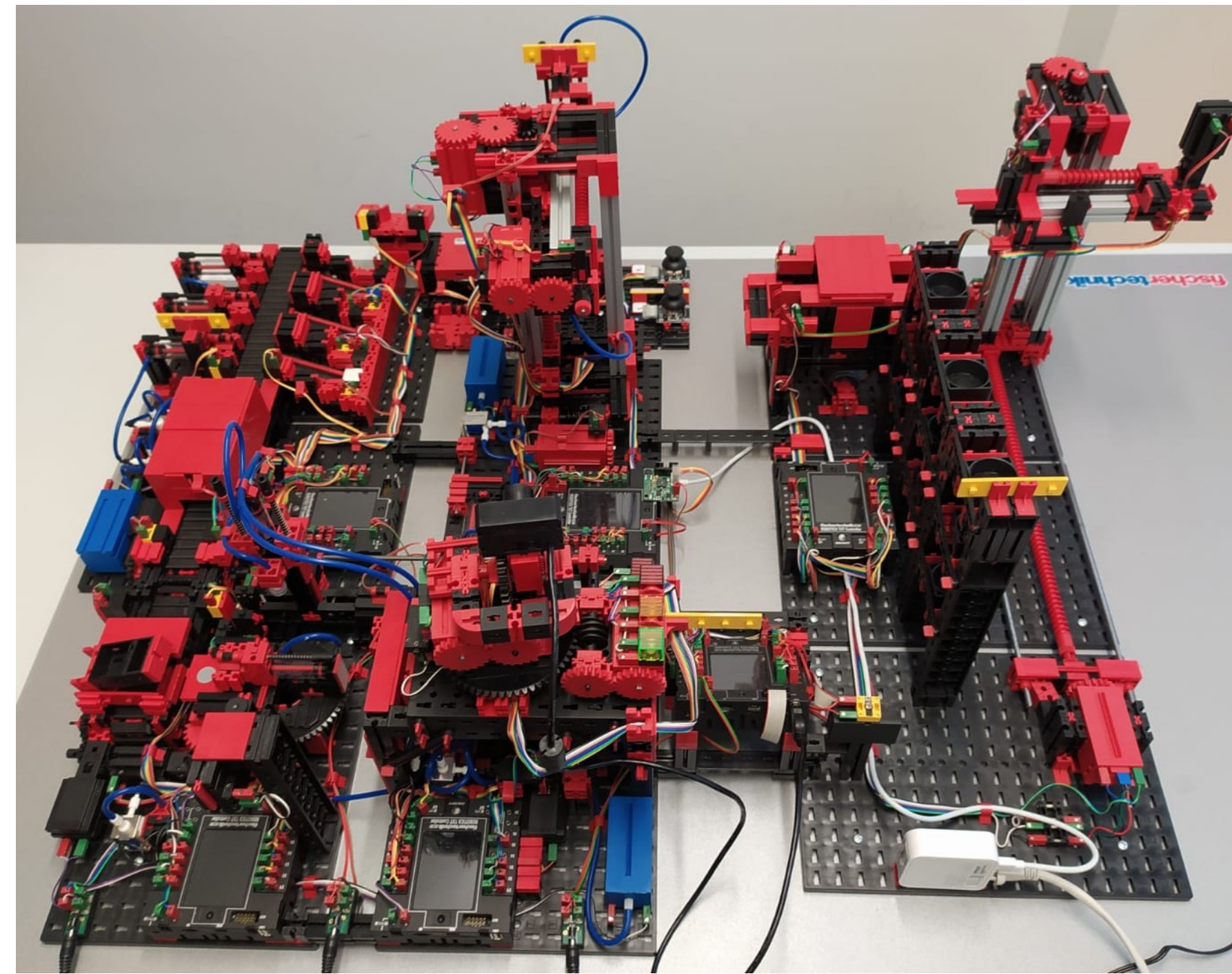
Angewandte Mathematik // Graphische Systeme
Brandenburgische Technische Universität Cottbus-Senftenberg



