ACM Multimedia 2026
Rio de Janeiro, Brazil

DFF-2026

Recovering Forensic Traces Through JPEG Artifact Removal for Synthetic Image Detection

Juan Carlos Pérez-Novoa, David Vázquez-Padín, and Fernando Pérez-González
(atlanTTic, Universidade de Vigo)



atlanTTic research center for Telecommunication Technologies
Universidade de Vigo
XUNTA DE GALICIA
CITIC
Cofinanciado por la Unión Europea
Fondos Europeos

The Real-World Effect:

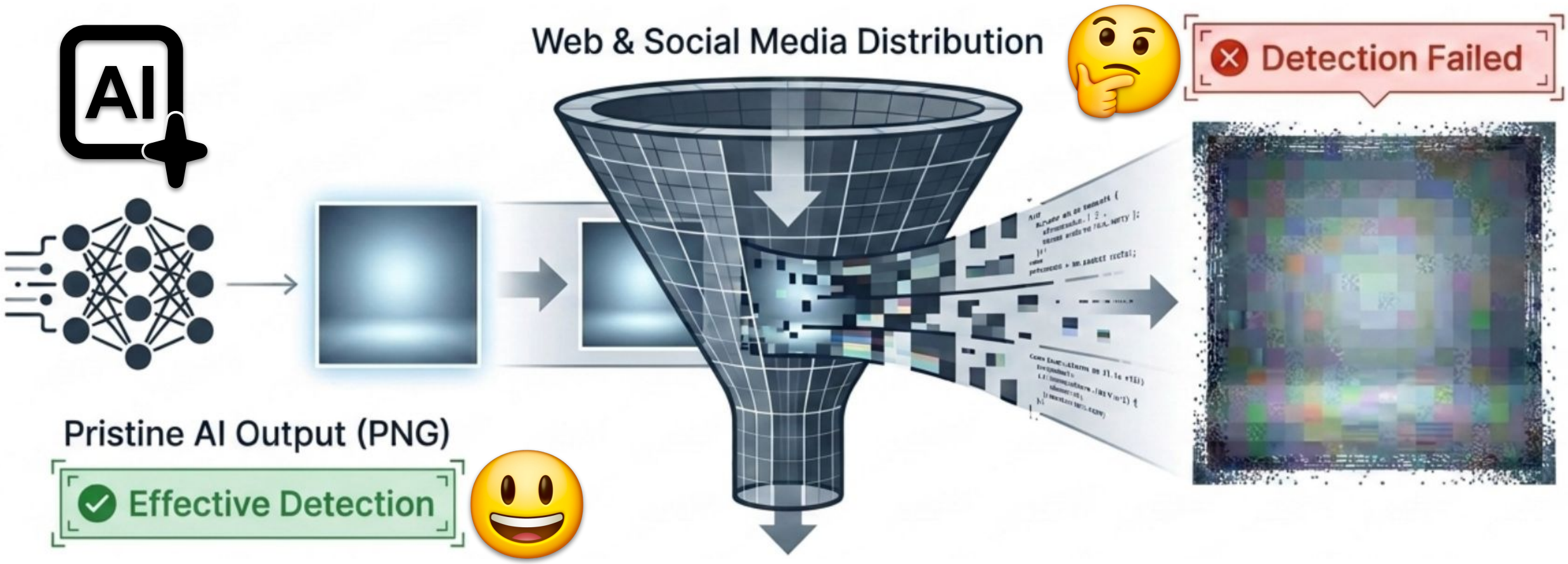


Image content survives compression...



AI

FLUX.2-dev

JPEG



Original Synthetic Image

JPEG (QF=40) Version

But forensic traces change:









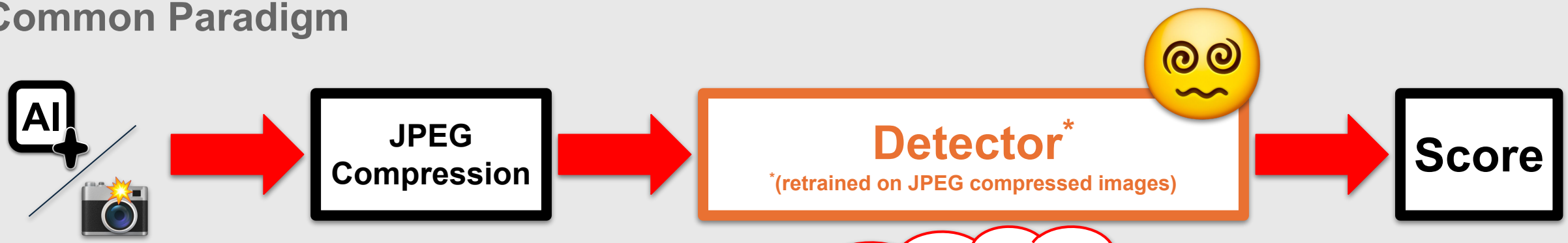
Same origin, but altered forensic traces

