Beyond Jagged Edges: Content-Controlled Spectral and Edge Forensics for AI-Generated Face Detection*

Poster Abstract

Maoyuan Xu
Department of Computer Science and Mathematics
Arcadia University
Glenside, PA 19038
mxu@arcadia.edu

Photorealistic AI-generated faces can closely resemble real photographs, while detectors may rely on incidental differences in image source, compression, or composition. Such shortcuts may fail after resizing or JPEG recompression. This study tests whether interpretable spectral and edge properties differ between AI-generated and real face images after controlling for measurable content. A YuNet-filtered subset of WildFake/FaceGuard supplied 2,000 balanced 200 × 200 JPEG face crops. Hungarian assignment paired images using low-frequency luminance, color, symmetry, center position, and coarse spatial layout; balance screening retained 776 pairs. Eight handcrafted measures quantified spectral energy, sharpness, local high-pass variation, and edge structure. Recompression, blur, and resizing were examined to assess feature stability.

The strongest differences occurred from 0.575 to 0.775 of the Nyquist frequency, where AI-generated faces tended to have less energy than real faces (Cliff's delta approximately −0.36 to −0.37). Direct staircase-edge measures were weaker and less consistent. Thus, reduced fine-detail energy should not be interpreted as evidence of more jagged contours. A model combining frozen, pretrained ResNet18 features with handcrafted measures achieved a locked 400-image blind-test ROC-AUC of 0.7482 (95% bootstrap confidence interval 0.7001–0.7952) and balanced accuracy of 0.6875. These results support moderate discrimination within this dataset, but they cannot establish the origin of an individual image. Evaluation on independently collected sources remains necessary.

*Copyright is held by the author/owner.