

Trainingsmethoden für Super-Resolution Vergleich

Für den Vergleich der drei SRCNN-Architekturen werden verschiedene Trainingsmethoden eingesetzt, um sowohl quantitative Leistung (PSNR/SSIM) als auch qualitative Wahrnehmung (LPIPS, visuelle Qualität) zu bewerten. Diese Methoden unterscheiden sich in der Art der Loss-Funktion und damit im Lernziel.

1. Klassisches Training mit L1-Loss

Die L1-Loss (Mean Absolute Error) misst die durchschnittliche Abweichung zwischen den Pixeln des Super-Resolution- und des Ground-Truth-Bildes. Sie ist robuster als MSE und führt zu schärferen Ergebnissen. Dieses Training optimiert PSNR und SSIM und bildet die Basis für klassische Super-Resolution-Netzwerke wie SRCNN oder EDSR.

2. Wahrnehmungsbasiertes Training (Perceptual Loss)

Beim perceptual-basierten Training wird die Ähnlichkeit der hochskalierten Bilder im Merkmalsraum eines vortrainierten Netzwerks (z. B. VGG19) bewertet. Dadurch kann das Netzwerk lernen, semantisch und visuell ähnliche Texturen zu erzeugen, auch wenn die Pixelwerte leicht abweichen. Das Ergebnis sind realistischere, detailreichere Bilder. Die Loss-Funktion kombiniert L1- und Perceptual-Komponenten, um ein Gleichgewicht zwischen Genauigkeit und Wahrnehmungsqualität zu erreichen.

3. Adversariales Training (GAN-basiert)

Das adversariale Training führt zusätzlich ein zweites Netzwerk (Diskriminator) ein, das echte und künstliche Bilder unterscheidet. Das Generator-Netzwerk (SRCNN) versucht, den Diskriminator zu täuschen. Dadurch werden hochfrequente Texturen realistischer rekonstruiert. Die resultierenden Bilder sind visuell überzeugender, können aber leicht von den Originalen abweichen. Diese Methode verbessert die subjektive Bildqualität und wird häufig in modernen SRGAN- oder ESRGAN-Ansätzen genutzt.

Zusammenfassend ermöglicht dieser Dreiklang an Trainingsmethoden eine umfassende Bewertung der Modelle: von präziser Pixelgenauigkeit (L1) über visuelle Wahrnehmung (Perceptual) bis hin zu realistischer Texturqualität (Adversarial).