

Projektübersicht: Vergleich von Super-Resolution-Methoden

Dieses Projekt untersucht mehrere Super-Resolution-Netzwerke (SRCNN-Varianten) mit unterschiedlichen Trainingsmethoden und Skalierungsfaktoren. Ziel ist ein umfassender Vergleich hinsichtlich Bildqualität, Trainingsstabilität und Effizienz.

1. Vergleichsebenen

Die Modelle werden auf drei Hauptebenen bewertet: Trainingsverhalten, Bildqualität und Effizienz. Zusätzlich wird der Einfluss verschiedener Skalierungsfaktoren (x2–x6) betrachtet.

Ebene	Ziele und Metriken
Trainingsverhalten	Analyse des Konvergenzverlaufs (Loss-Kurven), Zeit pro Epoche und Stabilität des Lernprozesses
Bildqualität	Bewertung der rekonstruierten Bilder mittels PSNR, SSIM und LPIPS (visuelle Wahrnehmung)
Effizienz	Vergleich von Parametern, FLOPs, Inferenzzeit und Speicherbedarf (praktische Nutzbarkeit)
Skalierungsfaktor	Untersuchung der Leistung bei x2, x3, x4 und x6, um den Einfluss des Upscaling-Faktors zu messen

2. Netzwerkarchitekturen

Drei Varianten des SRCNN werden eingesetzt, um Tiefe und Komplexität zu vergleichen:

Modell	Eigenschaften
SRCNN_low	4 Residualblöcke, keine Channel Attention, ca. 0.5 Mio Parameter. Schnell, einfach, Basis-Modell.
SRCNN_medium	10 Residualblöcke, SE-Attention alle 2 Blöcke, ca. 1.2 Mio Parameter. Balance zwischen Qualität und Rechenaufwand.
SRCNN_high	20 Residualblöcke, SE-Attention alle 2 Blöcke, ca. 2.4 Mio Parameter. Höchste Genauigkeit, aber rechenintensiv.

3. Trainingsmethoden

Die drei Trainingsmethoden repräsentieren unterschiedliche Lernstrategien mit verschiedenen Zielen:

Methode	Beschreibung und Ziel
L1-Training	Verwendet L1-Loss (MAE) für präzise Pixelrekonstruktion. Hoher PSNR, stabiler Lernverlauf.
Perceptual Training	Kombiniert L1- und Perceptual-Loss (VGG-Features). Fördert realistischere, detailreichere Ergebnisse.
Adversarial Training	Fügt GAN-Loss hinzu, um den Realismus zu erhöhen. Erzeugt lebendige Texturen, jedoch langsamer.

4. Skalierungsfaktoren

Unterschiedliche Skalierungsfaktoren bestimmen, wie stark ein Bild vergrößert wird. Je höher der Faktor, desto schwieriger die Rekonstruktion:

Faktor	Eigenschaften
x2	Geringe Vergrößerung, viele Details erhalten. Typisch 32–36 dB PSNR, leicht rekonstruierbar.
x3	Mittlere Vergrößerung, moderate Verluste. Typisch 30–33 dB PSNR.

x4	Stärkere Vergrößerung, schwieriger. Typisch 28–31 dB PSNR.
x6	Sehr starke Vergrößerung, deutliche Informationsverluste. Typisch 25–28 dB PSNR.

Das Projekt vergleicht damit drei Netzwerkarchitekturen, drei Trainingsmethoden und vier Skalierungsfaktoren über mehrere Bewertungsdimensionen (Qualität, Stabilität, Effizienz). Das Ziel ist die Identifikation des besten Trade-offs zwischen Rechenaufwand und wahrgenommener Qualität.