Original Synthetic Image

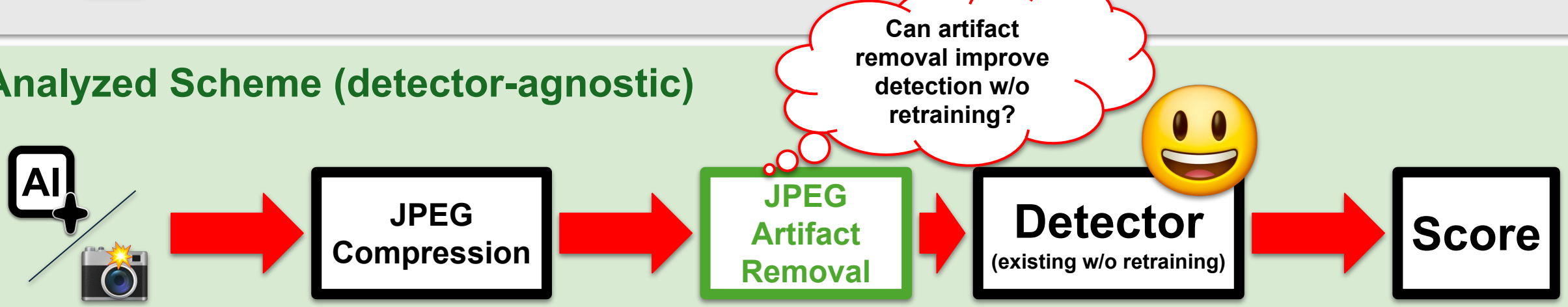
JPEG (QF=40) Version

Recovering Traces via Pre-Processing

Common Paradigm



Analyzed Scheme (detector-agnostic)



JPEG Artifact Removal

JPEG Artifact Removal Methods



DnCNN 1

- Trained on QF=75
- **Moderate recovery**
  - ↳ PSNR 37.32 dB (QF=75)
  - ↳ PSNR 34.37 dB (QF=40)

FBCNN 2

- Blind JPEG restoration
- **Best mid-compression**
  - ↳ PSNR **38.38 dB** (QF=75)
  - ↳ PSNR 35.69 dB (QF=40)

SwinIR 3

- Transformer QF=40
- **Best high-compression**
  - ↳ PSNR 37.21 dB (QF=75)
  - ↳ PSNR **35.97** dB (QF=40)


JPEG (QF=40)


DnCNN


JPEG (QF=40)


FBCNN


JPEG (QF=40)


SwinIR

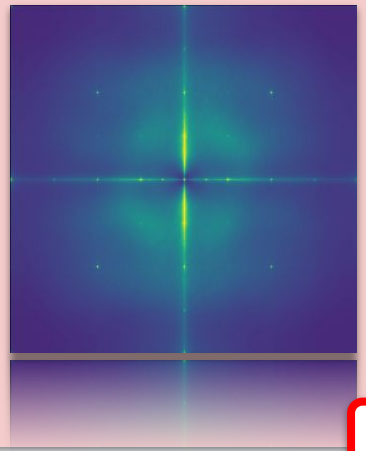
ACM Multimedia 2026
Rio de Janeiro, Brazil
atlanTTic research center for Telecommunication Technologies
Universidade de Vigo
XUNTA DE GALICIA
CITIC gain
Cofinanciado por la Unión Europea
Fondos Europeos
DFF-2026

Detector
(existing w/o retraining)

Synthetic Image Detector Families

Fingerprint-based

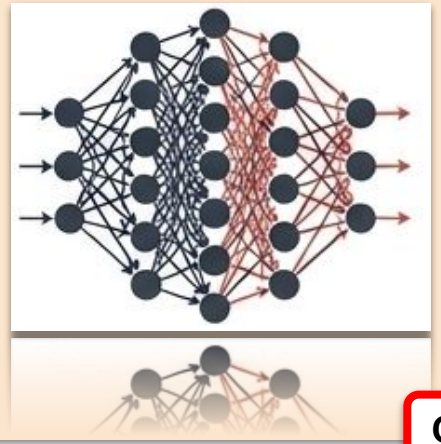
1. Extract **noise residuals**
2. Average spectra via FFT
3. Match source fingerprints via PCE



Fingerprint

ML-based residual

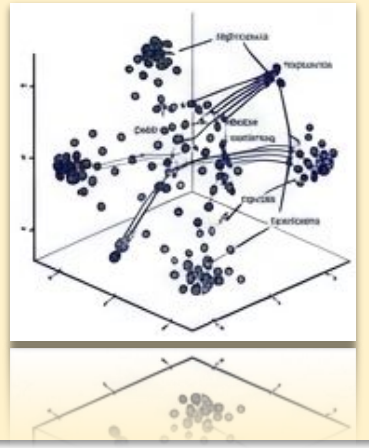
- ResNet50 architecture
- Trained on **low-level Latent Diffusion noise** patterns