Motivation

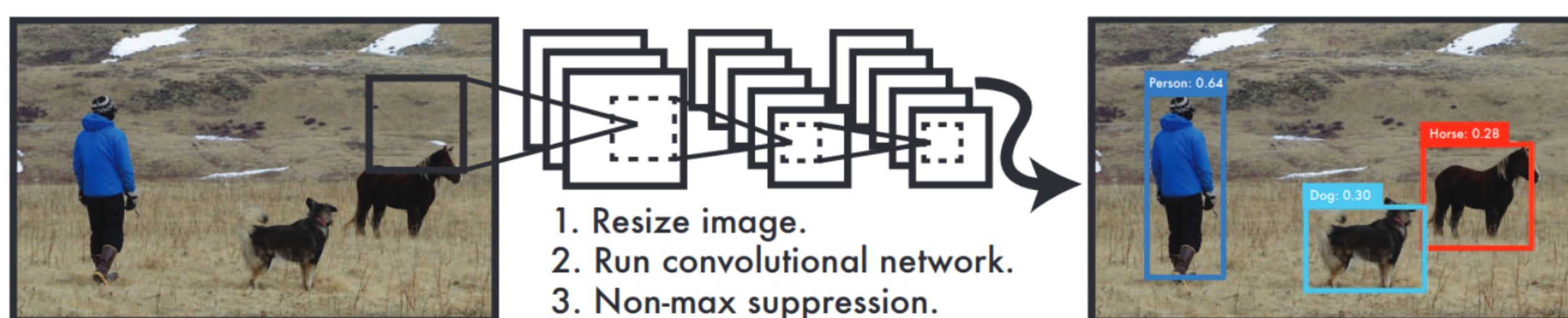
Zielsetzung und Lernfabrik

- Anwendung und spezifische Untersuchung von KI-basierter Echtzeit-Objekterkennung in einer Lernfabrikumgebung
- Lernfabrik: Simulation von Industrieprozessen im Miniaturmaßstab
- Erkennung von farbigen Werkstücken (zylindrische Kreisscheiben)
- Grundlage für weitere Forschung im Bereich (z.B.):
 - > Objekterkennungsbasierte Prozesssteuerung
 - > Prädiktive Instandhaltung

