The Zero-Shot Illusion in the Wild: Diagnosing Boundary Failures in AIGC Detection

Anonymous authors
Paper under double-blind review

Abstract

Detecting AI-generated images remains a significant challenge due to the rapid proliferation of diverse generative models and workflows. While many detectors report strong zero-shot performance in curated benchmarks, they struggle to generalize to the more complex distribution of real-world AI images, where outputs are from workflows combining multiple models, LoRAs, adapters, post-processing, etc. We expose this *zero-shot illusion* with **WildGen**, an in-the-wild benchmark built from 342K community-created AI images spanning 342 sources, together with 347K real images from 8 authentic datasets. On WildGen, detectors that perform well on existing benchmarks exhibit source-dependent failures: 11 of 12 training-based methods remain below 58% balanced accuracy, and the best training-free method reaches only 75.21% AUC. The evidence is consistent with substantial decision-boundary mismatch rather than a complete absence of discriminative signal in current representations. Within WildGen, few-shot source-level adaptation of off-the-shelf detectors reaches 97.20% mean balanced accuracy across five runs (95.69% using only 18 examples per source) on held-out validation images for AI-vs-real classification, and real-source identity is predictable at 84.33% balanced accuracy. Detectors trained on WildGen can also be extended to GenImage and ForenSynths through balanced joint few-shot fine-tuning that uses examples from both WildGen and the target benchmark. Rather than implying a static universal detector, our evidence supports reporting each detector's valid distributional boundary and treating few-shot updates as a practical maintenance tool as generators and content-creation workflows evolve.

1 Introduction

Rapid advancements in generative models like FLUX.2 (Black Forest Labs, 2025), Qwen-Image (Wu et al., 2025a), and Gemini (Team et al., 2023) now yield high-quality images that are often indistinguishable from reality. While this progress supports creativity, it also increases the risks of disinformation, impersonation, and provenance laundering. Robust AI-image detection therefore requires more than a high score on a static benchmark: it requires evidence that a detector's claimed generalization survives user-driven community conditions where generated images actually circulate, including tuned checkpoints, LoRAs, adapters, and post-processing workflows.

Existing defense mechanisms, categorized into training-based and training-free approaches, have reported high success rates on standard benchmarks like ForenSynths (Wang et al., 2020) and GenImage (Zhu et al., 2023). We argue that these results can create a *zero-shot illusion*: methods appear to generalize because the benchmark distribution is narrow or shares shortcuts with the training data, but their decision rules do not survive a heterogeneous wild setting. In our evaluation, several detectors collapse toward one class or depend on source-specific artifacts when exposed to community generators and traceable real-image sources. This failure reveals that the central question is not whether a detector is universally zero-shot, but where its learned *decision boundary* is valid, how source bias affects it, and when that boundary must be updated.

The generative ecosystem itself is far from static. Real-world generation pipelines have diversified to include granular modifications of base checkpoints, model merging, and modulation via Low-Rank Adaptations

Pipeline	[Base] SDXL Unstable Diffusers - YamerMIX · [LoRA] Detail Tweaker XL:0.85, extremely detailed:1.0 · [Embeds] Fast Negative Embedding, Beyond SDXL Negative
Prompt	In this mythical night scene viewed from high above, the panorama unfolds with breathtaking grandeur... *[truncated]*
Neg. Prompt	FastNegativeV2, low res, poor quality, BeyondSDXLv3
Config	Steps: 40 · CFG: 6.0 · Sampler: DPM++ 3M SDE Karras · Seed: 4004999179

Figure 1: A representative example illustrating the composite nature of the benchmark. The generation pipeline involves a complex combination of a base checkpoint and multiple LoRAs, modulated by prompt embeddings. This workflow demonstrates the high complexity of prompts and resource chaining in real-world content creation.

(LoRAs) (Hu et al., 2022). Images are rarely the raw output of a single model; they are the result of workflows involving text prompts, prompt embeddings, adapters, and post-processing (Figure 1). At the same time, the "Real" side of a detector is not homogeneous: real-image datasets have their own source-specific semantics, resolutions, compression patterns, and curation biases. A benchmark that ignores either side of this boundary can overstate progress.

To study these effects, we introduce WildGen, a large-scale in-the-wild benchmark constructed from community-generated images and traceable real-image sources. The current corpus indexes 1,005,490 AI images from 447 mapped generative sources and 347,207 real images from eight identifiable real-data sources. The benchmark split includes 342 AI sources with at least 1,000 images each, reserving 100 validation images per AI source and 1,000 validation images per real source. This design lets us expose zero-shot failure in the wild, examine feature-space separability and boundary mismatch, and test how efficiently the boundary can be maintained under cross-benchmark, real-source, and environmental shifts.

Our contributions are summarized as follows:

- We introduce WildGen, an in-the-wild benchmark for AIGC detection that better reflects both sides of the deployment distribution. On the AI side, WildGen contains community-generated images produced by real workflows involving base checkpoints, LoRAs, prompt engineering, and post-processing. On the Real side, it uses 347k traceable images from eight diverse sources, reducing the risk that detectors obtain apparent generalization by overfitting to a narrow real-image distribution.

- We benchmark representative existing detectors on WildGen and expose a zero-shot illusion in current AIGC detection. Methods that report strong performance on curated benchmarks often fail on heterogeneous community generators and diverse real sources: most training-based detectors remain low-performing or collapse toward one class, and the best training-free method reaches only 75.21% AUC. Source-dependent false positives reveal an additional real-side failure mode: detectors trained or calibrated on narrow real datasets can misclassify unfamiliar real sources as fake.

- We show that few-shot adaptation is a practical mechanism for maintaining detector boundaries under distribution shift. Within WildGen, frozen CLIP features already support strong AI-vs-real separation with very few examples per generator, while cross-benchmark transfer to GenImage and ForenSynths reveals larger boundary mismatch that benefits from target-domain adaptation. The best adaptation strategy depends on the distance of the target distribution and the capacity of the detector, rather than following a single universal recipe.

2 Related Work

2.1 AI-Generated Image Detection

Existing methods can be categorized into two primary paradigms: *training-based* and *training-free* approaches (Table 1).

Training-based Methods. Early detection methods primarily targeted CNN-based generators. Wang et al. (2020) demonstrated that binary classifiers (e.g., ResNet-50, He et al., 2016) trained on specific GANs could generalize to generators outside their training set through robust data augmentation. This sparked research into cross-generator generalization, including gradient-domain analysis (LGrad, Tan et al., 2023) and multi-branch architectures (HiFi-Net, Guo et al., 2023). With the rise of diffusion models, newer strategies have emerged: DIRE (Wang et al., 2023) exploits reconstruction error as a discriminative clue, RINE (Koutlis & Papadopoulos, 2024) and FatFormer (Liu et al., 2024b) leverage foundation model features (e.g., CLIP, Radford et al., 2021) with specialized adapters, and NPR (Tan et al., 2024) targets upsampling artifacts via nearest-neighbor representations.

Training-free Methods. Recent works explore zero-shot detection by leveraging intrinsic properties of pretrained foundation models without fine-tuning. One direction exploits feature robustness: WaRPAD (Choi et al., 2025) and RIGID (He et al., 2024b) observe that real images maintain consistent DINOv2 (Oquab et al., 2023) representations under perturbations (e.g., wavelet removal or Gaussian noise), whereas AI-generated images exhibit instability. Another perspective leverages reconstruction fidelity: AEROBLADE (Ricker et al., 2024) measures reconstruction error through latent diffusion VAEs, finding that generated images, originating from the same latent space, reconstruct with significantly lower error than real photographs. Furthermore, ManifoldBias (Brokman et al., 2025) analyzes manifold biases by examining Stable Diffusion's UNet response to latent noise. While versatile, these methods often incur high computational overhead and, as we show in Section 4, struggle with the evolving artifacts of modern generators.

2.2 Benchmarks

Benchmarks have transitioned from early GAN-based collections like ForenSynths (Wang et al., 2020) to modern diffusion-based sets, including DiffusionForensics (Wang et al., 2023) and GenImage (Zhu et al., 2023). As summarized in Table 2, existing benchmarks often exhibit two recurrent limitations. First, they often suffer from a "closed-set" bias by focusing on a limited number of standard architectures rather than the high-visibility community-tuned models that dominate public sharing platforms. Second, they frequently rely on legacy real datasets like ImageNet or LSUN, creating a "real-world gap" due to historical compression and sampling artifacts. We introduce WildGen to address these limitations through a multi-modal, community-driven distribution of high-fidelity synthetic and real data.

Table 1: Summary of open-sourced AI image detection methods. Training-based methods require labeled real/fake pairs, while training-free methods use pretrained features directly.

Method	Type	Backbone
CNNDetection	Training	ResNet-50
UnivFD	Training	CLIP ViT-L/14
LGrad	Training	ResNet-50
HiFi-Net	Training	HRNet
DIRE	Training	ResNet-50
RINE	Training	CLIP ViT-L/14
NPR	Training	ResNet-50
AEROBLADE	Free	SD VAE
FatFormer	Training	CLIP ViT-L/14
AIDE	Training	ResNet-50+ConvNeXt
RIGID	Free	DINOv2 ViT-L/14
WaRPAD	Free	DINOv2 ViT-L/14
DeeCLIP	Training	CLIP ViT-L/14
Effort	Training	CLIP ViT-L/14
ManifoldBias	Free	SD+CLIP+LLaVA
AlignedForensics	Training	ResNet-50

Table 2: Benchmarks for AI-generated image detection. The AI Generators column lists the generative models used to create synthetic images.

Benchmark	Size	Real Source	AI Generators
ForenSynths	720K	LSUN	ProGAN, StyleGAN, StyleGAN2, BigGAN, CycleGAN, StarGAN, GauGAN, Deepfake
DiffusionForensics	100K	ImageNet, LSUN, CelebA-HQ	ADM, DDPM, IDDPM, LDM, PNDM, Stable Diffusion, VQ-Diffusion
GenImage	1.3M+	ImageNet	Midjourney, Stable Diffusion, ADM, GLIDE, Wukong, VQDM, BigGAN
Chameleon (Yan et al., 2024a)	10K	Online sources	Professionally edited AI images
WildGen (Ours)	342K AI + 347K real	8 traceable sources	342 diverse sources

2.3 Emerging Paradigms and Perspectives

From Zero-shot to Few-shot. The proliferation of few-shot generation algorithms (He et al., 2024a; Zhao et al., 2023) has accelerated the customization of generative models, making few-shot adaptation increasingly vital for tracking evolving threats. Recent strategies like FTNet (Yao et al., 2025), FAMSeC (Xu et al., 2024), OmniDFA (Wu et al., 2025b), and OCC-CLIP (Liu et al., 2024a) rely on retrieval caches, learnable prompts, or parameter-heavy feature-reshaping, typically validated on small-scale benchmarks. Similarly, Urueña et al. (2025) utilize supervised contrastive learning followed by k-nearest neighbor classification to address open-set detection and attribution. In contrast, we target a heterogeneous, popularity-filtered set of community generators and evaluate when lightweight boundary updates are sufficient, when full fine-tuning is more effective, and when zero-shot transfer remains reliable.

Large Multimodal Models for Detection. Some works (Li et al., 2025; Ji et al., 2025) have attempted to use Large Multimodal Models (LMMs) for detection. However, their utility is constrained by computational cost and the difficulty of reproducible benchmarking as hosted models are continuously updated. Appendix D provides a qualitative case study illustrating how general-purpose assistants reason about image provenance, but we do not treat it as quantitative evidence.

Data-Centric Detection. Finally, our work connects to the data-centric AI movement (Gadre et al., 2023). We posit that the "Generalization Gap" in detection is often a boundary and data-coverage problem rather than purely an architectural failure. By shifting the focus from static models to measured distribution shift and targeted updates, we emphasize that maintaining robustness requires traceable data, real-source controls, and adaptation protocols rather than claims of static universality.

3 WildGen Benchmark

To construct a benchmark that faithfully reflects the complexity of the current generative landscape, we implemented a multi-stage data acquisition and curation pipeline targeting a leading community-driven model-sharing platform **Civitai** (Civitai, 2025). To promote reproducibility and future research, the WildGen benchmark is included in the Supplementary Material and will be made publicly available together with all associated code upon publication.

3.1 Data Acquisition and Model Definition

We systematically scanned the platform to harvest a diverse collection of AI-generated images. For each sample, we extracted available generation metadata, including positive and negative prompts, and the detailed model configuration used. Unlike simple templates, these prompts reflect intricate user engineering (e.g., negative embeddings, weighting), capturing the complexity of real-world usage (see Appendix A for additional pipeline examples).

To reflect authentic detection challenges, we define a "generative source" as the underlying base checkpoint. In modern modular generation pipelines, images are frequently conditioned by auxiliary components such as LoRA modules, ControlNet (Zhang et al., 2023) guidance, or *generative post-processing* utilities (e.g., ADetailer, Bing-su, 2023). We systematically treat these modifications as intra-class variations, assigning all images generated from the same foundational checkpoint to a unified source label. It is important to distinguish this *generative* post-processing–which occurs within the model's workflow–from the *distributional artifacts* (e.g., JPEG compression, resizing) that images undergo after publication; while WildGen captures both, our analysis focuses on the former's impact on architectural fingerprints. This aggregation strategy serves two critical purposes: it prevents label fragmentation in the benchmark and compels detectors to capture stable, architectural fingerprints rather than overfitting to transient stylistic or structural artifacts. Consequently, WildGen provides a more rigorous test of detection robustness against the heterogeneous and evolving landscape of generative modeling.