Corvi2023

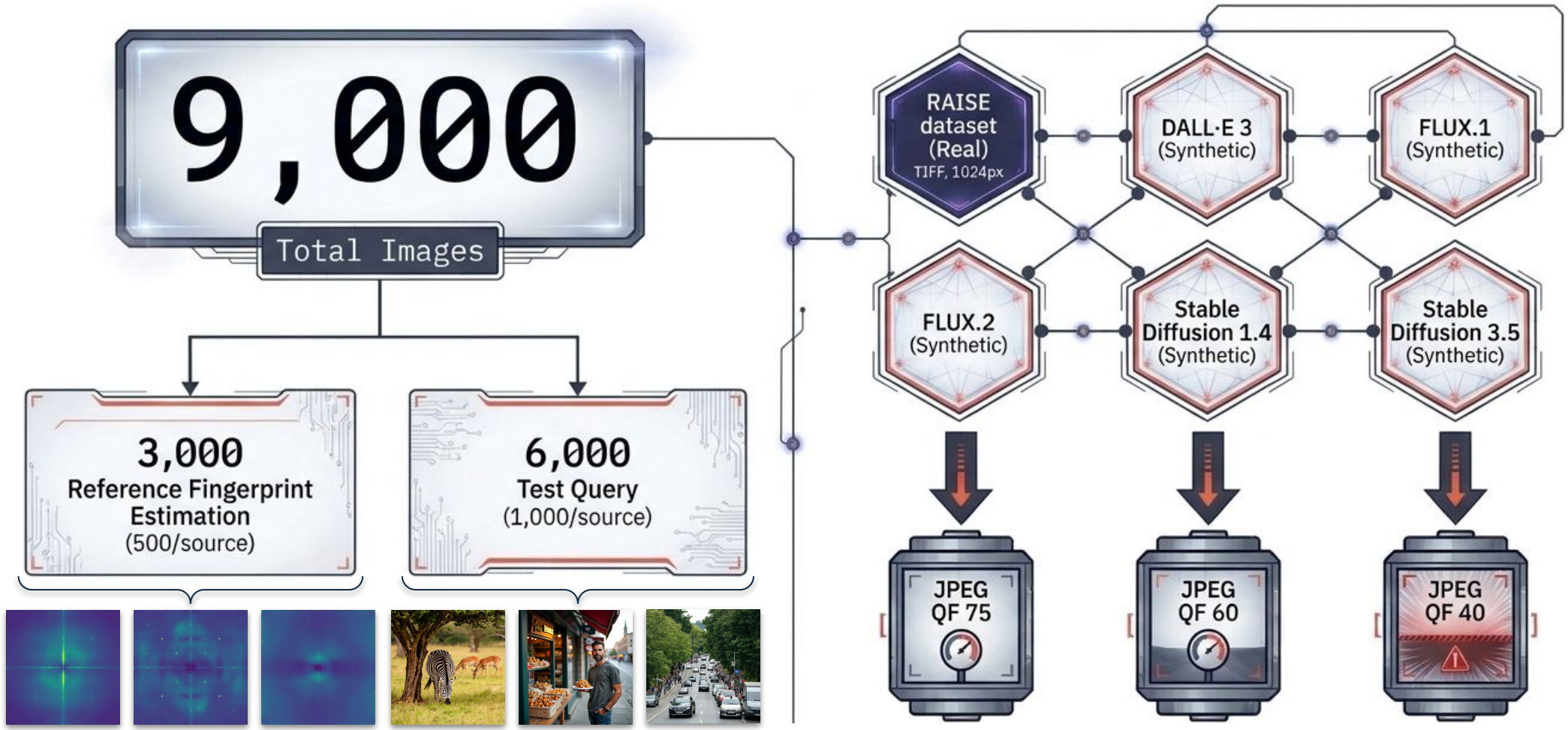
Semantic-based

- High-level vLLM representation extraction
- OpenCLIP ViT-L-14
- Analyze latent positioning

