

Pose Estimation

Mert Özmeral & Jerome Habanz

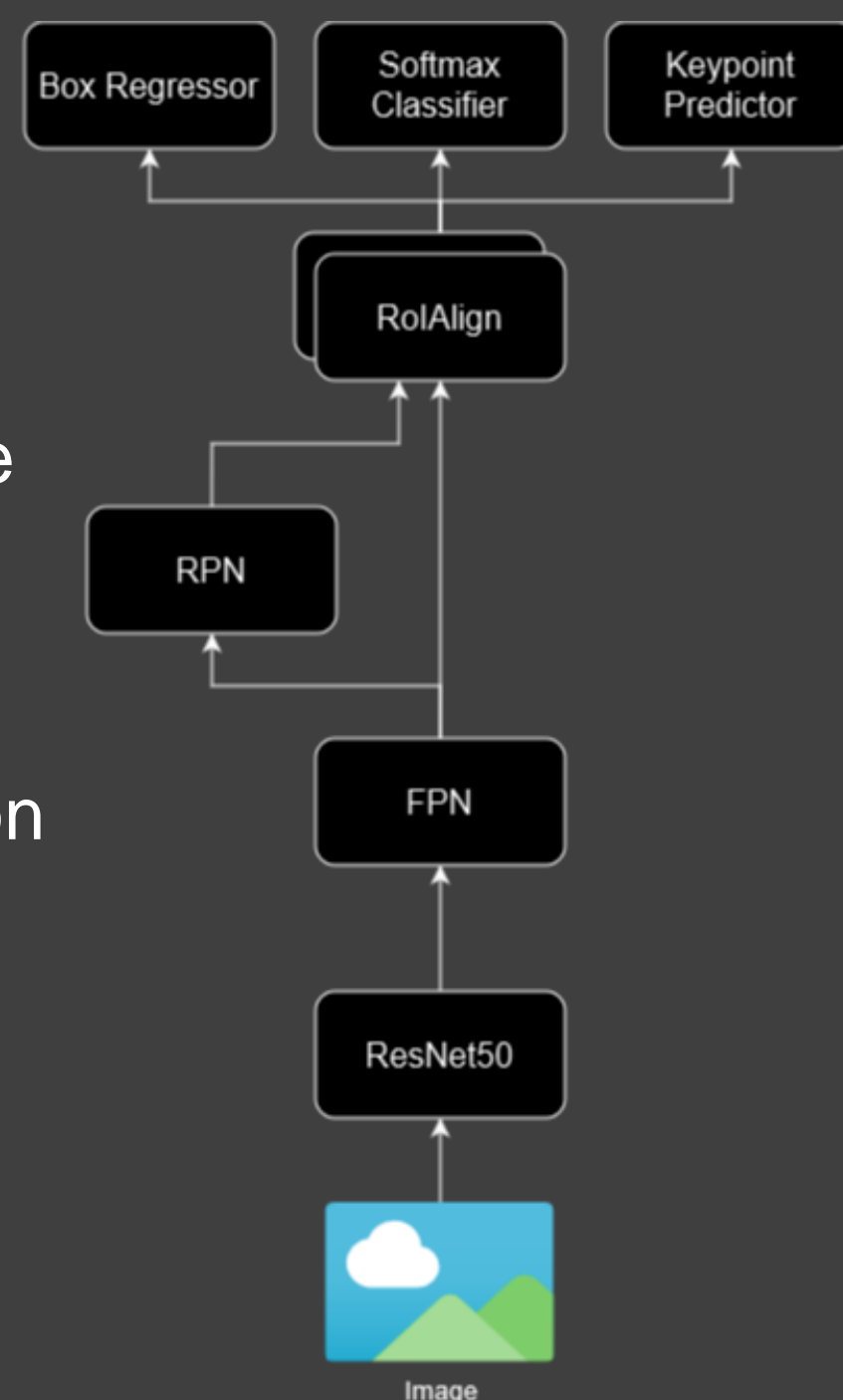
Beschreibung des Projekts

Die steigende Komplexität von Bild- und Videodaten erfordert fortschrittliche Methoden wie die Keypoint-Detektion mittels Convolutional Neural Networks (CNNs), insbesondere das Mask RCNN-Modell. Diese Untersuchung zielt darauf ab, die Leistung von Mask RCNN in der Multi-Object Pose Estimation zu verbessern. Solche Modelle sind entscheidend für Anwendungen in der Computer Vision, Robotik und autonomen Systemen, da sie komplexe Szenarien präzise interpretieren können, was zu verbesserten Entscheidungsprozessen führt.



Architektur

Das Keypoint RCNN nutzt ein ResNet Feature Pyramid Network (FPN) als Backbone, um Merkmale aus dem Bild zu extrahieren und eine hierarchische Merkmalspyramide zu erzeugen. Diese Pyramide wird an das Region Proposal Network (RPN) weitergegeben, das Region of Interests (ROIs) identifiziert. Die ROIs werden dann an Box-Regressor, Softmax Classifier und Keypoint Predictor weitergeleitet.