3.2 Two-Stage Curation Pipeline

To ensure that WildGen remains representative of the most relevant generative threats, we implement a two-stage curation pipeline. First, we filter for image-level quality by exclusively retaining samples that meet a strict verifiable community engagement threshold (likes + comments ≥ 1). A prompt vocabulary analysis in Appendix B reveals a dominance of quality-boosting tags such as "masterpiece" and "best quality", reflecting the aesthetic preferences of the community. This engagement-based filtering ensures that the dataset aligns with prevalent community trends, prioritizing the high-quality regime where AI content is most abundant and thus where the potential for confusion with authentic imagery is greatest. Regardless of this stylistic clustering, the semantic content remains diverse, covering a broad spectrum of subjects.

Second, we filter for **model-level significance** by identifying generators that have produced at least 1,000 such samples. This threshold intentionally focuses the fixed benchmark split on the popular head of Civitai, where enough images are available for balanced per-source evaluation. For each eligible source, we sample 1,000 images without replacement using a fixed random seed of 42 to avoid distribution imbalance while preserving the larger indexed corpus for metadata analysis. The resulting split contains 342 eligible AI sources spanning different creators, base families, and aesthetic goals (the full catalog is provided in Appendix F), but it should not be interpreted as coverage of truly rare generators.

3.3 Benchmark Statistics and Split

We initially collect 1,005,490 raw AI-image records from 447 mapped generative sources. The fixed experimental split focuses on the popular head: 342 sources meet the $\geq$1,000-image eligibility threshold, while 105 lower-count sources are excluded because they do not support balanced per-source validation. Within each eligible source, the seeded procedure assigns 900 images to the training pool and 100 to validation. This yields 342,000 AI images: 307,800 for training and 34,200 for validation. The resulting training pool and validation

indices are stored and reused across binary detection and source-identification evaluations. The seed-42 reference experiments use deterministic smaller-N subsamples of the fixed training pool; at $N < 900$, the repeated-run study in Table 6 additionally draws independent training subsamples for each seed while keeping the validation split fixed. At $N = 900$, every run uses the full per-source AI training pool. Assignment is randomized rather than chronological and is therefore not coupled to upload time; the split evaluates coverage of the current distribution rather than temporal forecasting of future generators. To keep WildGen aligned with the current generative landscape, it also includes recently released closed-source generators such as OpenAI 4o Image Gen, Imagen 4, and Nano Banana [1].

3.4 Real Data Integration

Another critical limitation in existing benchmarks is sampling bias within the "Real" class. Liu & He (2024) demonstrates that dataset bias persists even in billion-scale regimes; models can identify image sources (e.g., LAION vs. YFCC) with >85% accuracy by exploiting systematic semantic and stylistic differences. Thus, the distribution of the "Real" class is as critical as that of the "AI" one. When existing benchmarks rely exclusively on legacy datasets like ImageNet (Russakovsky et al., 2015) or LSUN (Yu et al., 2015), detection models risk learning dataset-specific artifacts (see Figure 2) rather than universal properties of authentic imagery. Consequently, these detectors often treat high-quality real images in the wild as out-of-distribution data. This is particularly problematic for training-free methods that calibrate detection hypotheses on specific real distributions; as confirmed in Section 4, these "universal" physical rules often fail when applied to our higher-fidelity real-world distribution.

To avoid unclear provenance on the authentic side, we build a traceable real-image index containing 347,207 images from eight sources: OpenImages (Kuznetsova et al., 2020), YFCC/Flickr CC (Thomee et al., 2016), Places365 (Zhou et al., 2017a), Unsplash (Unsplash, 2026), SUN397 (Xiao et al., 2010), ADE20K (Zhou et al., 2017b), Food101 (Bossard et al., 2014), and Google Landmarks v2 (Weyand et al., 2020). For each source, we sample 1,000 validation images without replacement using the same fixed seed and use the remainder as a training pool, yielding 8,000 real validation images and 339,207 real training images. In binary evaluations, real images are pooled as the authentic class; in source-aware controls, each real source is represented as a separate class. This design lets us measure how much of a detector's boundary is driven by authenticity versus real-source identity.

3.5 Scope and Limitations

WildGen represents the popular, engagement-filtered head of Civitai rather than the full generator ecosystem. Community styles such as anime, furry, fantasy, and photorealistic portraiture are prominent; prompts are not balanced by language or culture, and DALL-E, Midjourney, and Adobe Firefly are absent despite the inclusion of several recent closed-source generators. Platform metadata identifies 136 of the 342 sources (39.8%) as SDXL-based, 106 (31.0%) as SD1.5-based, and 9 (2.6%) as FLUX-based. The remaining 91 (26.6%) belong to other or unspecified families, including 51 Pony and 8 Illustrious sources. These skews limit our claims to the covered community distribution; rare checkpoints, omitted commercial systems, and adversarially prompted generations remain outside the benchmark's validated scope. The main manuscript reports adaptation down to $N = 18$, so lower-shot behavior is not used for the paper's primary claims.

4 Exposing the Zero-Shot Illusion

To ensure valid comparisons and reproducibility, we strictly evaluate representative methods with official open-source implementations and pretrained weights. Concurrent works lacking public inference code or checkpoints are excluded from this benchmark study. Evaluation uses the full fixed validation split: **100 AI images for each of the 342 eligible sources** and **1,000 real images for each of the eight real sources**. All detectors are evaluated from image pixels only; prompts, model names, source IDs, usernames, licenses, and other metadata are used only for dataset construction and analysis, not detector inference. Despite reported state-of-the-art performance on standard closed-set benchmarks, existing methods across

[1] Unlike the rest of the AI corpus, the Nano Banana subset is drawn from the public HuggingFace dataset `bitmind/nano-banana`

(a) GenImage (AI-generated) (b) ImageNet (Real)

Figure 2: Legacy-benchmark samples from GenImage (left) and ImageNet (right), not WildGen samples, often exhibit low resolution (< 256px) and significant compression artifacts. Because these artifacts are tied to benchmark sources rather than authenticity itself, detectors can exploit them as source shortcuts instead of learning robust discriminative features.

multiple paradigms show substantial boundary failures on the WildGen popularity-filtered community-source split. This gap is the empirical basis of the zero-shot illusion: benchmark-reported generalization does not necessarily transfer to wild generators and controlled real-image sources.

At this validation-set size, a worst-case binomial approximation gives a standard error of approximately 0.31 percentage points for aggregate balanced accuracy. Performance gaps of tens of points therefore cannot be explained by finite-sample variation alone. All confidence intervals are computed from the corresponding experimental outputs: balanced accuracy uses Bonferroni-adjusted class-wise Clopper–Pearson bounds, source-identification metrics use binomial intervals, and AP/AUC use bootstrap intervals.

4.1 Training-based methods

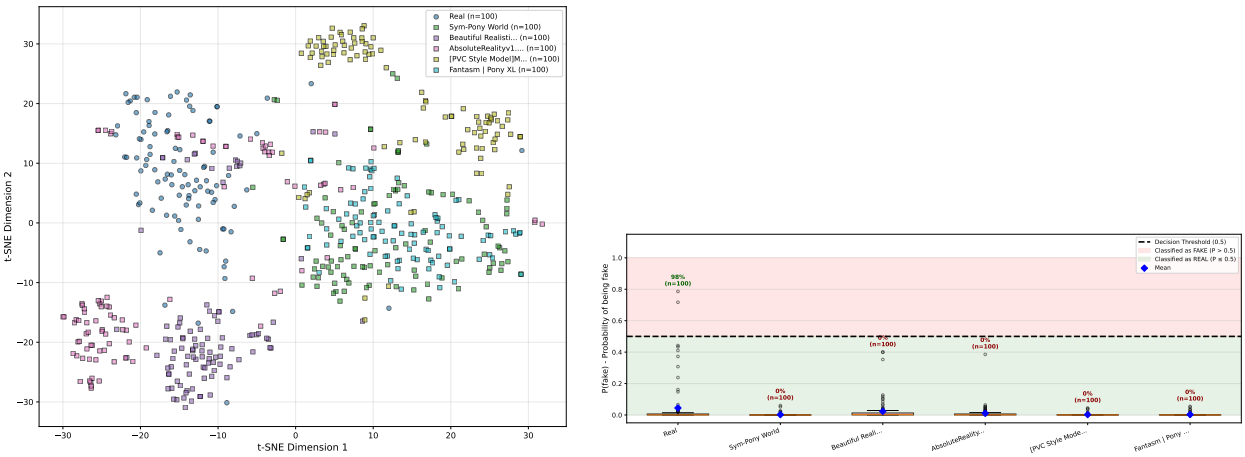
We categorize the evaluated training-based methods into three structural paradigms: (1) **Frozen Backbone Adapters** (UnivFD, Ojha et al., 2023; RINE, Koutlis & Papadopoulos, 2024; DeeCLIP, Keita et al., 2025; FatFormer, Liu et al., 2024b; Effort, Yan et al., 2024b), which adapt a fixed foundation-model backbone with lightweight trainable components (linear heads, adapters, or low-rank/orthogonal updates); (2) **Fully Trainable Classifiers** (CNNDetection, Wang et al., 2020; LGrad, Tan et al., 2023; DIRE, Wang et al., 2023; NPR, Tan et al., 2024; AlignedForensics, Rajan et al., 2025), which fine-tune backbones end-to-end; and (3) **Hybrid/Multi-Branch Architectures** (HiFi-Net, Guo et al., 2023; AIDE, Yan et al., 2024a) that fuse multiple streams. This categorization allows us to assess whether generalization capability correlates with specific architectural choices; for instance, whether large-scale pretraining (CLIP) offers an advantage over specialized training from scratch. Our evaluation shows that methods across all paradigms struggle to transfer to the WildGen community-source split, as shown in Table 3.

Observing Table 3 and the score distributions in Figure 3b, we identify a distinct pattern in failure modes indicative of boundary overfitting. Methods tend to collapse toward a single class prediction when their learned boundary is misaligned with the target distribution. Several detectors achieve high recall on Real images but fail on unfamiliar WildGen AI sources, effectively defaulting to the "Real" label when encountering unfamiliar artifacts. Conversely, DIRE and Effort (Yan et al., 2024b) exhibit the opposite pathology: they correctly identify most AI images (DIRE: 99.3%, Effort: 94.2%), but their Real recall drops sharply (DIRE: 5.6%, Effort: 60.9%). These failures show that the zero-shot illusion is not a small calibration error: many static detectors appear to rely on source-specific artifacts from their training sets rather than learning a boundary that transfers across real and generative distributions. The final two columns (Worst Source / Worst FPR) further show that these real-side errors are often concentrated on a few specific real sources—a shortcut we analyze in Section 5.4.

However, feature-space analysis shows that this failure is not simply caused by an absence of discriminative information. A qualitative t-SNE projection in Figure 3a visualizes the frozen-CLIP feature space, but the quantitative evidence comes from explicit probes: a binary logistic regression probe reaches 98.15% balanced accuracy, and a 343-way kNN classifier (342 AI sources plus a single pooled Real class) reaches 89.06%/95.56%

Table 3: Performance of training-based detectors on the full fixed WildGen zero-shot validation split. Balanced accuracy is reported with its conservative 95% confidence interval, computed via Bonferroni-adjusted class-wise Clopper–Pearson bounds. We also report AI recall, Real recall, and the real source with the highest false-positive rate (FPR).

Method	**Bal. Acc. [95% CI] ↑**	**Rec-AI (%) ↑**	**Rec-Real (%) ↑**	**Worst Source**	**Worst FPR (%) ↓**
DIRE	52.45 [52.12, 52.80]	99.33	5.58	SUN397	99.90
Effort	77.54 [76.77, 78.29]	94.16	60.91	Unsplash	97.60
NPR	43.41 [42.60, 44.22]	17.46	69.36	YFCC/Flickr CC	76.00
HiFi-Net	46.86 [46.39, 47.32]	2.88	90.84	YFCC/Flickr CC	17.40
FatFormer	50.05 [49.70, 50.39]	3.48	96.62	YFCC/Flickr CC	9.30
AlignedForensics	56.87 [56.55, 57.17]	14.28	99.46	YFCC/Flickr CC	0.90
DeeCLIP	48.78 [48.40, 49.14]	2.53	95.03	YFCC/Flickr CC	16.80
LGrad	44.62 [43.99, 45.23]	6.19	83.04	YFCC/Flickr CC	45.30
UnivFD	49.99 [49.80, 50.16]	0.87	99.11	Places365	1.80
RINE	49.81 [49.61, 49.99]	0.76	98.86	Unsplash	3.60
CNNDetection	49.69 [49.55, 49.82]	0.16	99.22	Unsplash	2.20
AIDE	57.80 [56.89, 58.69]	48.65	66.94	YFCC/Flickr CC	76.60



(a) Qualitative t-SNE projection using CLIP features.

(b) Score distribution of zero-shot UnivFD.

Figure 3: Qualitative visual analysis of UnivFD's performance on WildGen. Quantitative feature separability is reported by the probe and kNN results in the text.

top-1/top-5 accuracy. Consequently, many observed failures are consistent with decision boundaries that do not cover the target distribution. This finding motivates the adaptation and generalization studies in Section 5, where we test when lightweight realignment is sufficient and when larger target-domain shifts require more capacity.