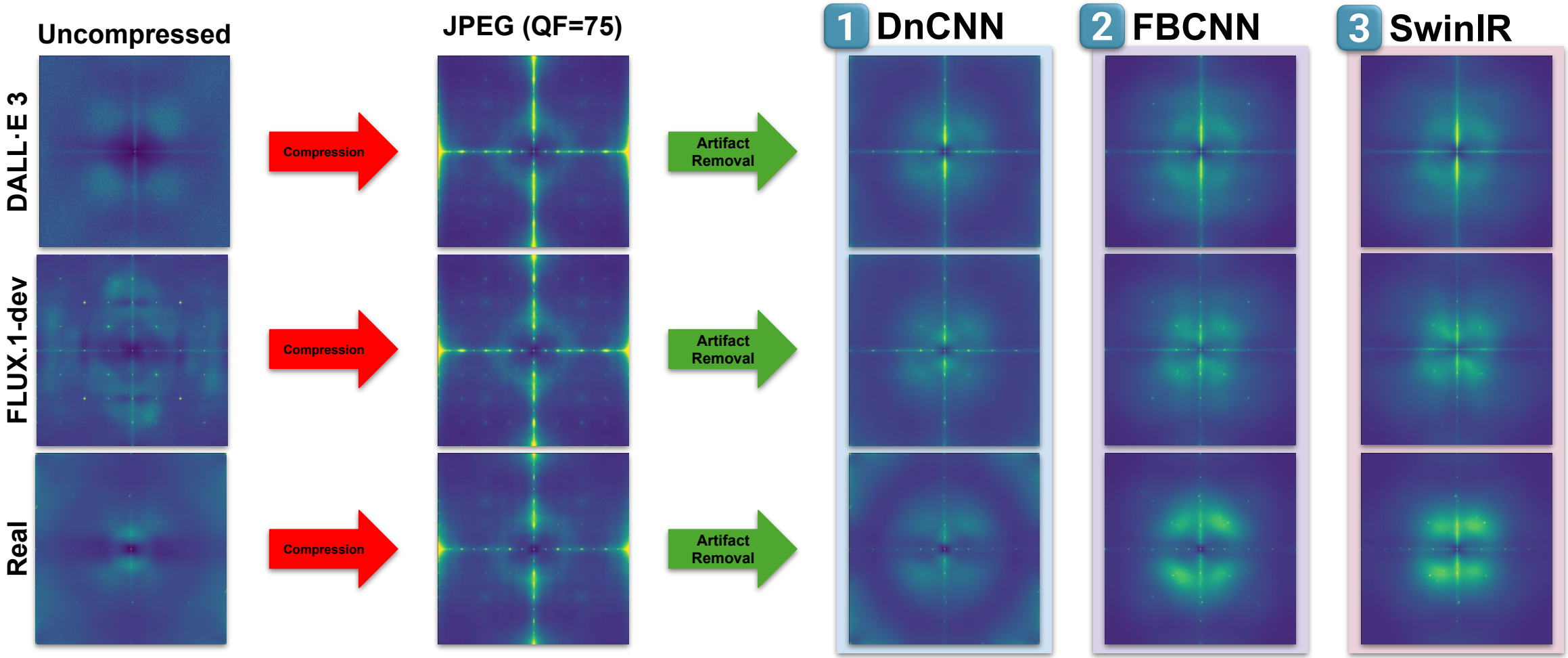


CLIPDet

Experimental Setup: Dataset & Parameters

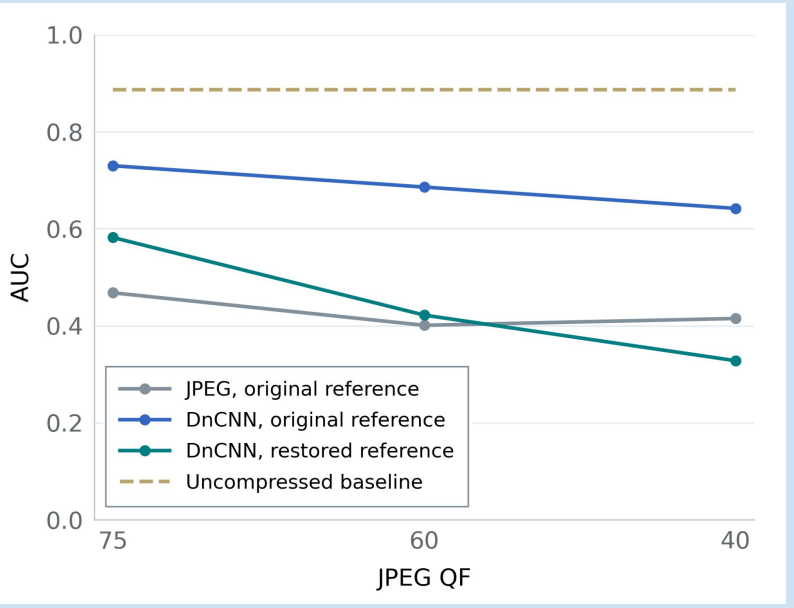


Fingerprint-based: JPEG Artifact Removal Effect

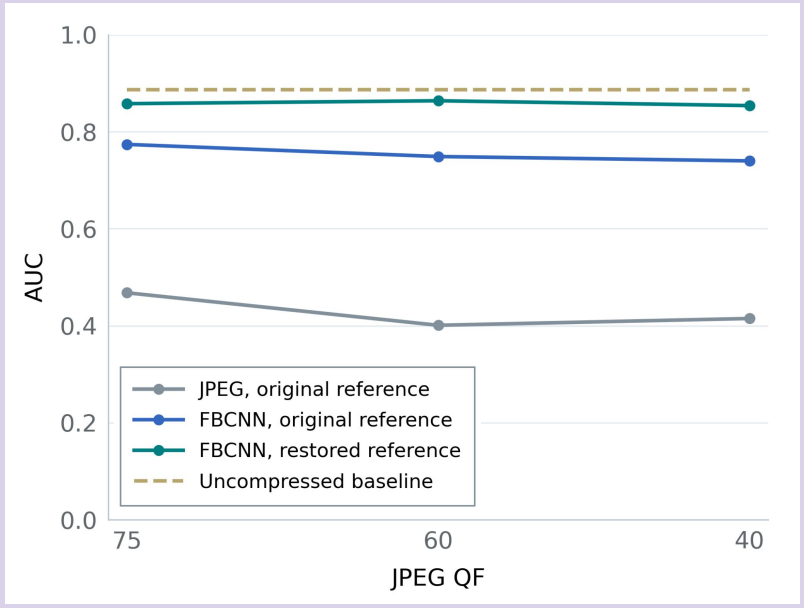


Fingerprint-based: Real vs. Synthetic Performance