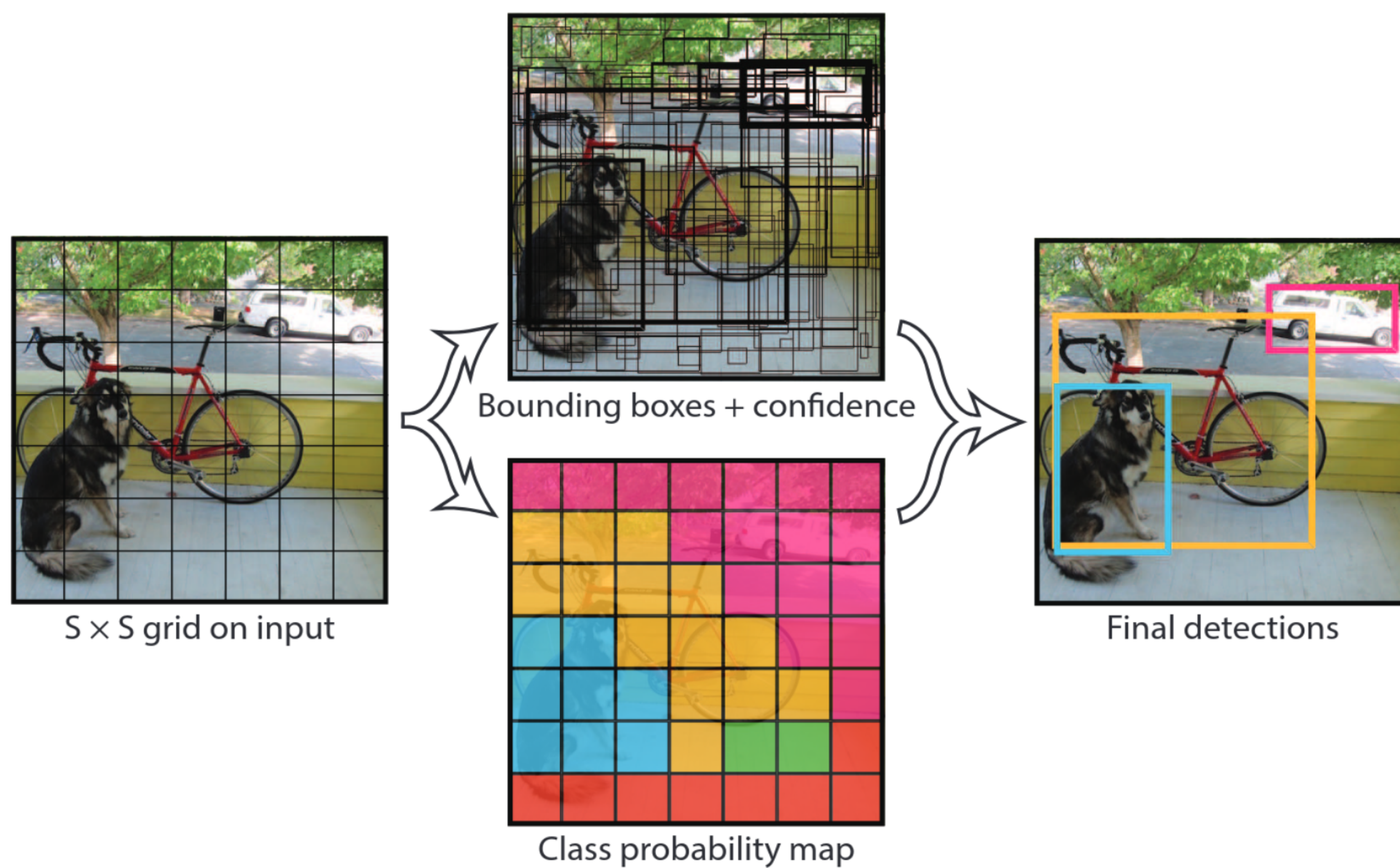


YOLO: You Only Look Once!

Grundidee der Funktionsweise: Das Eingabebild wird auf eine geeignete Größe skaliert, durch ein Convolutional Neural Network verarbeitet und die Ergebnisse mit den größten Wahrscheinlichkeiten werden ausgegeben



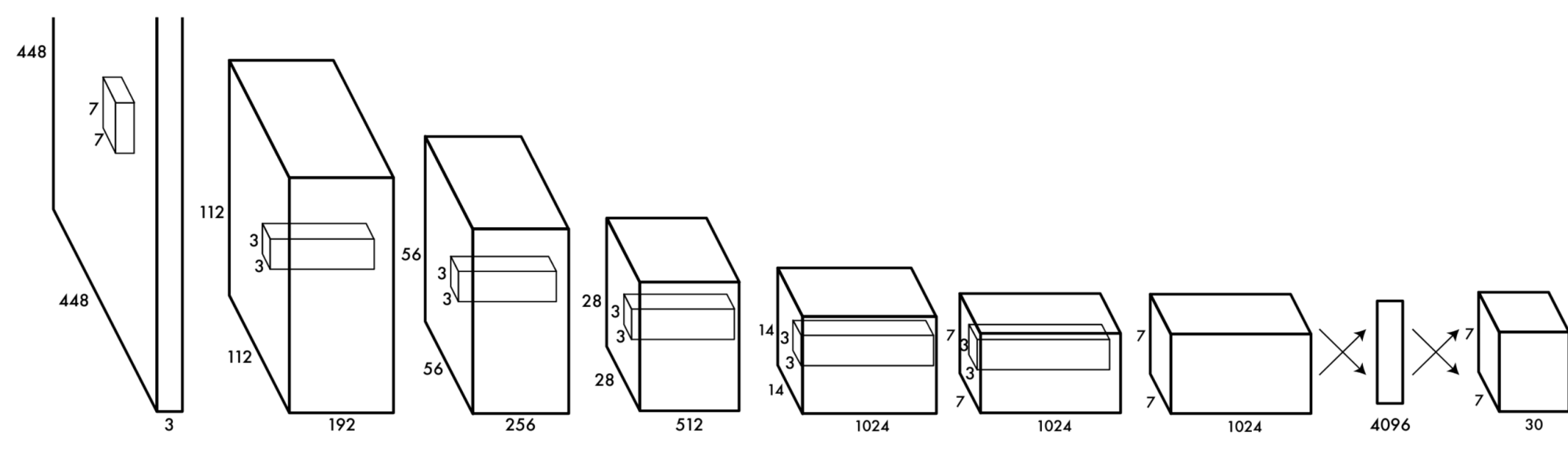
Lokalisierung und Klassifizierung: Jede Gitterzelle trifft eine Vorhersage über mehrere *bounding boxes* und alle *Klassen*, die sich möglicherweise in der Zelle befinden