4.2 Training-free methods

Training-free methods operate by leveraging the intrinsic properties of foundation models, typically building upon hypotheses related to feature robustness or reconstruction fidelity (see Section 2). While these approaches eliminate the need for task-specific training and offer theoretical generalization, their practical efficacy on modern, high-fidelity community models remains largely unvalidated. Because these methods produce scores without a fixed decision threshold, we evaluate them with threshold-independent metrics (per-source balanced AP/AUC; Table 4) rather than the pooled balanced accuracy used for training-based detectors, so the two families' numbers are not directly comparable.

Taking RIGID (He et al., 2024b), the highest-performing method in our evaluation (75.21% macro AUC; Table 4), as a case study, we examine its core premise: AI-generated content is more fragile than real content under minor Gaussian-noise perturbations in the DINOv2 feature space. This hypothesis was originally established using LDM (SD1.5) (Rombach et al., 2022) as the representative model to demonstrate a sharp drop in feature similarity under noise. In the source-level example shown in Figure 4, however, AI and real

Table 4: Performance of training-free methods on the current WildGen validation split. For each of the 342 eligible AI sources, we pair its 100 validation images with a deterministic 100-image real subset drawn from the eight traceable real validation sources, then macro-average AP/AUC over sources and report 95% bootstrap confidence intervals.

Method	**AP (%) ↑**	**AUC (%) ↑**
RIGID	70.21 [68.89, 72.19]	75.21 [73.50, 76.61]
AEROBLADE	66.14 [64.70, 67.70]	70.88 [69.11, 72.51]
WaRPAD	57.14 [55.74, 59.24]	61.32 [59.58, 62.48]
ManifoldBias	42.76 [41.16, 44.36]	37.13 [35.64, 38.95]

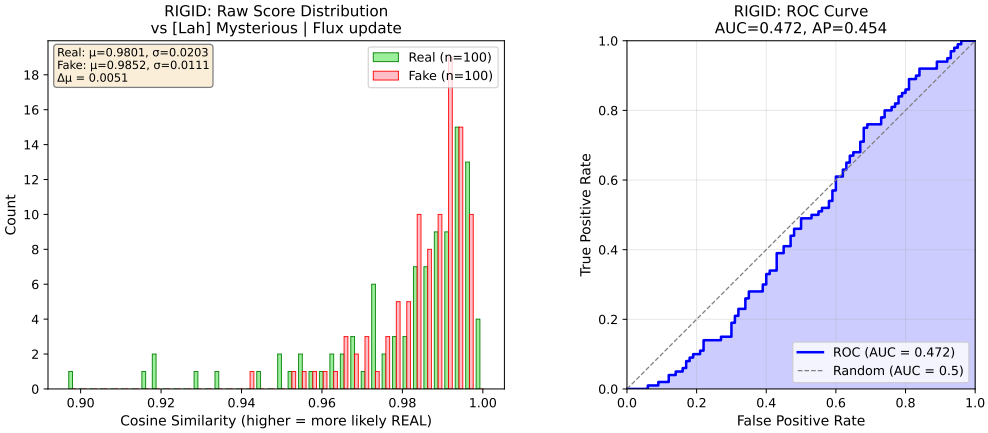


Figure 4: Illustrative RIGID failure on the "[Lah] Mysterious | Flux update" source: (a) raw score histogram and (b) ROC curve.

scores overlap substantially (47.2% AUC), illustrating that this "fragility" does not hold uniformly across WildGen sources. Consistent with this observation, Table 5 reveals a stark architectural disparity: the top-3 best-detected models are all fine-tuned from SD1.5, while the bottom-3 are based on SDXL (Podell et al., 2023).

We treat this architectural disparity as aggregate evidence that training-free assumptions do not hold uniformly across WildGen sources. A frequency-domain case study in Appendix C provides a complementary visualization of one selected RIGID failure. The poorly detected SDXL source exhibits a structured average spectral residual relative to the Real reference, showing that visible frequency-domain differences do not necessarily translate into a high RIGID score. This comparison is descriptive rather than causal: RIGID measures perturbation-induced feature stability, whereas the appendix visualizes static average spectra. The main evidence remains the aggregate AP/AUC in Table 4, the source-level best/worst breakdown in Table 5, and the illustrative score-overlap example in Figure 4.

5 Boundary Adaptation and Generalization

The zero-shot illusion documented in Section 4 raises a fundamental question: do detectors fail in the wild because current representations lack useful AI fingerprints, or because their decision boundaries are misaligned with the target distribution? In this section, we test that question with supervised boundary updates, source-identification probes, and external benchmark transfer. The results support a boundary-maintenance view: within WildGen, frozen features are highly separable and small updates are effective; across external benchmarks, the best adaptation strategy depends on the size and nature of the shift.

5.1 Evaluation Setup

To address the bottleneck of generalization, we investigate adaptation using two representative methods: CNNDetection (full fine-tuning of a pretrained ResNet-50 backbone) and UnivFD (linear probing on frozen

Table 5: Top-3 and Bottom-3 performing models on WildGen based on AUC and AP scores for RIGID, with 95% bootstrap confidence intervals.

Rank	Model Name	AUC ↑	AP ↑	Base Model
		Best 3 Models		
1	qchan Mix v2.0	0.8839 [0.8429, 0.9329]	0.8548 [0.8048, 0.8948]	SD 1.5
2	Toon Babes v1.0	0.8456 [0.7856, 0.8856]	0.8398 [0.7948, 0.8953]	SD 1.5
3	AnyLoraCleanLinearMix-ClearVAE v1.0	0.8359 [0.7859, 0.8959]	0.7736 [0.7086, 0.8186]	SD 1.5
		Worst 3 Models		
1	Pixel Art Diffusion XL	0.1795 [0.1095, 0.2295]	0.3404 [0.2904, 0.4104]	SDXL
2	SVSN001 v1.0	0.3462 [0.2662, 0.4162]	0.3937 [0.3237, 0.4737]	SDXL
3	ProtoVision XL - High Fidelity 3D	0.3762 [0.2862, 0.4462]	0.4072 [0.3372, 0.4972]	SDXL

Table 6: Binary adaptation metrics over five runs. Balanced accuracy and Real-class F1 report mean±sample standard deviation. Bracketed balanced-accuracy intervals are the original seed-42 fixed-validation 95% confidence intervals, retained unchanged to distinguish finite-validation-set uncertainty from run-to-run variation. At N=900, all 900 AI training images per source are used in every run, so only the real subset and training stochasticity vary.

		UnivFD (Linear)		**CNNDet (Fine-tune)**	
N	E	Bal. Acc. mean±SD [95% CI]	F1(Real) mean±SD	Bal. Acc. mean±SD [95% CI]	F1(Real) mean±SD
900	10	97.20±0.09 [96.89,97.52]	96.74±0.11	94.78±0.11 [94.38,95.22]	93.65±0.15
90	100	96.92±0.15 [96.63,97.29]	96.36±0.20	93.58±0.30 [93.29,94.23]	90.53±0.41
45	200	96.69±0.14 [96.33,97.02]	96.09±0.17	91.22±0.35 [90.56,91.63]	88.68±0.50
18	500	95.69±0.17 [95.37,96.14]	94.73±0.29	85.20±0.46 [84.61,85.88]	80.89±0.67

CLIP ViT-L/14 features). We use a data-scaling protocol where each AI model contributes N samples for training, with $N \in \{900, 90, 45, 18\}$.

Unless otherwise specified, the main split uses a real-to-AI ratio of 0.1: for each N, the number of real training images is 10% of the total generative samples. This ratio is a data-efficient compromise rather than a claim of optimality; Section 5.4 evaluates alternatives. We inversely scale training epochs E with sample size N so that each source contributes the same total number of training exposures (Table 6). Both detectors use batch size 64 and learning rate 10^{-4} (Adam for CNNDetection's full ResNet-50; AdamW on the frozen-CLIP linear head for UnivFD). Across all experiments in this paper that train or adapt a detector, we use fixed optimization schedules without early stopping or validation-guided checkpoint selection, and report the final-epoch model. Evaluation is conducted on the fixed validation split: 34,200 AI images and 8,000 real images. For Table 6, we repeat every model–N setting with seeds 7, 13, 21, 42, and 84. At $N < 900$, each run independently resamples the per-source AI training examples and the real training subset, in addition to varying initialization, data order, and augmentation. At $N = 900$, every run necessarily uses the full 900-image AI training pool for each source, so the AI subset is fixed while the real subset and training stochasticity still vary. We report the mean and sample standard deviation across the five runs. For balanced accuracy, the bracketed 95% confidence intervals are retained unchanged from the seed-42 fixed-validation evaluation: we compute conservative intervals using a Bonferroni correction by first obtaining 97.5% class-wise Clopper–Pearson intervals for AI and Real recall and then averaging the corresponding endpoints. These intervals quantify finite-validation-set uncertainty and should not be interpreted as run-to-run confidence intervals.

5.2 Binary Adaptation

Table 6 shows that both models reach high binary separation performance after adaptation. Since the validation split is imbalanced (34,200 AI vs. 8,000 Real images), we emphasize balanced accuracy and Real-class F1 rather than raw accuracy. Across five runs, with independent training-subset resampling at $N < 900$, UnivFD is particularly stable in low-data regimes, reaching $95.69 \pm 0.17\%$ balanced accuracy at $N = 18$ and $97.20 \pm 0.09\%$ at $N = 900$. CNNDetection benefits more from scale, improving from $85.20 \pm 0.46\%$ at $N = 18$ to $94.78 \pm 0.11\%$ at $N = 900$. At $N = 900$, the standard deviations reflect the resampled real subset and training stochasticity because the complete per-source AI training pool is used in every run.

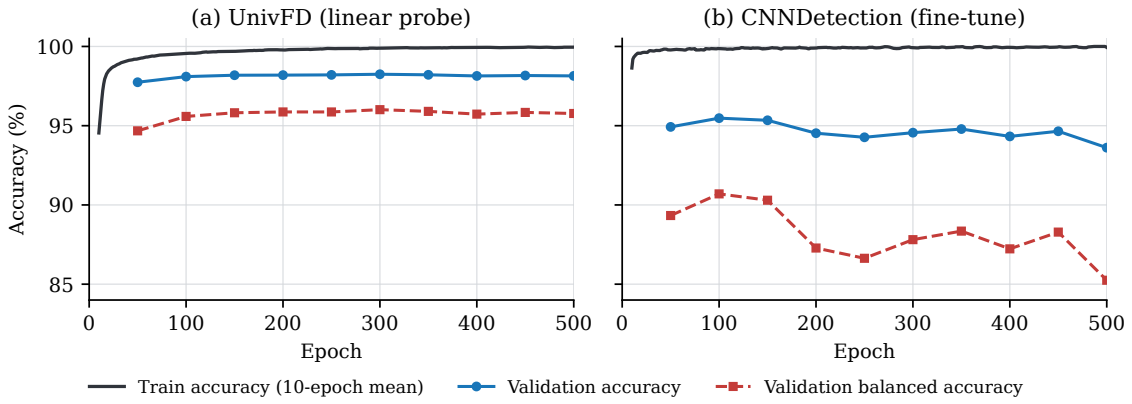


Figure 5: Training accuracy (10-epoch moving average), validation accuracy, and validation balanced accuracy at $N = 18$ for UnivFD and CNNDetection; validation metrics are evaluated every 50 epochs.

These results are consistent with the boundary-maintenance view: small amounts of target-domain data are sufficient to place a useful AI-vs-real boundary within WildGen.

The strong low-shot performance should not be read as universal cross-domain recovery: it reflects a shared train/validation feature geometry that a few examples can realign, whereas later cross-benchmark evaluations show that larger shifts may require more target-domain data or full fine-tuning.

Figure 5 shows the representative seed-42 runs, in which both detectors nearly saturate training accuracy. UnivFD's validation balanced accuracy remains between 94.68% and 96.01%, whereas CNNDetection peaks at 90.70% at epoch 100 and ends at 85.25%, indicating greater low-shot overfitting. To keep the scaling comparison controlled, we use the same optimization recipe and equal total per-source exposures across N. We neither retune hyperparameters for individual data scales nor use early stopping to select favorable checkpoints, since either would introduce scale-specific validation guidance. Reporting the fixed final epoch therefore isolates data scale and model stability rather than each setting's tuned upper bound.

Table 7: Source-identification metrics (%) with 95% binomial confidence intervals on the 350-class source-multiclass split: 342 AI source classes plus 8 real-source classes. Rec-R is exact real-source recall.

N	**E**	**Top-1 (%)**	**Top-5 (%)**	**Rec-R (%)**
		UnivFD (Linear Probe)		
900	10	43.57 [43.12,44.07]	63.85 [63.37,64.28]	82.78 [81.98,83.64]
90	100	37.17 [36.73,37.65]	56.59 [56.09,57.05]	78.28 [77.35,79.18]
45	200	32.61 [32.15,33.08]	51.44 [50.95,51.94]	75.38 [74.41,76.32]
18	500	27.09 [26.65,27.52]	44.51 [44.01,44.98]	70.61 [69.59,71.60]
		CNNDetection (Fine-tune)		
900	10	53.94 [53.45,54.41]	78.46 [78.05,78.84]	72.14 [71.18,73.13]
90	100	33.74 [33.28,34.21]	56.62 [56.13,57.09]	60.03 [58.98,61.11]
45	200	26.59 [26.18,27.03]	46.27 [45.78,46.73]	53.94 [52.83,55.02]
18	500	22.43 [22.05,22.84]	40.49 [40.04,40.98]	53.78 [52.72,54.88]

5.3 Source Identification

Beyond binary detection, the 350-class identification task reveals the nuances of feature disentanglement. Unlike the binary task, real images are split into eight real-source classes, so Rec-R measures exact recovery of the real source rather than simply recognizing authenticity. As shown in Table 7, source identification is substantially harder than binary detection. CNNDetection achieves the strongest high-data source localization (53.94% top-1 and 78.46% top-5 at $N = 900$), while UnivFD preserves stronger real-source recall.