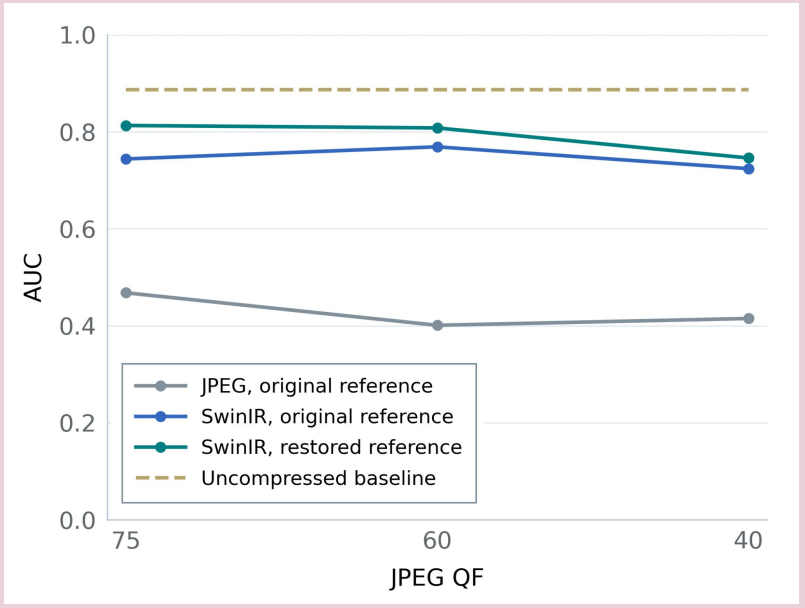
1 DnCNN



2 FBCNN



3 SwinIR



Original Reference

JPEG (Compressed Images) → Fingerprint (Extraction Process)

Restored Reference

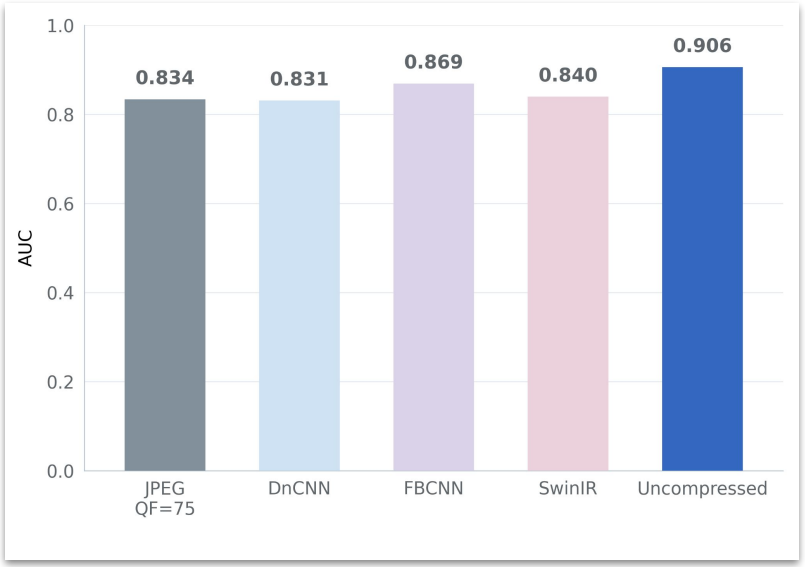
JPEG (Compressed Images) → JPEG (Artifact Removal) → Fingerprint (Extraction Process)