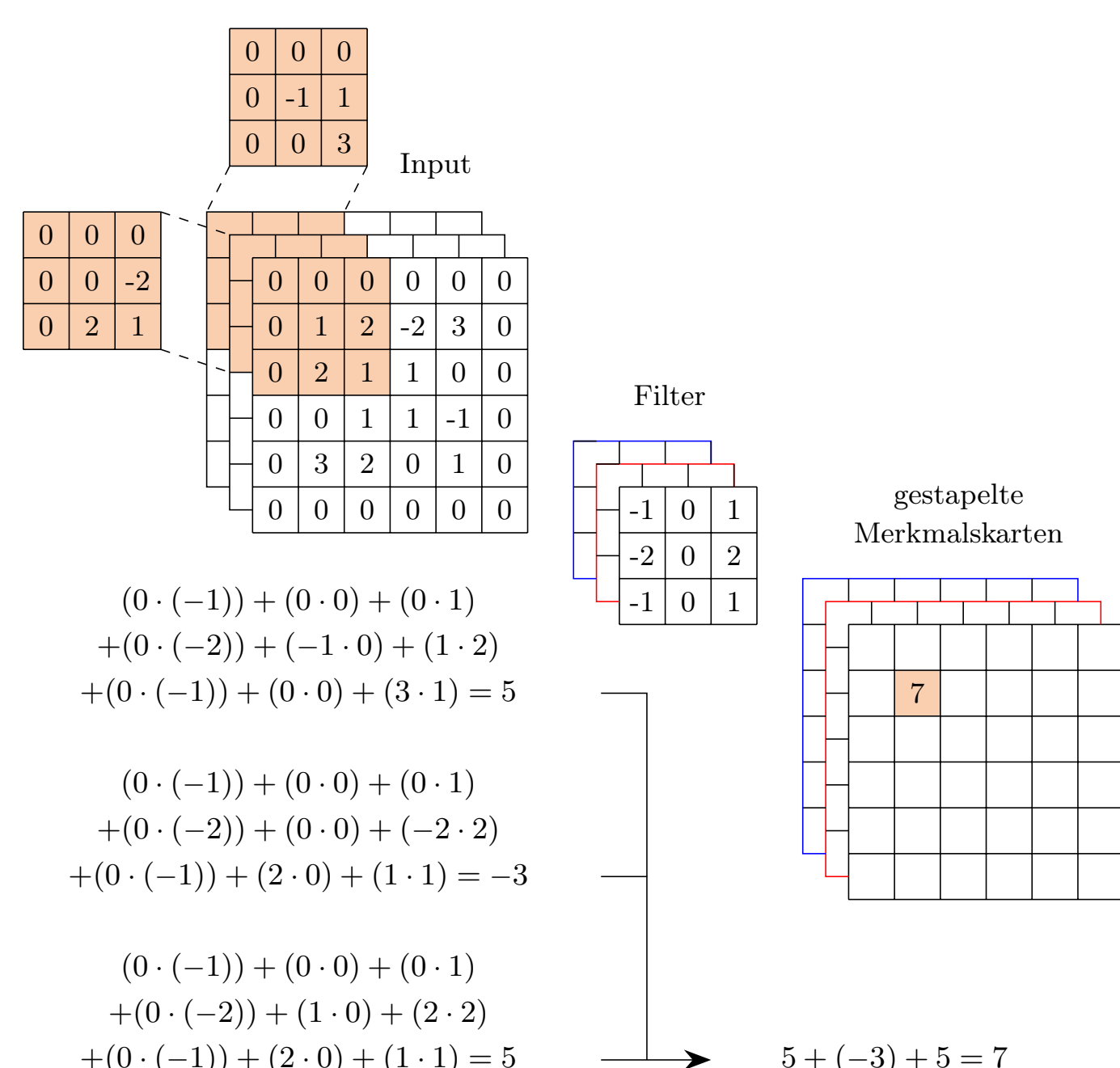
Vorhersage: Lokalisierung und Klassifizierung werden separat vorhergesagt und am Ende zusammengefügt. Je öfter eine bounding box vorhergesagt wird, desto größer ist die Sicherheit dort ein Objekt zu finden. Die Klasse des Objekts wird vom Neuronalen Netz gesondert vorhergesagt und am Ende mit den bounding boxes verbunden