This complexity gap has two causes. First, the increase from 2 classes to 350 classes creates a crowded decision space. Second, many AI sources are phylogenetically related, often sharing base architectures and

differing only by training data, fine-tuning, or adapters. We use source identification as forensic and diagnostic evidence rather than treating it as a solved exact-attribution task.

The optimization dynamics vary with the available data. In the data-rich regime ($N = 900$), CNNDetection's full parameter updates yield a Top-5 accuracy of 78.46%, which is forensically useful even when the exact checkpoint cannot be pinpointed. At $N = 18$, UnivFD instead has higher top-1 accuracy than CNNDetection (27.09% vs. 22.43%), reinforcing frozen features as a stable low-data representation. Thus $N = 18$ separates AI from Real in the binary task, but precise source characterization requires substantially more data.

5.4 Real-Source Bias and Real-Anchor Ratio

The preceding binary results pool all real images into one class, but the real distribution itself is heterogeneous. To quantify this effect, we train an 8-way classifier on frozen CLIP features to predict the real-image source. It reaches 84.33% balanced accuracy (95% CI: 83.51–85.12), confirming that real-source identity is highly learnable (per-source precision/recall in Table 10). This does not invalidate binary detection, but it means that Real cannot be treated as a homogeneous background distribution. This compounds the narrow-Real benchmark problem discussed in Section 3: when the fake generator varies but the Real class stays fixed (e.g., LSUN or ImageNet), a detector can learn the benchmark's specific Real distribution rather than a general notion of authenticity. The source-specific FPR columns in Table 3 test this risk directly by evaluating pretrained legacy detectors on the eight real validation sources in WildGen.

Table 3 provides direct evidence that real-side shortcuts contribute to the zero-shot illusion. The effect is not uniform across methods: AlignedForensics, CNNDetection, UnivFD, and RINE remain conservative on these real sources, while DIRE, NPR, AIDE, Effort, and LGrad misclassify a large fraction of specific real domains as fake. The concentration on YFCC/Flickr CC, Unsplash, and SUN397 suggests that detectors trained on narrow real distributions can treat unfamiliar real-world photography as out-of-distribution fake content. This makes real-source diversity a core benchmark requirement rather than a cosmetic dataset expansion.

Table 8: Real-to-AI ratio sensitivity for the seed-42 reference runs. Each cell reports binary pooled-real balanced accuracy (%) with its fixed-validation 95% confidence interval; the 0.10 rows reproduce the seed-42 values underlying the unchanged confidence-interval brackets in Table 6.

Model	**Real Ratio**	$N = 18$	$N = 900$
UnivFD	0.01	85.10 [84.50,85.69]	91.20 [90.69,91.68]
UnivFD	0.10	95.77 [95.37,96.14]	97.22 [96.89,97.52]
UnivFD	1.00	98.61 [98.37,98.82]	98.88 [98.66,99.07]
CNNDet	0.01	62.70 [62.13,63.27]	86.67 [86.07,87.26]
CNNDet	0.10	85.25 [84.61,85.88]	94.81 [94.38,95.22]
CNNDet	1.00	95.95 [95.57,96.30]	98.64 [98.41,98.85]

Table 8 shows why the main protocol uses a 0.1 real-to-AI ratio: it provides a useful real anchor at much lower data cost than fully balanced training. However, lowering the real ratio to 0.01 substantially harms balanced accuracy, especially for CNNDetection at $N = 18$. We report real-source bias and ratio sensitivity as controls, and avoid interpreting the Real class as a monolithic concept.

5.5 Cross-Benchmark Transfer

Section 4 showed detectors trained on existing benchmarks fail on WildGen; we now test the converse. We adapt both families to GenImage and ForenSynths via *balanced joint few-shot fine-tuning*: warm-starting from the WildGen detector and fine-tuning on K images per source drawn from both WildGen and the target, with the target real anchor aligned to the real pool.

As in the controlled within-WildGen scaling study, the cross-benchmark experiment uses a fixed training budget and reports the final-epoch checkpoint for every model and K, without validation-guided checkpoint selection or early stopping. The same final-epoch checkpoint is used for the target-benchmark result and the WildGen retention evaluation. This keeps checkpoint reporting consistent across the paper and avoids selecting a favorable checkpoint on the evaluation split.

Table 9: Cross-benchmark joint few-shot adaptation using fixed final-epoch checkpoints. Each cell reports balanced accuracy (%) with its 95% confidence interval. At K=900 the WildGen side is the full training pool, i.e. a full-data reference rather than a few-shot point. The "WildGen" column re-evaluates each K=900 adapted model on the WildGen validation split to measure forgetting.

		Target balanced acc.				**WildGen**
Model	**Target**	Zero-shot	K=18	K=90	K=900	(K=900)
UnivFD	GenImage	61.11 [60.77,61.45]	80.01 [79.62,80.41]	83.92 [83.56,84.28]	85.02 [84.64,85.40]	96.31 [95.94,96.66]
UnivFD	ForenSynths	51.32 [50.81,51.83]	85.68 [85.31,86.05]	89.24 [88.91,89.57]	90.05 [89.68,90.40]	96.90 [96.55,97.22]
CNNDet	GenImage	60.68 [60.30,61.06]	72.43 [71.98,72.88]	79.11 [78.70,79.52]	85.94 [85.56,86.32]	94.94 [94.51,95.35]
CNNDet	ForenSynths	51.47 [51.11,51.84]	71.91 [71.44,72.38]	80.49 [80.06,80.91]	94.20 [93.91,94.48]	95.55 [95.14,95.94]

As shown in Table 9, both families recover from weak zero-shot transfer with only a few labelled examples per source. UnivFD is the most data-efficient—at K=18 it already reaches 80.0 (GenImage) and 85.7 (ForenSynths)—while CNNDetection needs more target data but scales higher on the most distant benchmark, overtaking UnivFD on ForenSynths at K=900 (94.2 vs. 90.1). The best strategy thus depends on target-domain distance and trainable capacity, not a single recipe.

Two controls shape these results. *Real-source anchoring is decisive*: under the default 0.1 real ratio the target's own real images are a negligible fraction of the mix and both detectors over-reject them as fake (e.g. UnivFD ForenSynths real recall 57%), capping balanced accuracy at 76–86%; aligning the target real count restores real recall to 91–94% and gives the numbers above. This is the real-source bias of Section 5.4 acting on the AI-vs-real boundary. *Adaptation maintains rather than replaces the boundary*: re-evaluated on WildGen, adapted models still score $\geq$ 96% (UnivFD) and $\geq$ 94% (CNNDetection) balanced accuracy at K=900. Thus, few-shot replay *extends* the valid boundary to a new benchmark while preserving the old one.

Cold-start remains a deployment challenge when a new generator has not yet been identified or labeled. Our binary results suggest a practical first-stage response: a source-agnostic AI-vs-real classifier can screen images without requiring the generator's identity and can be updated once verified examples accumulate. Extending this framework to multiclass attribution and novel-source discovery, including inferring that a group of images originates from a previously unseen generator, remains future work.

5.6 Environmental Robustness

Finally, we evaluate JPEG compression, resizing, and their combination at $N = 18$ and $N = 900$ using the seed-42 final-epoch checkpoints that underlie the fixed-validation confidence-interval brackets in Table 6. These perturbations are not the primary failure mode in our evaluation. For UnivFD, JPEG compression causes the largest drop—balanced accuracy falls from 95.8% to 93.6% at $N = 18$ and from 97.2% to 95.2% at $N = 900$ (about 2 points)—while resizing alone leaves it essentially unchanged or slightly higher (96.4% and 97.5%, respectively). CNNDetection shows no meaningful degradation: at $N = 18$, its balanced accuracy rises from 85.2% clean to 86.2–88.2% under the three perturbation settings; at $N = 900$, it remains within 0.3 points of the 94.8% clean baseline (94.6–94.7%). We treat environmental robustness as supporting evidence: distribution shift and real-source confounding remain the more important failure modes.

6 Conclusion

We expose the zero-shot illusion in AIGC detection: detectors that appear to generalize on curated benchmarks fail sharply in the wild setting of WildGen. The evidence points to a substantial boundary-mismatch component rather than a complete lack of useful features—frozen CLIP features remain highly separable within WildGen—but two factors complicate this view: narrow Real distributions act as shortcuts, producing real-source bias and source-dependent false positives, and transfer to external benchmarks such as GenImage and ForenSynths exposes large distribution shifts. Few-shot adaptation is thus a maintenance mechanism, not a universal guarantee: frozen-feature probes suffice for same-domain updates, full fine-tuning absorbs larger cross-benchmark shifts, and both need real-source controls. The resulting recommendation is to report the distributional boundary within which a detector is valid; WildGen offers a reproducible foundation for that practice as generators and real-image sources evolve.

7 Ethics, Licenses, and Broader Impact

Provenance, licenses, and governance. The AI portion of WildGen is derived primarily from public Civitai image pages and generation metadata. Civitai's Terms allow public user content to be viewed and automated access through provided interfaces, subject to applicable licenses, rate limits, and the remaining Terms.[2] Public availability does not imply unrestricted redistribution: we release URL, hash, and metadata indices rather than image binaries, and users remain responsible for source- and item-level terms. Appendix F reports the audited license of every retained AI source. The real-image sources are governed by their respective terms: OpenImages (CC BY 2.0), YFCC/Flickr CC (image-level Creative Commons), Places365 (CC0), SUN397 (CC BY 4.0), Food101 (non-commercial research agreement), and Unsplash Lite, ADE20K, and Google Landmarks v2 (research or image-level licenses). Versioned manifests record source URLs, licenses, and the NSFW rule; platform removals or relicensing propagate to benchmark availability, and creator or rights-holder opt-outs are removed from later releases.

Sensitive content and privacy. Civitai includes adult-oriented and identity-specific content. We exclude images that Civitai flags as possibly NSFW and do not intentionally include adult content. Because metadata filters are imperfect, releases retain a sensitive-content warning; future versions may add manual or stronger filtering. Some images may use LoRA or related adapters trained on individuals' faces, so WildGen must not be used for face recognition, identity inference, impersonation, demographic profiling, or releasing biometric labels.

Dual use and bias. WildGen supports defensive evaluation but could also help tune generators against known detectors. It is therefore a diagnostic benchmark, not a deployment guarantee, and results should be reported by source. Its Civitai-derived distribution overrepresents anime, fantasy, stylized, photorealistic, and English-prompted workflows while underrepresenting enterprise pipelines, private deployments, and non-English prompt cultures. Detectors trained or adapted on WildGen must be re-evaluated on their target cultural, demographic, and operational distributions.

References

Bing-su. Adetailer. `https://github.com/Bing-su/adetailer`, 2023. Accessed: 2026-01-10.

Black Forest Labs. FLUX.2: Frontier Visual Intelligence. `https://bfl.ai/blog/flux-2`, 2025.

Lukas Bossard, Matthieu Guillaumin, and Luc Van Gool. Food-101 – mining discriminative components with random forests. In *European Conference on Computer Vision*, 2014.

Jonathan Brokman, Amit Giloni, Omer Hofman, Roman Vainshtein, Hisashi Kojima, and Guy Gilboa. Manifold induced biases for zero-shot and few-shot detection of generated images. *arXiv preprint arXiv:2504.15470*, 2025.

Sungik Choi, Hankook Lee, and Moontae Lee. Training-free detection of ai-generated images via cropping robustness. *arXiv preprint arXiv:2511.14030*, 2025.

Civitai. Civitai: The home of open-source generative ai. `https://civitai.com/`, 2025.

Samir Yitzhak Gadre, Gabriel Ilharco, Alex Fang, Jonathan Hayase, Georgios Smyrnis, Thao Nguyen, Ryan Marten, Mitchell Wortsman, Dhruba Ghosh, Jieyu Zhang, et al. Datacomp: In search of the next generation of multimodal datasets. *Advances in Neural Information Processing Systems*, 36:27092–27112, 2023.

Xiao Guo, Xiaohong Liu, Zhiyuan Ren, Steven Grosz, Iacopo Masi, and Xiaoming Liu. Hierarchical fine-grained image forgery detection and localization. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 3155–3165, 2023.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[2] `https://civitai.com/content/tos`

Xiaosheng He, Fan Yang, Fayao Liu, and Guosheng Lin. Few-shot image generation via style adaptation and content preservation. *IEEE Transactions on Neural Networks and Learning Systems*, 2024a.

Zhiyuan He, Pin-Yu Chen, and Tsung-Yi Ho. Rigid: A training-free and model-agnostic framework for robust ai-generated image detection. *arXiv preprint arXiv:2405.20112*, 2024b.

Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, Weizhu Chen, et al. Lora: Low-rank adaptation of large language models. *ICLR*, 1(2):3, 2022.

Yikun Ji, Hong Yan, Jun Lan, Huijia Zhu, Weiqiang Wang, Qi Fan, Liqing Zhang, and Jianfu Zhang. Interpretable and reliable detection of ai-generated images via grounded reasoning in mllms. *arXiv preprint arXiv:2506.07045*, 2025.