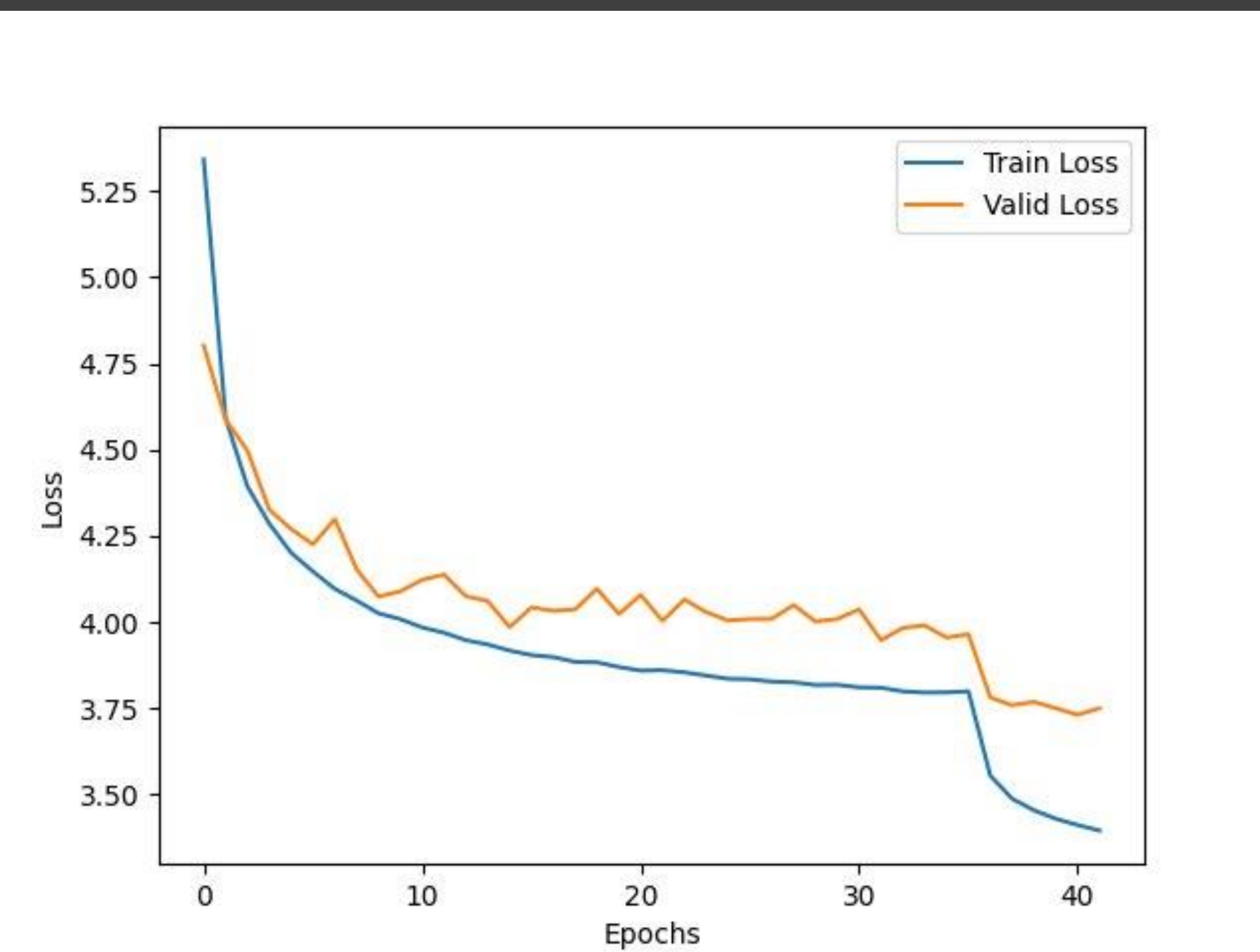
ResNet50 FPN R-CNN

Hyperparameter

- Epochen: 46
- Batchgröße: 8
- SGD:
 - Lernrate: 0.02
 - Momentum: 0.9
 - Gewichtsabnahme: 1e-4
- MultiStepLR:
 - Milestones: (36, 43)
 - Gamma: 0.1

Training

Die Lernstrategie für die Pose Estimation vereint eine Vielzahl von Maßnahmen, um das Training des Modells zu optimieren. Diese umfassen die Filterung des Datensatzes, Transformationen der Bilder und Annotationen sowie die Anpassung der Lernrate durch Optimierungsverfahren. Diese Maßnahmen arbeiten zusammen, um die Leistungsfähigkeit des Modells zu verbessern und seine Fähigkeit zur präzisen Keypoint-Erkennung zu steigern.



Metriken

Das Modell wurde anhand verschiedener Metriken wie IoU, Average Precision (AP) und Average Recall (AR) bewertet. Diese Metriken zeigen die Genauigkeit der Objekterkennung und -lokalisierung bei unterschiedlichen IoU-Schwellenwerten.

Metrik	IoU	Fläche	%
AP	0.50:0.95	all	65.5
AP	0.50	all	86.0
AP	0.75	all	71.2
AP	0.50:0.95	M	62.1
AP	0.50:0.95	L	72.0
AR	0.50:0.95	all	72.3
AR	0.50	all	90.9
AR	0.75	all	77.5
AR	0.50:0.95	M	67.8
AR	0.50:0.95	L	78.7