Domain Alignment:
Highest performance requires identical restoration pipeline

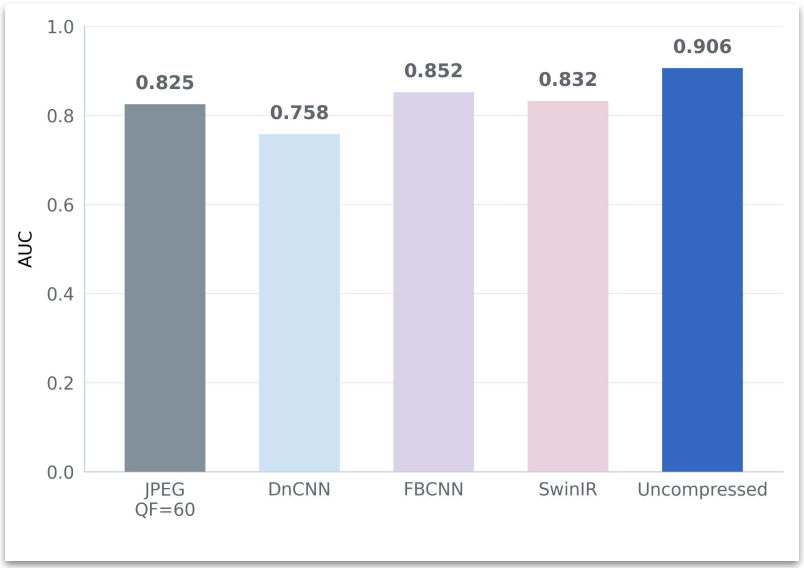
ACM Multimedia 2026
Rio de Janeiro, Brazil

atlanTTic research center for Telecommunication Technologies
Universidade de Vigo
XUNTA DE GALICIA
CITIC gain
Cofinanciado por la Unión Europea
Fondos Europeos
DFF-2026

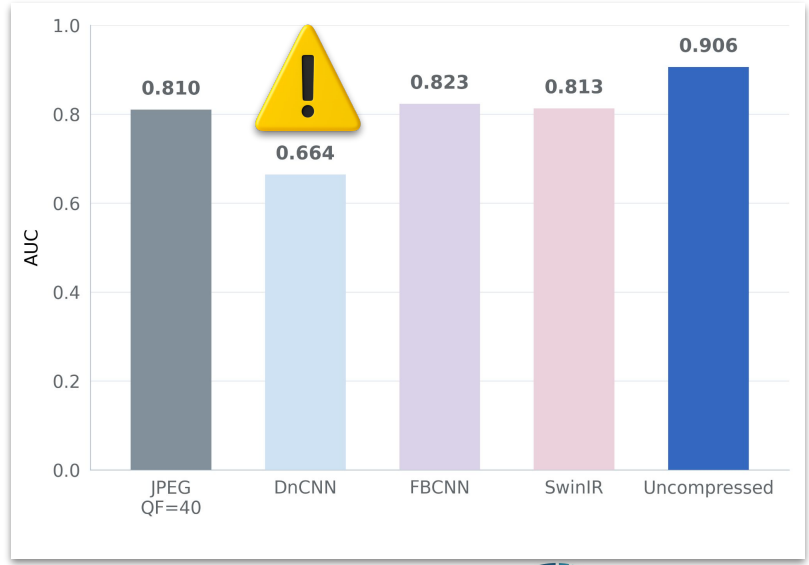
ML-based Residual: Detection Performance