Mamadou Keita, Wassim Hamidouche, Hessen Bougueffa Eutamene, Abdelmalik Taleb-Ahmed, and Abdenour Hadid. Deeclip: A robust and generalizable transformer-based framework for detecting ai-generated images. *arXiv preprint arXiv:2504.19876*, 2025.

Christos Koutlis and Symeon Papadopoulos. Leveraging representations from intermediate encoder-blocks for synthetic image detection. In *European Conference on Computer Vision*, pp. 394–411. Springer, 2024.

Alina Kuznetsova, Hassan Rom, Neil Alldrin, Jasper Uijlings, Ivan Krasin, Jordi Pont-Tuset, Shahab Kamali, Stefan Popov, Matteo Malloci, Alexander Kolesnikov, et al. The open images dataset v4: Unified image classification, object detection, and visual relationship detection at scale. *International journal of computer vision*, 128(7):1956–1981, 2020.

Yixuan Li, Yu Tian, Yipo Huang, Wei Lu, Shiqi Wang, Weisi Lin, and Anderson Rocha. Fakescope: Large multimodal expert model for transparent ai-generated image forensics. *arXiv preprint arXiv:2503.24267*, 2025.

Fengyuan Liu, Haochen Luo, Yiming Li, Philip Torr, and Jindong Gu. Which model generated this image? a model-agnostic approach for origin attribution. In *European Conference on Computer Vision*, pp. 282–301. Springer, 2024a.

Huan Liu, Zichang Tan, Chuangchuang Tan, Yunchao Wei, Jingdong Wang, and Yao Zhao. Forgery-aware adaptive transformer for generalizable synthetic image detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 10770–10780, 2024b.

Zhuang Liu and Kaiming He. A decade's battle on dataset bias: Are we there yet? *arXiv preprint arXiv:2403.08632*, 2024.

Utkarsh Ojha, Yuheng Li, and Yong Jae Lee. Towards universal fake image detectors that generalize across generative models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 24480–24489, 2023.

Maxime Oquab, Timothée Darcet, Théo Moutakanni, Huy Vo, Marc Szafraniec, Vasil Khalidov, Pierre Fernandez, Daniel Haziza, Francisco Massa, Alaaeldin El-Nouby, et al. Dinov2: Learning robust visual features without supervision. *arXiv preprint arXiv:2304.07193*, 2023.

Dustin Podell, Zion English, Kyle Lacey, Andreas Blattmann, Tim Dockhorn, Jonas Müller, Joe Penna, and Robin Rombach. Sdxl: Improving latent diffusion models for high-resolution image synthesis. *arXiv preprint arXiv:2307.01952*, 2023.

Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. Learning transferable visual models from natural language supervision. In *International Conference on Machine Learning*, 2021.

Anirudh Sundara Rajan, Utkarsh Ojha, Jedidiah Schloesser, and Yong Jae Lee. Aligned datasets improve detection of latent diffusion-generated images. In *The Thirteenth International Conference on Learning Representations*, 2025. URL `https://openreview.net/forum?id=doBkiqESYq`.

Jonas Ricker, Denis Lukovnikov, and Asja Fischer. Aeroblade: Training-free detection of latent diffusion images using autoencoder reconstruction error. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9130–9140, 2024.

Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 10684–10695, 2022.

Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, et al. Imagenet large scale visual recognition challenge. *International journal of computer vision*, 115(3):211–252, 2015.

Chuangchuang Tan, Yao Zhao, Shikui Wei, Guanghua Gu, and Yunchao Wei. Learning on gradients: Generalized artifacts representation for gan-generated images detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 12105–12114, 2023.

Chuangchuang Tan, Yao Zhao, Shikui Wei, Guanghua Gu, Ping Liu, and Yunchao Wei. Rethinking the up-sampling operations in cnn-based generative network for generalizable deepfake detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 28130–28139, 2024.

Gemini Team, Rohan Anil, Sebastian Borgeaud, Jean-Baptiste Alayrac, Jiahui Yu, Radu Soricut, Johan Schalkwyk, Andrew M Dai, Anja Hauth, Katie Millican, et al. Gemini: a family of highly capable multimodal models. *arXiv preprint arXiv:2312.11805*, 2023.

Bart Thomee, David A Shamma, Gerald Friedland, Benjamin Elizalde, Karl Ni, Douglas Poland, Damian Borth, and Li-Jia Li. Yfcc100m: The new data in multimedia research. *Communications of the ACM*, 59 (2):64–73, 2016.

Unsplash. Unsplash lite dataset 1.4.0. `https://unsplash.com/data`, 2026.

Jaime Álvarez Urueña, David Camacho, and Javier Huertas Tato. Supervised contrastive learning for few-shot ai-generated image detection and attribution. *arXiv preprint arXiv:2511.16541*, 2025.

Sheng-Yu Wang, Oliver Wang, Richard Zhang, Andrew Owens, and Alexei A Efros. Cnn-generated images are surprisingly easy to spot...for now. In *CVPR*, 2020.

Zhendong Wang, Jianmin Bao, Wengang Zhou, Weilun Wang, Hezhen Hu, Hong Chen, and Houqiang Li. Dire for diffusion-generated image detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 22445–22455, 2023.

Tobias Weyand, Andre Araujo, Bingyi Cao, and Jack Sim. Google landmarks dataset v2-a large-scale benchmark for instance-level recognition and retrieval. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 2575–2584, 2020.

Chenfei Wu, Jiahao Li, Jingren Zhou, Junyang Lin, Kaiyuan Gao, Kun Yan, Sheng ming Yin, Shuai Bai, Xiao Xu, Yilei Chen, Yuxiang Chen, Zecheng Tang, Zekai Zhang, Zhengyi Wang, An Yang, Bowen Yu, Chen Cheng, Dayiheng Liu, Deqing Li, Hang Zhang, Hao Meng, Hu Wei, Jingyuan Ni, Kai Chen, Kuan Cao, Liang Peng, Lin Qu, Minggang Wu, Peng Wang, Shuting Yu, Tingkun Wen, Wensen Feng, Xiaoxiao Xu, Yi Wang, Yichang Zhang, Yongqiang Zhu, Yujia Wu, Yuxuan Cai, and Zenan Liu. Qwen-image technical report, 2025a. URL `https://arxiv.org/abs/2508.02324`.

Shiyu Wu, Shuyan Li, Jing Li, Jing Liu, and Yequan Wang. Omnidfa: A unified framework for open set synthesis image detection and few-shot attribution. *arXiv preprint arXiv:2509.25682*, 2025b.

Jianxiong Xiao, James Hays, Krista A Ehinger, Aude Oliva, and Antonio Torralba. Sun database: Large-scale scene recognition from abbey to zoo. In *2010 IEEE computer society conference on computer vision and pattern recognition*, pp. 3485–3492. IEEE, 2010.

Juncong Xu, Yang Yang, Han Fang, Honggu Liu, and Weiming Zhang. Famsec: A few-shot-sample-based general ai-generated image detection method. *IEEE Signal Processing Letters*, 2024.

Shilin Yan, Ouxiang Li, Jiayin Cai, Yanbin Hao, Xiaolong Jiang, Yao Hu, and Weidi Xie. A sanity check for ai-generated image detection. *arXiv preprint arXiv:2406.19435*, 2024a.

Zhiyuan Yan, Jiangming Wang, Peng Jin, Ke-Yue Zhang, Chengchun Liu, Shen Chen, Taiping Yao, Shouhong Ding, Baoyuan Wu, and Li Yuan. Orthogonal subspace decomposition for generalizable ai-generated image detection. *arXiv preprint arXiv:2411.15633*, 2024b.

Shibo Yao, Renshuai Tao, Xiaolong Zheng, Chao Liang, and Chunjie Zhang. Leveraging failed samples: A few-shot and training-free framework for generalized deepfake detection. *arXiv preprint arXiv:2508.09475*, 2025.

Fisher Yu, Ari Seff, Yinda Zhang, Shuran Song, Thomas Funkhouser, and Jianxiong Xiao. Lsun: Construction of a large-scale image dataset using deep learning with humans in the loop. *arXiv preprint arXiv:1506.03365*, 2015.

Z-Image Team. Z-image: An efficient image generation foundation model with single-stream diffusion transformer. *arXiv preprint arXiv:2511.22699*, 2025.

Lvmin Zhang, Anyi Rao, and Maneesh Agrawala. Adding conditional control to text-to-image diffusion models. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 3836–3847, 2023.

Yunqing Zhao, Chao Du, Milad Abdollahzadeh, Tianyu Pang, Min Lin, Shuicheng Yan, and Ngai-Man Cheung. Exploring incompatible knowledge transfer in few-shot image generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7380–7391, 2023.

Bolei Zhou, Agata Lapedriza, Aditya Khosla, Aude Oliva, and Antonio Torralba. Places: A 10 million image database for scene recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017a.

Bolei Zhou, Hang Zhao, Xavier Puig, Sanja Fidler, Adela Barriuso, and Antonio Torralba. Scene parsing through ade20k dataset. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 633–641, 2017b.

Mingjian Zhu, Hanting Chen, Qiangyu Yan, Xudong Huang, Guanyu Lin, Wei Li, Zhijun Tu, Hailin Hu, Jie Hu, and Yunhe Wang. Genimage: A million-scale benchmark for detecting ai-generated image. *Advances in Neural Information Processing Systems*, 36:77771–77782, 2023.

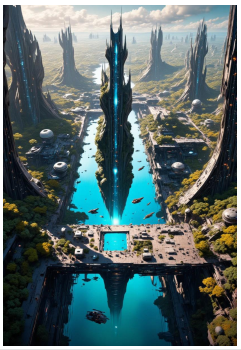
A Generated and Real Samples

In this appendix, we provide additional visual examples to illustrate the diversity of the WildGen. Figure 6 showcases images generated using complex synthesis pipelines, including high-fidelity models like SDXL and FLUX.dev, along with their detailed prompting and configuration metadata. Figure 7 presents examples of real images from our dataset, which serve as the authentic baseline for our binary classification and source-identification evaluations.

Pipe:	[Base] AAM XL (Anime Mix) · [LoRA] Perfect Eyes XL:0.65, Envy Mimic XL 01:0.35, extremely detailed (no trigger):0.65, Red Grunge:0.35, Midjourney mimic:0.65, Simon Stalenhag TFTL SDXL:0.45, Pink Pearl:0.35 · [Embeds] Ultimate Text Embeddings SDXL Pack
Prompt:	Cinematic. Dynamic. Ancient, rusty and tall mech warrior. Cut away diagram. Many details everywhere... [truncated]
Neg:	35mm film, deformed, glitch, noisy, signature, watermark. (worst quality, low quality:1.4), (easynegative, negative_hand-neg), (crown:1.5), horns, long ears, stars.
Cfg:	Steps: 50 · CFG: 7.0 · Sampler: Euler A

Pipe:	[Base] epiCRealism
Prompt:	score_9, score_8_up, score_7_up, realistic lighting, photo, photorealistic Blonde red hair 30 year old woman with green eyes, leaning out a window looking at the star field night sky. A lake, valley, cabin porch... [truncated]
Neg:	score_6, score_5, score_4, source_pony, source_anime, source_furry, source_cartoon, (worst quality:2), (low quality:2), (normal quality:2), (lowres, error), username, watermark, ugly face
Cfg:	Steps: 50 · CFG: 4.0 · Sampler: DPM++ 2M Karras

Pipe:	[Base] Pixel Art Diffusion XL
Prompt:	Gaulish warrior, full body, 16bit, in european forest
Neg:	bad quality, bad anatomy, worst quality, low quality, low resolution, extra fingers, blur, blurry, ugly, wrong proportions, watermark, image artifacts, lowres, ugly, jpeg artifacts, deformed, noisy image
Cfg:	Steps: 20 · CFG: 3.0 · Sampler: DPM++ 2M



Pipe:	[Base] FLUX dev · [LoRA] Velvet's Mythic Fantasy Styles:1.0, Sci-fi Environments:0.85, Artify's Fantastic Flux Landscape Lora:1.0
Prompt:	bustling sci-fi metropolis, chrome towers, a bridge running over a clear, fast-flowing large river with azure waters, mythpOrt, sunset, vivid vibrant colours, yellow vibrant trees, flying vehicles
Cfg:	Steps: 25 · CFG: 3.5

Figure 6: Examples of complex generation pipelines in the WildGen. Each example highlights the detailed configuration and prompting strategies used to synthesize high-quality images.

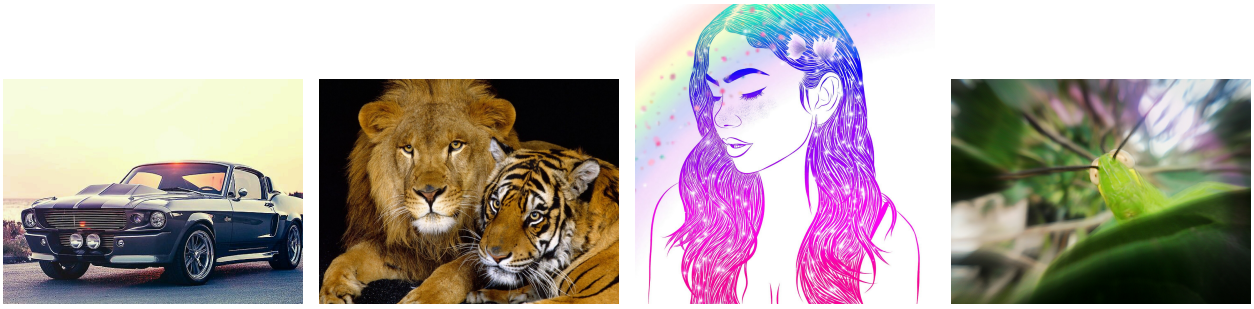


Figure 7: Examples of real (authentic) images in WildGen.

