

# Wissenschaftliche Projektübersicht: Vergleich von Super-Resolution-Methoden

Das vorliegende Projekt untersucht verschiedene Convolutional Neural Network (CNN)-Architekturen zur Bild-Superauflösung (Super-Resolution, SR) mit dem Ziel, die Zusammenhänge zwischen Netzwerkkomplexität, Trainingsstrategie, Skalierungsfaktor und resultierender Bildqualität systematisch zu analysieren. Hierzu werden drei SRCNN-Varianten mit unterschiedlichen Tiefen, drei Trainingsmethoden mit verschiedenen Optimierungszielen und vier Skalierungsfaktoren (x2, x3, x4, x6) miteinander verglichen.

## 1. Vergleichsebenen

Der Vergleich erfolgt entlang mehrerer Evaluationsdimensionen, die sowohl quantitative als auch qualitative Aspekte der Modelle berücksichtigen. Dabei werden neben der reinen Genauigkeit auch Stabilität, Effizienz und generalisierende Leistungsfähigkeit betrachtet.

Ebene	Beschreibung
Trainingsverhalten	Analyse des Konvergenzverhaltens anhand des Trainings- und Validierungsverlusts über
Bildqualität	Bewertung der Bildrekonstruktion anhand etablierter Metriken. Neben der klassischen PSNR
Effizienz	Beurteilung der rechnerischen und speicherbezogenen Anforderungen eines Modells. Daz
Skalierungsfaktor	Vergleich der Leistungsfähigkeit der Modelle bei unterschiedlichen Upscaling-Faktoren (x2

## 2. Netzwerkarchitekturen

Zur Untersuchung des Zusammenhangs zwischen Modellkomplexität und Leistungsfähigkeit werden drei Varianten eines SRCNN-basierten Super-Resolution-Netzwerks implementiert. Diese unterscheiden sich in der Tiefe (Anzahl der Residualblöcke) und der Nutzung von Kanalaufmerksamkeitsmechanismen (Squeeze-and-Excitation).

Architektur	Beschreibung
SRCNN_low	Umfasst 4 Residualblöcke ohne Channel Attention. Dieses Modell stellt die Baseline dar, z
SRCNN_medium	Besteht aus 10 Residualblöcken mit Squeeze-and-Excitation-Attention alle zwei Blöcke. D
SRCNN_high	Umfasst 20 Residualblöcke und SE-Attention in regelmäßigen Abständen. Dieses tiefe Ne

Alle drei Modelle bestehen aus denselben Hauptkomponenten: einem großkernigen Eingabefilter (9×9) zur Feature-Extraktion, einem konvolutionalen Rumpf aus Residualblöcken mit PReLU-Aktivierungen sowie einer Ausgangsschicht (3×3). Eine globale Skip-Verbindung verbindet Eingabe- und Ausgabebild direkt, was zu stabilerer Gradientenweitergabe führt. Zur Stabilisierung des Trainings wird Residual Scaling (Faktor 0.1) eingesetzt; auf Batch Normalization wird bewusst verzichtet, um Detailverlust zu vermeiden.

### 3. Trainingsmethoden

Die drei Trainingsmethoden unterscheiden sich hinsichtlich ihrer Optimierungsziele. Während klassische Verfahren auf maximale Pixelgenauigkeit abzielen, fokussieren perceptual- und adversariale Ansätze stärker auf visuelle Wahrnehmung und Texturrealismus.

Methode	Beschreibung
L1-Training	Das Modell wird mit der L1-Loss-Funktion (Mean Absolute Error) trainiert. Diese misst die
Perceptual Training	Hier wird der Verlust im Feature-Raum eines vortrainierten VGG-Netzes berechnet (Perce
Adversarial Training	Das Netzwerk wird zusätzlich durch einen Diskriminator (GAN-Ansatz) trainiert, der zwisch

Diese drei Trainingsstrategien ermöglichen einen Vergleich zwischen rein quantitativer (PSNR-orientierter) Optimierung und qualitativ-wahrnehmungsbasierten Zielen, die auf visuelle Kohärenz und Realismus abzielen.

## 4. Skalierungsfaktoren

Die Super-Resolution-Aufgabe wird für vier Skalierungsfaktoren untersucht (x2, x3, x4, x6). Jeder Faktor stellt eine unterschiedliche Rekonstruktionsschwierigkeit dar. Mit zunehmender Vergrößerung steigt der Informationsverlust im Eingabebild, wodurch die Modelle stärker auf erlernte Texturrepräsentationen angewiesen sind.

Faktor	Beschreibung
x2	Geringe Vergrößerung, hoher Informationsgehalt. Modelle erzielen typischerweise 32–36 dB PSNR.
x3	Mittlere Vergrößerung, moderate Informationsverluste. Erwartete PSNR-Werte zwischen 30–33 dB.
x4	Deutlich anspruchsvoller, feine Details gehen verloren. Typische PSNR-Werte 28–31 dB.
x6	Sehr hohe Skalierung, starker Informationsverlust. PSNR 25–28 dB. Erfordert perceptual loss.

Insgesamt werden somit drei Modellarchitekturen, drei Trainingsstrategien und vier Skalierungsfaktoren in einem faktoriellen Vergleich untersucht. Diese Kombination ermöglicht eine umfassende Bewertung der Trade-offs zwischen Modellkomplexität, Effizienz, Bildqualität und wahrgenommener Realitätsnähe.