YOLO-Architektur und Convolution

Architektur von YOLOv1:



- Farbbild ist eine dreidimensionale Matrix
- Filter laufen über die Matrix und "extrahieren" Merkmale
- Pooling-Schichten komprimieren die extrahierten Informationen
- Filtereinträge sind veränderlich und werden gelernt

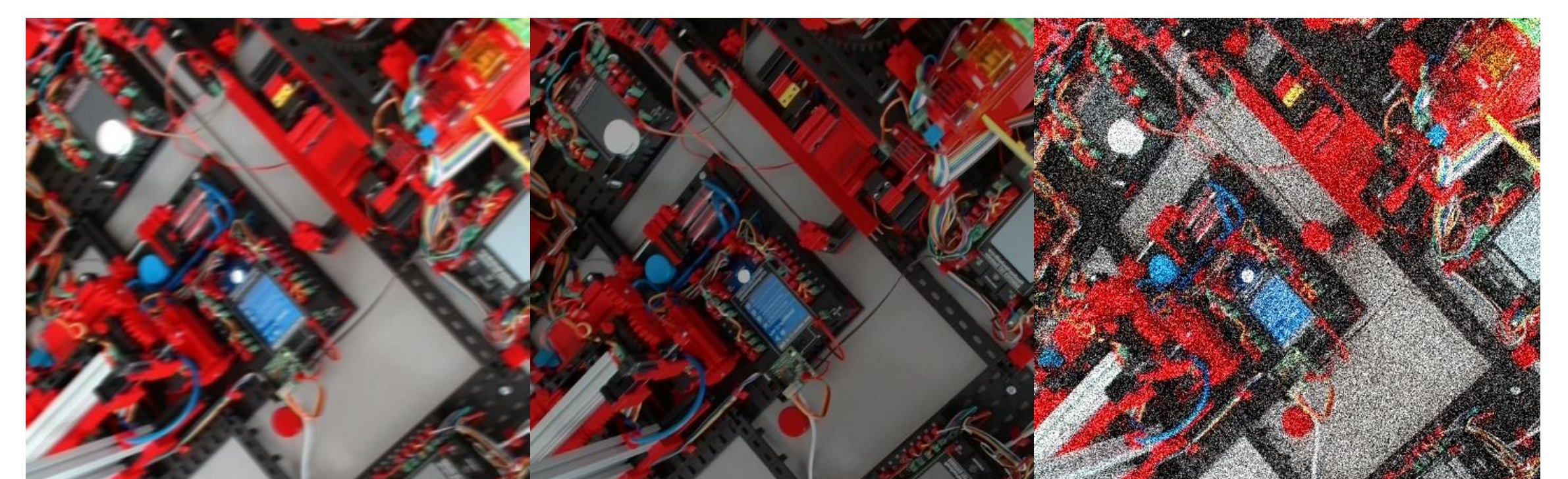


Datensatzgenerierung

Datengewinnung: Aufnahme von 700 Trainingsbildern der Werkstücke in der Lernfabrik, jedes Werkstück ist auf 400 Bildern zu sehen und gelabelt