B Prompt Vocabulary Analysis

This appendix provides the prompt vocabulary analysis referenced in Section 3. The word cloud is metadata analysis only: evaluated detectors receive image pixels, not prompts, model names, source IDs, usernames, licenses, or other metadata.

Figure 8: Word frequency analysis of prompts in WildGen. The prominence of quality- and style-related modifiers reflects the benchmark's focus on high-engagement community generations.

C Frequency-Domain Case Study for RIGID

This appendix expands the selected RIGID failure discussed in Section 4. It should be read as a diagnostic case study rather than primary quantitative evidence. The main training-free results are the aggregate AP/AUC in Table 4, the source-level best/worst breakdown in Table 5, and the illustrative score-overlap example in Figure 4.

To provide a descriptive frequency-domain comparison alongside this selected RIGID failure, we examine one best-detected SD1.5-based source (qchan Mix v2.0) and one worst-detected SDXL-based source (Pixel Art Diffusion XL). We compute a canonical Real spectrum by averaging the log-magnitude FFTs of 100 randomly sampled real images after grayscale conversion and resizing to 512×512. We compute the corresponding average spectra for 100 images from each AI source and visualize AI-minus-Real residual maps.

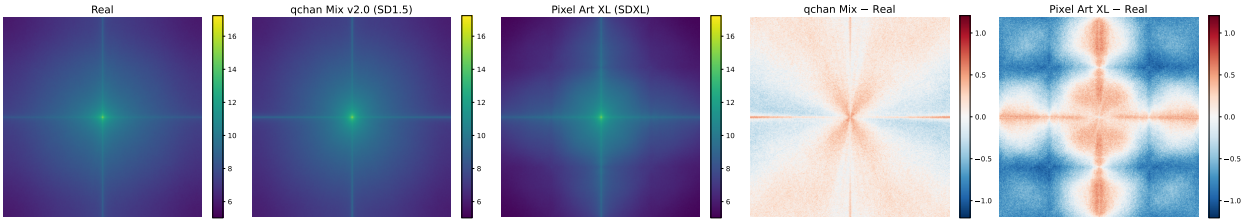


Figure 9: Illustrative frequency-domain comparison for one selected RIGID failure. The first three panels show average log-magnitude FFT spectra for Real images, qchan Mix v2.0, and Pixel Art Diffusion XL. The last two panels show AI-minus-Real residual maps. Although the poorly detected Pixel Art Diffusion XL source exhibits a structured spectral residual, this static frequency difference does not by itself explain RIGID performance, which depends on perturbation-induced feature stability.

In this selected comparison, qchan Mix v2.0 shows diffuse high-frequency residual energy, whereas Pixel Art Diffusion XL exhibits a more structured geometric residual pattern with a larger RMS difference from the Real spectrum. The latter nevertheless receives poor RIGID scores. This comparison illustrates that a visually prominent static spectral residual need not correspond to the perturbation-instability signal used by RIGID. We therefore do not interpret the residual map as evidence of spectral robustness or as a general explanation of detection difficulty; it is only a descriptive comparison accompanying the source-level RIGID result.

D Qualitative LMM Case Study

This appendix presents a qualitative case study of Large Multimodal Models (LMMs) for AI-image detection. It is not a systematic benchmark and should be read only as an illustration of how general-purpose assistants reason about provenance. We queried Gemini and ChatGPT in January 2026 using a sample image released alongside Z-Image (Z-Image Team, 2025) in December 2025 (Figure 10).

Figure 10: The sample image from Z-Image (December 2025) used to evaluate LMM detection capabilities.

D.1 Gemini Evaluation

We first queried Gemini. Its default response, which relies partly on digital watermark detection, was as follows:

> *Based on a digital watermark check, this image was not made with Google's AI tools.*
> *While it doesn't contain a Google AI watermark, it's difficult to say for certain if it was created using other AI platforms or if it's a real photograph. The lighting and textures, such as the knit of the scarf and the glaze on the croissant, look very natural. However, many modern AI generators are capable of producing highly realistic results.*

When explicitly instructed not to rely on built-in watermark detection tools, Gemini provided the following analysis:

> *It's difficult to say for certain if an image is AI-generated just by looking at it, but there are a few common signs you can check for:*
> ***Inconsistencies in Detail:*** *AI often struggles with complex details. Look closely at things like hands, eyes, and hair. You might see extra fingers, asymmetrical features, or hair that seems to blend into the skin.*
> ***Strange Textures:*** *Some AI images have a "too perfect" or plastic-like texture, while others may have blurry or smeared areas in the background that don't make sense.*
> ***Logical Errors:*** *Look for things that defy physics or logic, like objects melting into each other, shadows going in different directions, or text that is nonsensical or uses unrecognizable characters.*
> ***Metadata and Watermarks:*** *Sometimes the image file itself will have metadata indicating it was created with an AI tool, or there might be a small watermark from a service like DALL-E or Midjourney.*
> *If you have a high-resolution version of the image, these details are often easier to spot. You can also try using online AI detection tools, though they are not always 100% accurate.*

D.2 ChatGPT Evaluation

We also queried ChatGPT, which provided a detailed visual inspection:

> *I can't be 100% certain, but based on visual inspection alone, this image looks more like a real photograph than an AI-generated one.*
>
> *Here's why:*
>
> ***Natural lighting and reflections:*** *The light coming through the train window, the soft shadows, and the reflections on the glass look physically consistent and subtle in a way that current AI images often struggle with.*
>
> ***Material realism:*** *The textures of the scarf, paper bag, croissant layers, and coffee cup sleeve appear coherent and physically plausible, without the smeared or overly perfect details often seen in AI outputs.*
>
> ***Composition and imperfections:*** *The framing is slightly off-center and casual, like a candid travel photo, rather than overly stylized or symmetric.*
>
> ***No obvious AI artifacts:*** *I don't see common signs such as warped objects, inconsistent edges, strange text, or anatomically odd details.*
>
> *That said, modern AI models can produce images of this quality, especially for lifestyle scenes like this. Without metadata (EXIF), provenance, or known source context, the safest conclusion is:*
>
> ***Likely a real photo, but high-quality AI generation cannot be ruled out.***
>
> *If you want, I can: Point out specific AI telltale signs to double-check Analyze it assuming a particular model family (e.g., diffusion vs. GAN) Help assess likelihood using context such as where the image came from*

D.3 Analysis and Implications

These examples suggest that LMMs often reason through visible artifacts, metadata, and watermark availability, rather than applying a calibrated detector. This behavior can be useful for qualitative triage, but it is not a substitute for a controlled detection benchmark.

The case study is consistent with the main paper's broader message: provenance judgments depend on distributional context and available evidence. However, because this appendix contains only a small number of prompts and model responses, we do not use it as quantitative evidence for or against LMM-based detection. A systematic LMM evaluation would require fixed model versions, repeated prompts, a balanced image set, and a predefined scoring protocol.

E Real-Source Separability Breakdown

This appendix expands the real-source bias control in Section 5.4. We train a logistic-regression classifier on frozen CLIP ViT-L/14 features to predict which of the eight real-image sources an authentic image comes from, using 5,000 training and 1,000 validation images per source (seed 42). Because every source has equal validation support, balanced accuracy equals pooled accuracy; the observed 84.33% accuracy has an exact 95% Clopper–Pearson interval of 83.51–85.12%. Table 10 reports the per-source precision, recall, and F1.

The sources are far from equally separable. Curated, visually distinctive collections such as Food101 and Unsplash are almost perfectly recovered (99.50% and 95.50% recall), whereas broad web-photo sources overlap more: OpenImages (67.10% recall) and SUN397 (72.60%) are the hardest to distinguish. This confirms that "Real" is not a homogeneous background class, and that a detector trained on any single real source can inherit that source's idiosyncrasies rather than a general notion of authenticity.

Table 10: Per-source performance of the frozen-CLIP eight-way real-source classifier on the WildGen real validation split (1,000 images per source). Overall balanced accuracy is 84.33% (95% CI: 83.51–85.12).

Real Source	**Precision (%)**	**Recall (%)**	**F1 (%)**
Food101	96.88	99.50	98.17
Unsplash	91.30	95.50	93.35
Places365	89.61	92.30	90.94
Google Landmarks v2	83.69	86.20	84.93
ADE20K	81.36	81.60	81.48
YFCC/Flickr CC	83.21	79.80	81.47
SUN397	74.46	72.60	73.52
OpenImages	72.15	67.10	69.53
Macro average	84.08	84.33	84.17

F Generative Model Catalog

This appendix lists the 342 AI sources included in the main WildGen benchmark split. Each source has at least 1,000 indexed images; the split protocol selects 1,000 images per source, with 900 images assigned to the train pool and 100 images assigned to validation. We also report the source-level license name recorded during our Civitai audit. Model IDs 438–447 correspond to recently released generators and are included in this default split on the same terms as all other sources.

Table 11: AI sources with $\geq$ 1,000 indexed images included in the main WildGen split, with source-level license names.

ID	**Source Name**	**License**	**count**
1	+RealisticAnimeHentMix666+v1.0	CreativeML Open RAIL-M	1,415
2	287641	CreativeML Open RAIL-M	1,026
3	2DN-Pony	Illustrious License	3,218
4	2DN-Ponyv1	Illustrious License	1,854
6	[Pony] Geekpower AI Styles	CreativeML Open RAIL++-M	1,188
7	Meichidark_Mix	CreativeML Open RAIL++-M	1,166
8	AAA(AAAAAAAAAAAAAAAAAAAA) | Finetune mix on whatever model i want at that point which is Illustrious XL right now,but i will keep Pony one for on-site as well	CreativeML Open RAIL-M	3,498
10	AAM XL (Anime Mix)	CreativeML Open RAIL++-M	4,337
11	AAM XL (Anime Mix)v1.0	CreativeML Open RAIL++-M	3,480
12	AbsoluteReality	CreativeML Open RAIL-M	4,008
14	AbsoluteRealityv1.8.1	CreativeML Open RAIL-M	4,660

Continued on next page

Table 11, continued from previous page

ID	Source Name	License	count
15	AbyssOrangeMix2 - Hardcore	CreativeML Open RAIL-M	1,248
16	AbyssOrangeMix2 - HardcoreAbyssOrangeMix2_hard	CreativeML Open RAIL-M	2,140
17	AlbedoBase XL	CreativeML Open RAIL++-M	4,718
18	AlbedoBase XLv1.3	CreativeML Open RAIL++-M	2,097
19	AlbedoBase XLv2.0	CreativeML Open RAIL++-M	3,339
20	AlbedoBase XLv2.1	CreativeML Open RAIL++-M	4,112
22	Analog Madness - Realistic model	CreativeML Open RAIL-M	3,196
24	Analog Madness - Realistic modelv7.0	CreativeML Open RAIL-M	3,914
26	AniMerge	CreativeML Open RAIL-M	1,964
29	AniMesh	CreativeML Open RAIL-M	1,492
30	AniMeshAnimesh-Pruned V2.2	CreativeML Open RAIL-M	1,776
32	AniVerse	CreativeML Open RAIL-M	3,797
33	AniVerseV1.5 - Pruned	CreativeML Open RAIL-M	1,628
34	AniVerseV2.0 (HD) - Pruned	CreativeML Open RAIL-M	1,367
36	Animagine XL V3.1	CreativeML Open RAIL++-M	4,681
37	Animagine XL V3.1v3.0	CreativeML Open RAIL++-M	4,740
38	Animagine XL V3.1v3.1	CreativeML Open RAIL++-M	4,689
39	Anime Illust Diffusion XL	CreativeML Open RAIL++-M	1,723
41	Anime Model OtakuReliableEnable (AMORE)	CreativeML Open RAIL++-M	2,378
42	AnyLoRA - Checkpoint	CreativeML Open RAIL-M	3,694
43	AnyLoRA - CheckpointbakedVae (blessed) fp16 NOT-PRUNED	CreativeML Open RAIL-M	4,565
44	AnyLoRA - CheckpointbakedVae ftMse fp16 NOT-PRUNED	CreativeML Open RAIL-M	3,496
45	AnyLoraCleanLinearMix-ClearVAE	CreativeML Open RAIL-M	4,028
46	AnyLoraCleanLinearMix-ClearVAEv1.0	CreativeML Open RAIL-M	4,883
47	AnyOrangeMix - Anything + AbyssOrangeMix	CreativeML Open RAIL-M	1,962
48	AnythingElse V4v4.5	CreativeML Open RAIL-M	2,383
49	Arthemy Comics	CreativeML Open RAIL-M	2,246
51	Artiumv2.0	CreativeML Open RAIL++-M	1,304
53	AutismMix SDXL	CreativeML Open RAIL++-M	4,739
55	AutismMix SDXLAutismMix_confetti	CreativeML Open RAIL++-M	4,919
56	AutismMix SDXLAutismMix_pony	CreativeML Open RAIL++-M	4,844
57	BB95 Furry Mix	CreativeML Open RAIL-M	3,496
58	Beautiful Realistic Asiansv7	CreativeML Open RAIL-M	1,042
59	Better than words	CreativeML Open RAIL++-M	3,243
60	Better than wordsv3.0	CreativeML Open RAIL++-M	1,232
61	BlenderMix	CreativeML Open RAIL++-M	1,744
62	Boltning - Realistic, Lightning, HYPERHYPER_D	CreativeML Open RAIL++-M	1,030
64	BriXL | A must in your toolbox	CreativeML Open RAIL++-M	1,562
66	CATalyst - Citron Anime Treasure Catalystv1.0	CreativeML Open RAIL-M	1,032
67	CG Realistic PonyV5	Illustrious License	3,292
68	CalicoMix	CreativeML Open RAIL-M	3,353
71	CashMoney (Anime)	CreativeML Open RAIL++-M	1,803
72	Cherry Picker XL	CreativeML Open RAIL++-M	1,592
73	ChilloutMixChilloutmix-Ni-pruned-fp32-fix	CreativeML Open RAIL-M	4,249
74	CineVisionXL by SoCalGuitaristRelease-v1.5.0-BakedVAE	CreativeML Open RAIL++-M	1,292
75	Cinenauts XL (A TRUE CINEMATIC DIFFUSION)	CreativeML Open RAIL++-M	1,957
76	Cinenauts XL (A TRUE CINEMATIC DIFFUSION)Cinenauts v1.0	CreativeML Open RAIL++-M	3,421