Benefits depend on the restoration method: DnCNN struggles at lower QFs

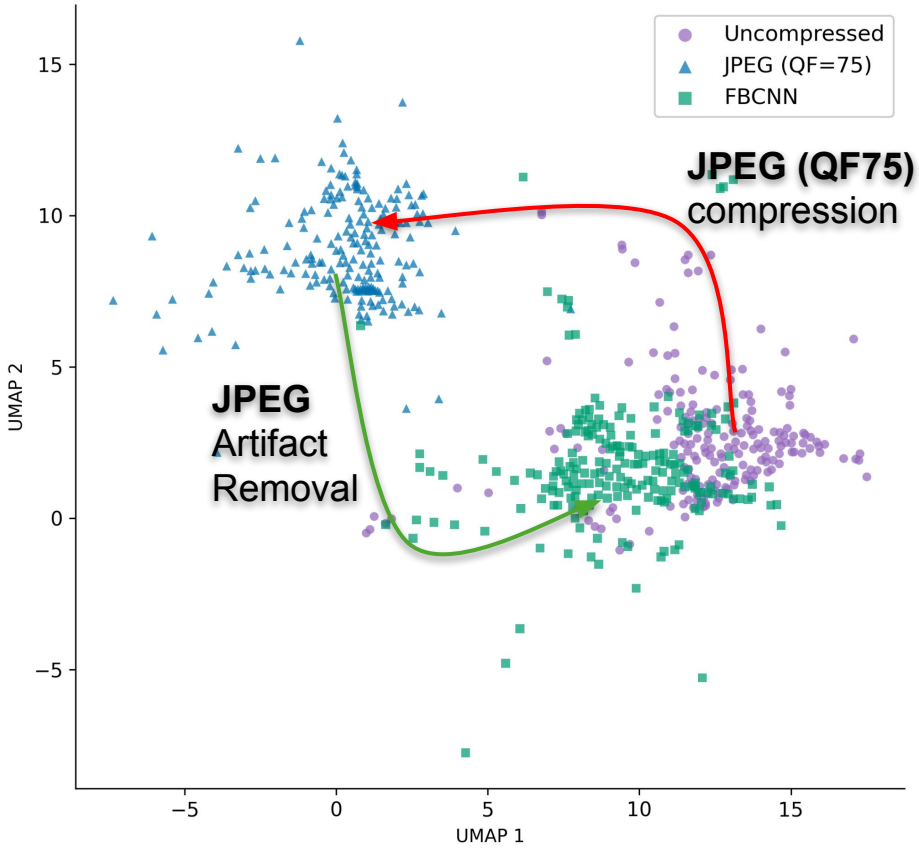


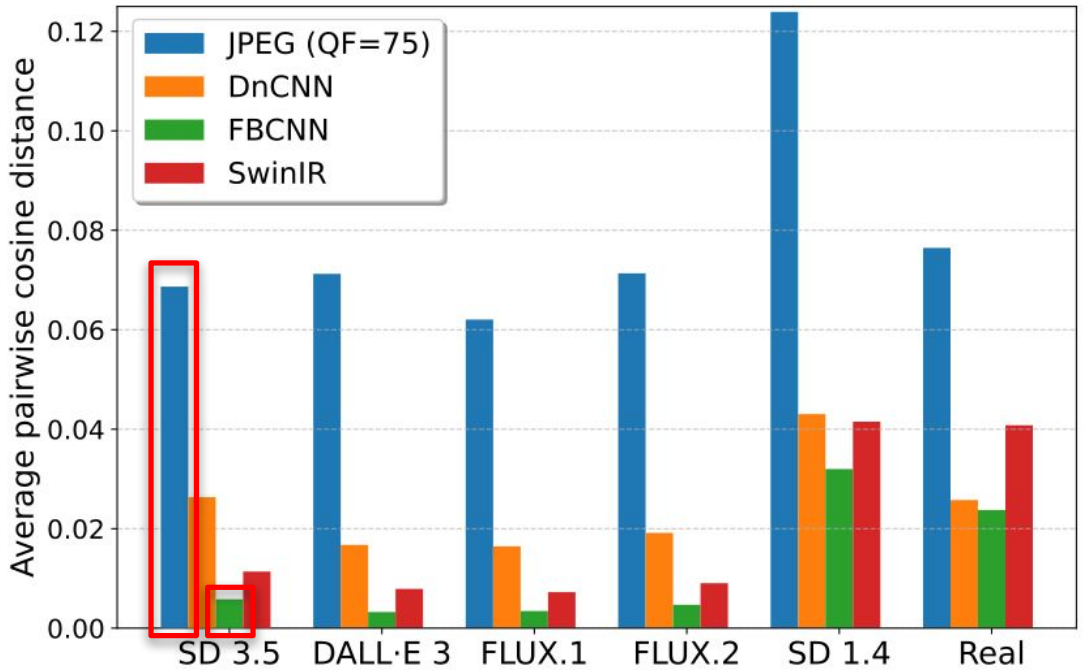
QF ↓↓



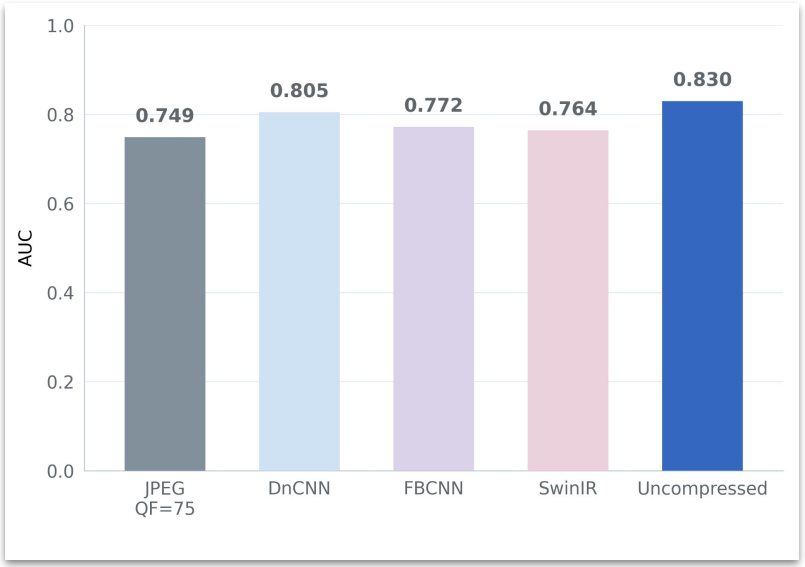
Semantic-based: JPEG Artifact Removal Effect