Datenerweiterung: Simulation verschiedener äußerer Einflüsse und Vergrößerung der Zahl von Trainingsbildern in Bezug auf:

- Helligkeit, Kontrast, Rauschen
- Schärfe, Gauß'sche Unschärfe, Bewegungsunschärfe

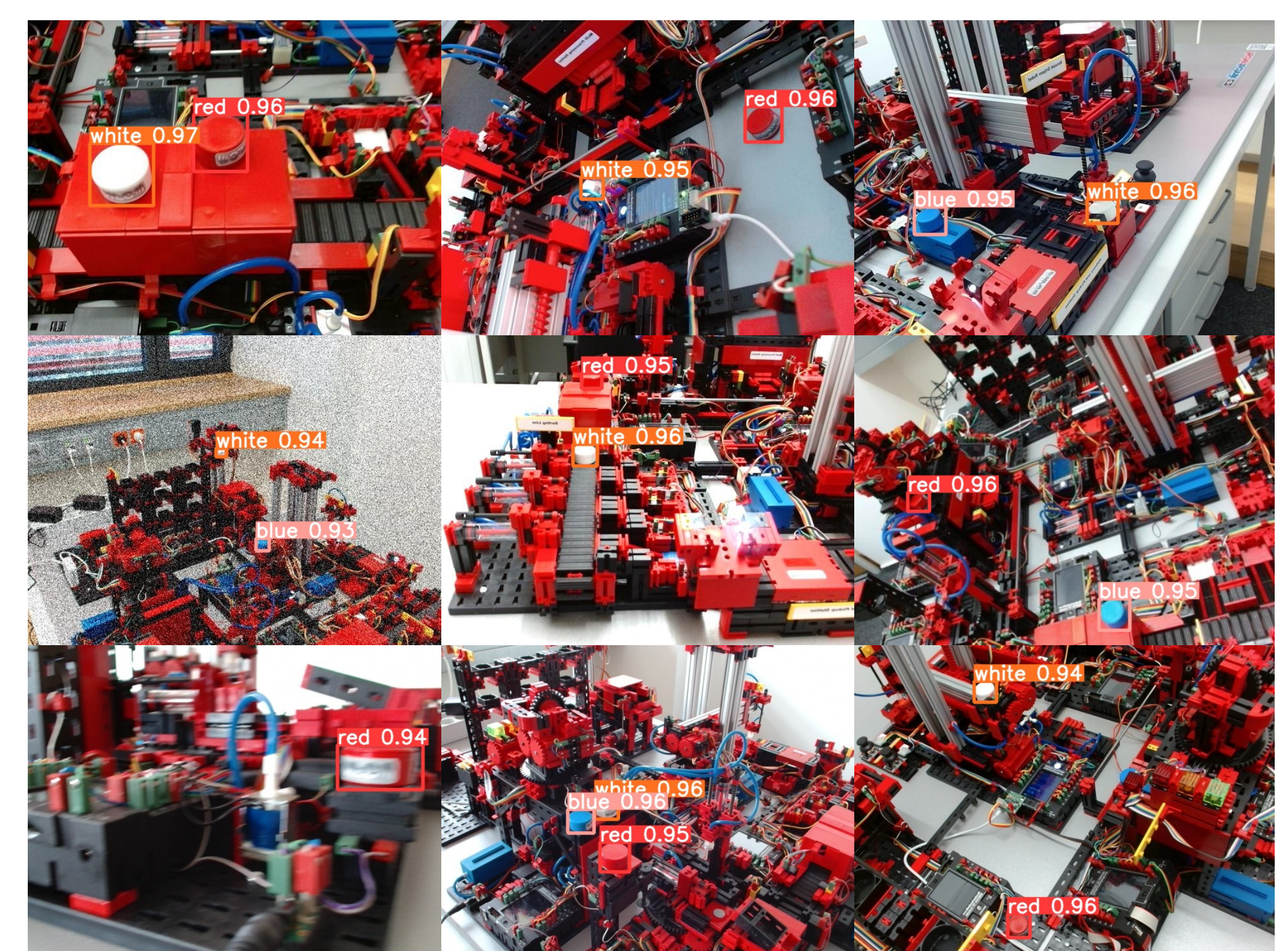


Beispielhafte Darstellung der Datensatzerweiterung für ein Trainingsbild mit (v.l.n.r.) hinzugefügter Bewegungsunschärfe, veränderter Helligkeit und hinzugefügtem Rauschen

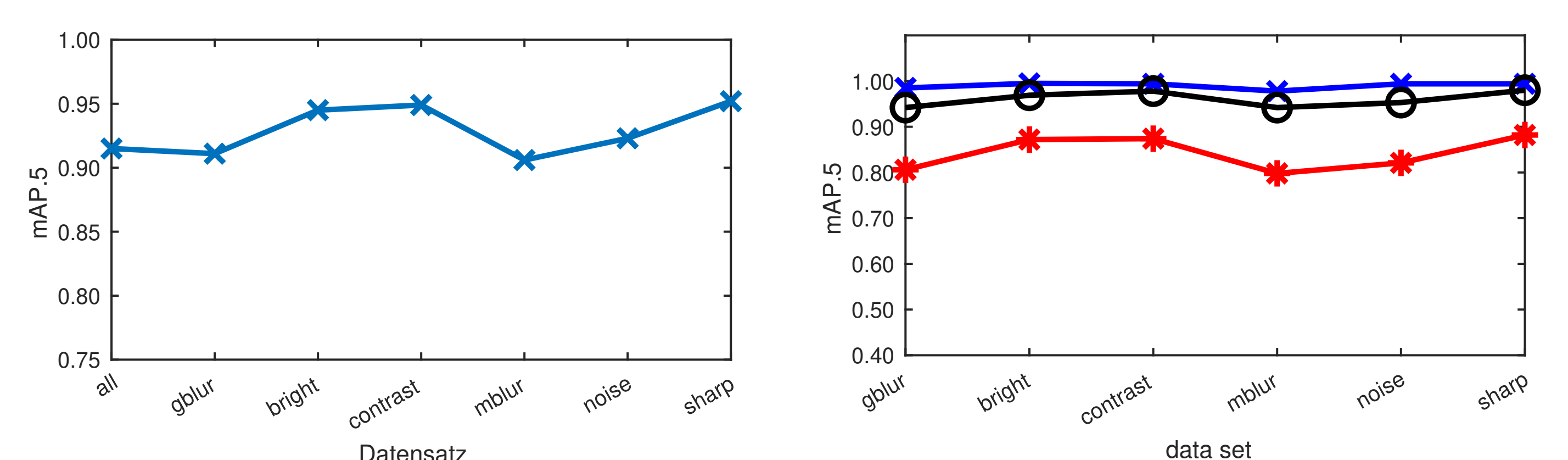
Datensatz: Trainingsdatensatz besteht aus 4900 (gelabelten) Trainingsbildern

Ergebnisse

Beispielbilder der Objekterkennung mit YOLOv3



Genauigkeit von YOLOv3 für Datensätze (links) und Werkstücke (rechts)



Quellen

Joseph Redmon et. al.: *You Only Look Once: Unified, Real-Time Object Detection*. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779-788.

Slavomira Schneidereit: *Untersuchung der Objekterkennung mit Neuronalen Netzen unter Verwendung von YOLO-Modellen in einer Lernfabrik*. Masterarbeit mit Abschluss M.Sc., 2022.