Continued on next page

Table 11, continued from previous page

ID	Source Name	License	count
77	Cinenauts XL (A TRUE CINEMATIC DIFFUSION)Cinenauts v2.0	CreativeML Open RAIL++-M	2,001
78	CoffeeBreakv2.0	CreativeML Open RAIL-M	1,426
79	Copax TimeLessXL	CreativeML Open RAIL++-M	3,517
80	Copax TimeLessXLV8	CreativeML Open RAIL++-M	2,319
83	Counterfeit-V3.0	CreativeML Open RAIL-M	3,811
84	Counterfeit-V3.0v3.0	CreativeML Open RAIL-M	4,258
85	Crystal Clear One_vS1	CreativeML Open RAIL++-M	1,434
86	Crystal Clear XL	CreativeML Open RAIL++-M	3,926
87	Crystal Clear XL - Primev1.0	CreativeML Open RAIL++-M	1,852
88	Crystal Clear XLCCXL	CreativeML Open RAIL++-M	4,442
89	CyberRealistic	CreativeML Open RAIL-M	4,402
90	CyberRealistic Pony	CreativeML Open RAIL++-M	3,191
91	CyberRealistic Ponyv6.3	CreativeML Open RAIL++-M	1,114
92	CyberRealisticv3.3	CreativeML Open RAIL-M	3,328
94	CyberRealisticv5.0	CreativeML Open RAIL-M	4,191
95	CyberSushi	CreativeML Open RAIL-M	1,798
96	CyberSushiv2.1	CreativeML Open RAIL-M	1,040
97	DENRealMergev1.0	CreativeML Open RAIL-M	4,019
98	Dark Sushi 2.5D 2.5D	CreativeML Open RAIL-M	1,851
99	Dark Sushi 2.5D 2.5Dv4.0	CreativeML Open RAIL-M	1,201
100	Dark Sushi Mix Mix	CreativeML Open RAIL-M	3,762
101	Dark Sushi Mix Mix2.25D	CreativeML Open RAIL-M	3,545
102	DarkSunv4.1	CreativeML Open RAIL-M	1,690
103	DevlishPhotoRealism SDXL	CreativeML Open RAIL++-M	2,009
105	DisneyRealCartoonMix	CreativeML Open RAIL++-M	2,245
106	Divine Diffusion by Dever (DDD) [SDXL]	CreativeML Open RAIL++-M	2,514
109	Doll-E XLv1	CreativeML Open RAIL++-M	1,721
110	DreamShaper	CreativeML Open RAIL-M	4,832
111	DreamShaper XL	CreativeML Open RAIL++-M	4,660
112	DreamShaper XLLightning DPM++ SDE	CreativeML Open RAIL++-M	2,097
114	DreamShaper XLalpha2 (xl1.0)	CreativeML Open RAIL++-M	4,698
115	DreamShaper XLv2.1 Turbo DPM++ SDE	CreativeML Open RAIL++-M	2,873
117	DreamShaper8	CreativeML Open RAIL-M	3,854
118	DucHaiten-AIart-SDXL	CreativeML Open RAIL++-M	3,312
119	DucHaiten-AIart-SDXLv2.0	CreativeML Open RAIL++-M	1,119
120	DucHaiten-AIart-SDXLv3.3.5.1.5	CreativeML Open RAIL++-M	1,167
121	DucHaiten-Pony-Real	CreativeML Open RAIL++-M	3,805
122	DucHaiten-Pony-XL (no-score)	CreativeML Open RAIL++-M	4,613
124	DucHaiten-Pony-XL (no-score)pony-no-score_v4.0	CreativeML Open RAIL++-M	2,831
126	DucHaiten-Pony-XL (no-score)v5.2	CreativeML Open RAIL++-M	1,456
129	DynaVision XL - All-in-one stylized 3D SFW and NSFW output, no refiner needed!	CreativeML Open RAIL++-M	4,032
130	DynaVision XL - All-in-one stylized 3D SFW and NSFW output, no refiner needed!Release_0.5.5.7.BakedVAE	CreativeML Open RAIL++-M	2,082
131	DynaVision XL - All-in-one stylized 3D SFW and NSFW output, no refiner needed!Release_v0.6.1.0-bakedvae	CreativeML Open RAIL++-M	1,589
136	Ether PDXL	CreativeML Open RAIL++-M	2,006
137	Everclear PNY by Zovya	CreativeML Open RAIL++-M	3,170
138	FLUX	FLUX.1 [dev] Non-Commercial License	4,886
139	FLUX.1 [dev]v1.0	FLUX.1 [dev] Non-Commercial License	2,645
140	FLUX.1DEV FP8 - Kijai [11 GB]	FLUX.1 [dev] Non-Commercial License	1,683
141	Fantasm | Pony XL	CreativeML Open RAIL++-M	1,453

Continued on next page

Table 11, continued from previous page

ID	Source Name	License	count
142	FenrisXL + Flux	FLUX.1 [dev] Non-Commercial License	3,493
143	Flat-2D Animerge	CreativeML Open RAIL-M	1,692
144	ForArtisticXLv0.1	CreativeML Open RAIL++-M	1,342
145	ForRealXL V1.0v0.1	CreativeML Open RAIL++-M	2,208
146	ForRealXL V1.0v0.3	CreativeML Open RAIL++-M	1,080
147	ForRealXL V1.0v0.5	CreativeML Open RAIL++-M	3,880
151	FurryBlend	CreativeML Open RAIL-M	3,130
156	GhostMix	CreativeML Open RAIL-M	3,418
157	GhostMixv2.0-BakedVAE	CreativeML Open RAIL-M	4,414
158	Hassaku (hentai model)	Anima	4,511
159	Hassaku (hentai model)v1.3	Anima	4,088
160	Hassaku XL (Hentai)	CreativeML Open RAIL++-M	4,156
161	Hassaku XL (Hentai)v1.3	CreativeML Open RAIL++-M	3,075
162	Hassaku XL (Hentai)v1.3(better eyes version)	CreativeML Open RAIL++-M	1,437
166	Honey Mix [XL] High contrast anime checkpoint	CreativeML Open RAIL++-M	1,190
167	Hoseki - LustrousMix [Pony XL]	CreativeML Open RAIL++-M	2,792
168	Hoseki - LustrousMix [Pony XL]v1.0	CreativeML Open RAIL++-M	1,743
170	ICBINP - "I Can't Believe It's Not Photography"Final	CreativeML Open RAIL-M	1,147
172	Incursio's Meme Diffusion (SDXL, Pony)	CreativeML Open RAIL++-M	4,542
173	Incursio's Meme Diffusion (SDXL, Pony)v1.6PDXL	CreativeML Open RAIL++-M	4,476
174	Incursio's Meme Diffusionv1.6 Pruned	CreativeML Open RAIL-M	1,022
175	Indigo Furry mix	CreativeML Open RAIL-M	4,074
180	Jib Mix Realistic XL	CreativeML Open RAIL++-M	4,211
181	Jib Mix Realistic XLv9.0 Better Bodies	CreativeML Open RAIL++-M	1,933
182	Juggernaut	CreativeML Open RAIL-M	3,323
183	Juggernaut XL	CreativeML Open RAIL++-M	4,882
184	Juggernaut XLJugg_XI_by_RunDiffusion	CreativeML Open RAIL++-M	1,051
185	Juggernaut XLJugg_X_by RunDiffusion	CreativeML Open RAIL++-M	4,026
186	Juggernaut XLV 7 + RunDiffusion	CreativeML Open RAIL++-M	4,317
187	Juggernaut XLV 8 + RunDiffusion	CreativeML Open RAIL++-M	4,300
188	Juggernaut XLV9 + RunDiffusionPhoto 2	CreativeML Open RAIL++-M	4,188
189	Juggernaut XLV9+RDPhoto2-Lightning_4S	CreativeML Open RAIL++-M	3,261
190	Juggernaut XLVersion 6 + RunDiffusion	CreativeML Open RAIL++-M	4,728
191	JuggernautReborn	CreativeML Open RAIL-M	2,326
192	Kohaku-XL beta	CreativeML Open RAIL++-M	1,444
193	Kohaku-XL betabeta7	CreativeML Open RAIL++-M	2,399
194	Koji	CreativeML Open RAIL-M	3,305
195	Kojiv2.1	CreativeML Open RAIL-M	1,891
196	LEOSAM's HelloWorld XL	CreativeML Open RAIL++-M	4,001
197	LEOSAM's HelloWorld XLHelloWorld XL 5.0 GPT4V	CreativeML Open RAIL++-M	1,881
198	LEOSAM's HelloWorld XLHelloWorld XL 6.0	CreativeML Open RAIL++-M	1,346
199	LEOSAM's HelloWorld XLHelloWorld XL 7.0	CreativeML Open RAIL++-M	2,932
201	LazyMix+ (Real Amateur Nudes)	CreativeML Open RAIL-M	1,146
202	Luminaverse [Pony XL]	CreativeML Open RAIL++-M	4,177
203	Luminaverse [Pony XL]v1.0	CreativeML Open RAIL++-M	1,253
206	MFCG PDXL	Apache 2.0	2,179
207	MFCG PDXLv1.0	CreativeML Open RAIL++-M	2,617
208	MHXL - Aventis Horizon (SDXL)	CreativeML Open RAIL++-M	1,307
209	MeichiDarkMix_Reload	CreativeML Open RAIL++-M	1,955
210	MeinaHentai	CreativeML Open RAIL-M	3,664
211	MeinaHentaiV4	CreativeML Open RAIL-M	2,843
212	MeinaMix	CreativeML Open RAIL-M	4,325

Continued on next page

Table 11, continued from previous page

ID	Source Name	License	count
213	MeinaMixMeina V11	CreativeML Open RAIL-M	4,851
214	MeinaPastel	CreativeML Open RAIL-M	2,365
215	MeinaPastelV6 (Pastel)	CreativeML Open RAIL-M	1,907
216	MeinaUnreal	CreativeML Open RAIL-M	1,017
217	MeinaUnrealV4.1	CreativeML Open RAIL-M	2,218
218	Mistoon_Anime	Illustrious License	3,571
220	Moxie Diffusion XLv1.32	CreativeML Open RAIL++-M	1,010
223	NeverEnding Dream (NED)	CreativeML Open RAIL-M	2,535
225	NightVisionXL	CreativeML Open RAIL++-M	4,578
226	NightVisionXLNightVisionXL_v9.0.0	CreativeML Open RAIL++-M	1,902
227	NightVisionXLRelease_0.7.7.0.BakedVAE	CreativeML Open RAIL++-M	1,538
228	NightVisionXLV0.7.9.1.BakedVAE	CreativeML Open RAIL++-M	2,896
229	NightVisionXLv0.8.1.1	CreativeML Open RAIL++-M	1,706
231	ONE FOR ALL Pony Fantasy DPO+VAE	CreativeML Open RAIL++-M	2,340
233	Objective Reality (Reloaded)v2.0 (original VAE)	CreativeML Open RAIL-M	2,802
235	Omnium	CreativeML Open RAIL++-M	2,472
236	Omniumv1.1	CreativeML Open RAIL++-M	1,080
238	Painter's Checkpoint (oil paint / oil painting art style)	CreativeML Open RAIL++-M	1,252
239	Paradox 2 SD XL 1.0Paradox 2 SD XL 1.0	CreativeML Open RAIL++-M	1,082
241	Perfect World	CreativeML Open RAIL-M	3,267
243	PerfectDeliberate	Illustrious License	1,422
244	PhotoPedia XL	CreativeML Open RAIL++-M	2,962
245	PhotoPedia XL4.5	CreativeML Open RAIL++-M	3,183
246	Photon	CreativeML Open RAIL-M	1,489
247	Photonv1	CreativeML Open RAIL-M	1,843
248	PicX_real	CreativeML Open RAIL-M	3,433
249	PicX_real1.0	CreativeML Open RAIL-M	3,981
250	Pixel Art Diffusion XL	CreativeML Open RAIL++-M	1,645
251	Pony Diffusion V6 XL	CreativeML Open RAIL++-M	4,944
252	Pony Diffusion V6 XLV6 (start with this one)	CreativeML Open RAIL++-M	4,652
254	Pony Realism	CreativeML Open RAIL++-M	4,490
255	Pony Realism v2.1 Main + VAE	CreativeML Open RAIL++-M	3,943
256	Pony FaeTality	CreativeML Open RAIL++-M	4,372
257	Pony FaeTalityv1.0	CreativeML Open RAIL++-M	2,088
258	Pony FaeTalityv1.1	CreativeML Open RAIL++-M	3,680
260	Prefect Pony XL	CreativeML Open RAIL++-M	4,569
261	Prefect Pony XLv1.0	CreativeML Open RAIL++-M	2,687
262	Prefect Pony XLv2 cleaned style	CreativeML Open RAIL++-M	2,622
263	Prefect Pony XLv3	CreativeML Open RAIL++-M	1,411
265	ProtoVision XL - High Fidelity 3D / Photorealism / Anime / hyperrealism - No Refiner Needed	CreativeML Open RAIL++-M	4,307
266	ProtoVision XL - High Fidelity 3D / Photorealism / Anime / hyperrealism - No Refiner NeededRelease_0.6.3.0.BakedVAE	CreativeML Open RAIL++-M	1,172
267	ProtoVision XL - High Fidelity 3D / Photorealism / Anime / hyperrealism - No Refiner NeededRelease_v6.6.0-BakedVAE	CreativeML Open RAIL++-M	1,917
270	RPG	CreativeML Open RAIL-M	1,630
272	Raemora XLv1.0	CreativeML Open RAIL++-M	1,134
274	ReV Animated	CreativeML Open RAIL-M	4,250
275	ReV AnimatedV2 Rebirth	CreativeML Open RAIL-M	1,164
276	ReV Animatedv1.2.2-EOL	CreativeML Open RAIL-M	4,804
277	Real Dream	CreativeML Open RAIL++-M	3,522
278	RealCartoon-Pixar	CreativeML Open RAIL-M	1,829