Intuition: (Stable Diffusion 3.5)

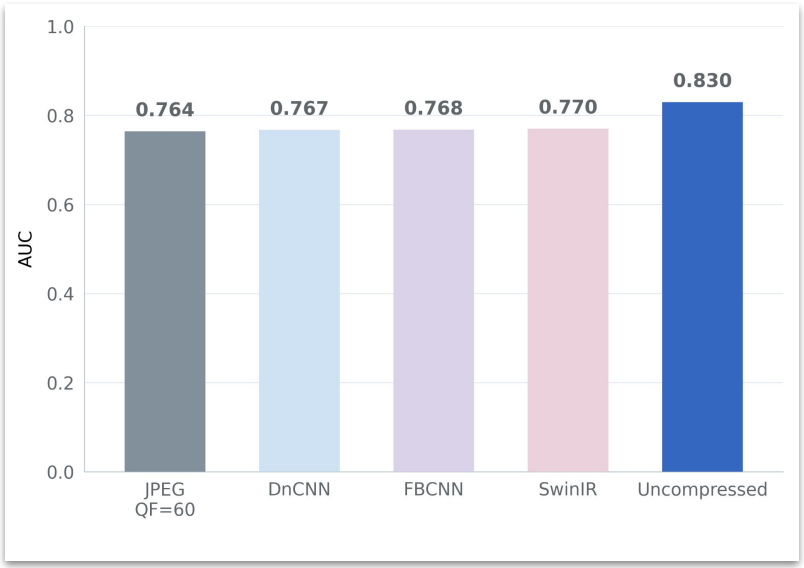




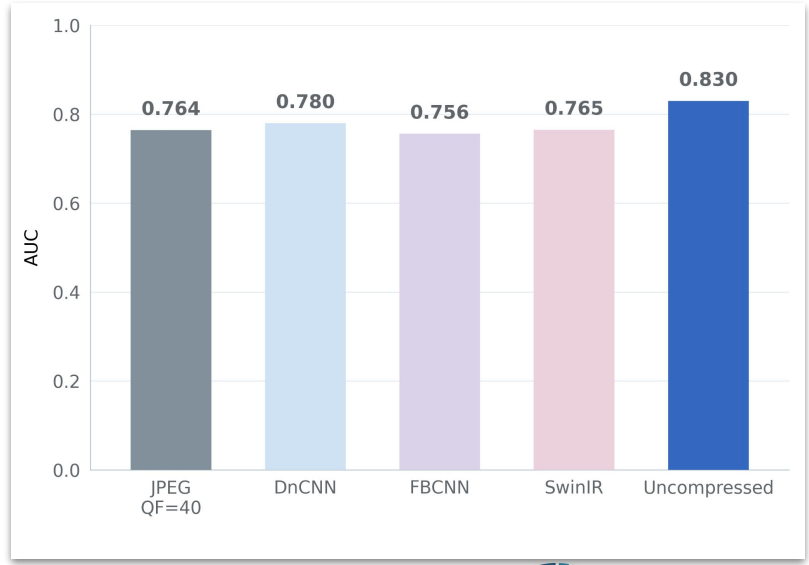
Semantic-based: CLIPDet Performance per QF



Best improvement for QF=75:
Reversing latent space drift gets harder as QF↓↓



QF ↓↓



Key Findings and Open Questions



Domain Matching Matters: with **fingerprints** the reference and query must share the same processing pipeline.

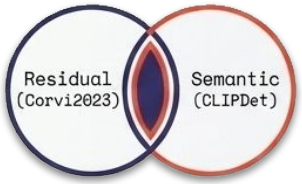
Github Repository



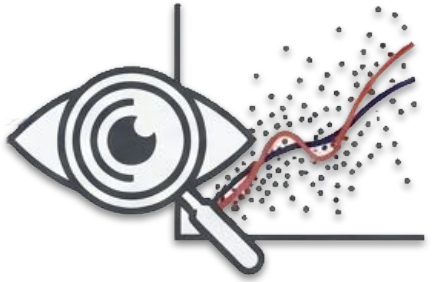


Specialists Beat Generalists:

- **DnCNN** retraining at a single QF limits generalization.
- **FBCNN** generalizes better across JPEG QFs.



Next Step: Explore **fusion** of residual- and semantic-based detectors.



Open Question: Which forensic cues does restoration recover?