Continued on next page

Table 11, continued from previous page

ID	Source Name	License	count
279	RealCartoon-Realistic	CreativeML Open RAIL-M	3,152
280	RealCartoon-XL	CreativeML Open RAIL++-M	4,557
281	RealCartoon-XLV5	CreativeML Open RAIL++-M	1,078
282	RealCartoon-XLV6	CreativeML Open RAIL++-M	3,454
283	RealCartoon3D	CreativeML Open RAIL-M	3,634
284	RealCartoon3DV15	CreativeML Open RAIL-M	1,126
285	RealVisXL V5.0	CreativeML Open RAIL++-M	4,615
286	RealVisXL V5.0V2.0 (BakedVAE)	CreativeML Open RAIL++-M	2,321
287	RealVisXL V5.0V3.0 (BakedVAE)	CreativeML Open RAIL++-M	2,300
288	RealVisXL V5.0V4.0 (BakedVAE)	CreativeML Open RAIL++-M	4,164
290	Realisian	CreativeML Open RAIL-M	3,057
291	Realism Engine SDXL	Apache 2.0	3,274
292	Realism Engine SDXLv3.0 VAE	CreativeML Open RAIL++-M	2,634
293	Realistic Stock Photov2.0	CreativeML Open RAIL-M	1,200
294	Realistic Vision V6.0 B1	CreativeML Open RAIL++-M	4,303
295	Realistic Vision V6.0 B1V5.1 (VAE)	CreativeML Open RAIL++-M	4,783
296	Realistic Vision V6.0 B1V5.1 Hyper (VAE)	CreativeML Open RAIL++-M	1,512
299	Rev Engine PonyXLv1.0	CreativeML Open RAIL++-M	1,523
300	SD XL	CreativeML Open RAIL++-M	4,842
301	SD XLv1.0	CreativeML Open RAIL++-M	4,907
302	SD XLv1.0 VAE fix	CreativeML Open RAIL++-M	4,481
303	SDVN6-RealXL	CreativeML Open RAIL++-M	1,903
304	SDVN6-RealXLDetailface	CreativeML Open RAIL++-M	2,047
305	SDVN7-NijiStyleXL	CreativeML Open RAIL++-M	3,829
306	SDVN7-NijiStyleXLv1	CreativeML Open RAIL++-M	1,347
308	SDXL Unstable Diffusers YamerMIX	CreativeML Open RAIL++-M	4,691
309	SDXL Unstable Diffusers YamerMIXNihilMania	CreativeML Open RAIL++-M	1,762
310	SDXL Unstable Diffusers YamerMIXV11 + RunDiffusion	CreativeML Open RAIL++-M	2,841
312	SDXL Unstable Diffusers YamerMIXV8 Heaven's Wrath+VAE	CreativeML Open RAIL++-M	4,039
313	SDXL Yamer's Anime Unstable Illustrator	CreativeML Open RAIL++-M	2,902
314	SDXL Yamer's Realism! - Realistic/Anime/3D	CreativeML Open RAIL++-M	1,896
315	SDXL Yamer's Realistic 5	CreativeML Open RAIL++-M	4,400
316	SDXL Yamer's Realistic 5 V5 + RunDiffusion	CreativeML Open RAIL++-M	1,854
317	SDXL_Niji_Seven	CreativeML Open RAIL++-M	4,174
318	SDXL_Niji_SevenSDXL_Niji_SE	CreativeML Open RAIL++-M	1,980
319	SDXL_Niji_SevenSDXL_Niji_Seven	CreativeML Open RAIL++-M	1,048
320	SDXL_Niji_SevenSDXL_Niji_v6	CreativeML Open RAIL++-M	1,220
321	SDXXXL	CreativeML Open RAIL++-M	1,582
322	SVSN001v1.0	CreativeML Open RAIL++-M	4,775
323	SereneXL	CreativeML Open RAIL++-M	1,605
324	SkibidiMix	CreativeML Open RAIL++-M	1,777
325	SoloMixv1.0	CreativeML Open RAIL-M	4,532
326	Starlight XL Animated	CreativeML Open RAIL++-M	4,507
327	Starlight XL Animatedv3	CreativeML Open RAIL++-M	3,771
328	Sym-Pony World	CreativeML Open RAIL++-M	2,818
329	Sym-Pony Worldv1.0	CreativeML Open RAIL++-M	1,109
330	T-ponynai3	CreativeML Open RAIL++-M	2,278
332	T-ponynai3v6.5	CreativeML Open RAIL++-M	1,028
334	TalmendoXL - SDXL Uncensored Full Model	CreativeML Open RAIL++-M	1,148
335	Tamarin_XLv1.0	CreativeML Open RAIL++-M	3,889
336	The Truality Engine	CreativeML Open RAIL-M	1,270
338	The Truality EngineThe Truality Engine V3	CreativeML Open RAIL-M	1,680
340	Toon Babes	CreativeML Open RAIL-M	1,419

Continued on next page

Table 11, continued from previous page

ID	Source Name	License	count
341	Toon Babesv1.0	CreativeML Open RAIL-M	4,151
342	ToonYou	CreativeML Open RAIL-M	2,497
344	Toonify	Illustrious License	3,060
348	Unleashed DiffusionWhimsical	CreativeML Open RAIL-M	1,717
349	Virtual Diffusion Pony XL	CreativeML Open RAIL++-M	1,616
350	WAI-ANI-NSFW-PONYXL	CreativeML Open RAIL++-M	3,600
351	WAI-ANI-NSFW-PONYXLv7.0	CreativeML Open RAIL++-M	2,182
352	WAI-ANI-NSFW-PONYXLv8.0	CreativeML Open RAIL++-M	1,394
355	WildCardX-XL LIGHTNINGWildCardX-XL+Rundiffusion	CreativeML Open RAIL++-M	2,603
356	WildCardX-XL-Fusion	CreativeML Open RAIL++-M	4,731
358	WildCardX-XL-FusionFUSION OG	CreativeML Open RAIL++-M	3,347
359	WildCardX-XLV4+RunDiffusion	CreativeML Open RAIL++-M	3,096
360	WoW_(XL+PD+Flux).	FLUX.1 [dev] Non-Commercial License	1,517
361	WoW_(XL+PD+Flux).v2	FLUX.1 [dev] Non-Commercial License	3,851
362	XL6 - HEPHAISTOS SD 1.0XL (SFW&NSFW)	CreativeML Open RAIL++-M	3,052
363	XXMix_9realisticSDXLv1.0	CreativeML Open RAIL++-M	2,182
365	YiffyMix	CreativeML Open RAIL-M	4,004
366	YiffyMixv36	CreativeML Open RAIL-M	1,038
367	ZavyChromaXL	CreativeML Open RAIL++-M	4,810
368	ZavyChromaXLv2.1	CreativeML Open RAIL++-M	3,391
369	ZavyChromaXLv3.0	CreativeML Open RAIL++-M	3,655
370	ZavyChromaXLv4.0	CreativeML Open RAIL++-M	2,612
371	ZavyChromaXLv5.0	CreativeML Open RAIL++-M	3,534
372	ZavyChromaXLv6.0	CreativeML Open RAIL++-M	3,643
373	ZavyChromaXLv7.0	CreativeML Open RAIL++-M	3,843
374	ZavyChromaXLv8.0	CreativeML Open RAIL++-M	3,708
375	ZavyChromaXLv9.0	CreativeML Open RAIL++-M	4,137
377	Zaxious_XLv2.0	CreativeML Open RAIL++-M	1,158
378	[Lah] Mysterious | Flux update	Apache 2.0	3,822
379	[Lah] Mysterious | Flux updatev4.0	Apache 2.0	1,523
380	[PVC Style Model]Movable figure model Pony	CreativeML Open RAIL++-M	2,209
381	_CHEYENNE_	CreativeML Open RAIL++-M	3,716
382	_CHEYENNE_v1.6	CreativeML Open RAIL++-M	1,565
384	blue_pencil-XL	CreativeML Open RAIL++-M	1,752
385	epiCPhotoGasm	CreativeML Open RAIL-M	3,903
386	epiCPhotoGasmLast Unicorn	CreativeML Open RAIL-M	4,108
387	epiCPhotoGasmUltimate Fidelity	CreativeML Open RAIL-M	1,713
388	epiCPhotoGasmZ - Universal	CreativeML Open RAIL-M	2,423
389	epiCRealism	CreativeML Open RAIL-M	4,078
390	epiCRealism XL	CreativeML Open RAIL++-M	3,189
392	epiCRealismNatural Sin RC1 VAE	CreativeML Open RAIL-M	3,504
394	hash_	FLUX.1 [dev] Non-Commercial License	4,112
395	hash_0685a29800	FLUX.1 [dev] Non-Commercial License	1,476
396	hash_06f96f89f6	FLUX.1 [dev] Non-Commercial License	1,545
397	hash_275ef623d3	FLUX.1 [dev] Non-Commercial License	3,236
398	hash_3f97fdc57a	FLUX.1 [dev] Non-Commercial License	2,085
404	hash_bea01d51bd	FLUX.1 [dev] Non-Commercial License	3,528
405	hash_c161224931	FLUX.1 [dev] Non-Commercial License	1,665
407	hash_d50bfd647c	CreativeML Open RAIL++-M	1,586
408	hash_de5612ec57	CreativeML Open RAIL-M	1,154
412	majicMIX realistic	CreativeML Open RAIL-M	1,255
414	majicMIX realistic v7	CreativeML Open RAIL-M	4,618
417	qchan Mixv2.0	CreativeML Open RAIL-M	2,060
418	qchan realistic Mixv2.0	CreativeML Open RAIL-M	1,888

Continued on next page

Table 11, continued from previous page

ID	Source Name	License	count
419	reweik_Uni020v20-1	CreativeML Open RAIL-M	1,097
420	richyrichMix	CreativeML Open RAIL-M	4,002
422	richyrichMixrichyrichMix-v2.fp16	CreativeML Open RAIL-M	4,255
423	seizaMixv2	CreativeML Open RAIL-M	1,402
424	supashymixv3.0-lite	CreativeML Open RAIL-M	1,105
425	Realities Edge XL LIGHTNING + Turbo!	CreativeML Open RAIL++-M	1,323
426	CheckpointYesMixv1.6 (original)	CreativeML Open RAIL-M	1,247
427	| Anything XL	CreativeML Open RAIL++-M	3,436
428	| Anything XLXL	CreativeML Open RAIL++-M	1,110
429	STOIQO NewReality | FLUX, SD, XL, Lightning	CreativeML Open RAIL++-M	4,091
430	STOIQO NewReality | FLUX, SD, XL, Lightning XL 1.3	CreativeML Open RAIL++-M	1,272
432	STOIQO NewReality | FLUX, SD, XL, Lightning XL 2.1	CreativeML Open RAIL++-M	2,598
433	STOIQO NewReality | FLUX, SD, XL, Lightning XL 4.0	CreativeML Open RAIL++-M	2,174
434	SDXL FaeTastic	CreativeML Open RAIL++-M	4,511
435	SDXL FaeTasticv1.6	CreativeML Open RAIL++-M	2,389
436	SDXL FaeTasticv2.0	CreativeML Open RAIL++-M	2,518
437	SDXL FaeTasticv24	CreativeML Open RAIL++-M	4,176
438	OpenAI 4o Image Gen 1	OpenAI	1,000
439	Imagen 4	Imagen4	1,000
440	Nano Banana Pro	Imagen4	1,000
441	Flux.2 Dev	FLUX.1 [dev] Non-Commercial License	1,000
442	Z Image Turbo	Apache 2.0	1,000
443	Flux.2 Klein 9B-base	FLUX.1 [dev] Non-Commercial License	1,000
444	Flux.2 Klein 9B	FLUX.1 [dev] Non-Commercial License	1,000
445	Z Image Base	Apache 2.0	1,000
446	Nano Banana 2	Imagen4	1,000
447	Nano Banana	MIT	1,000