



4KAgent:

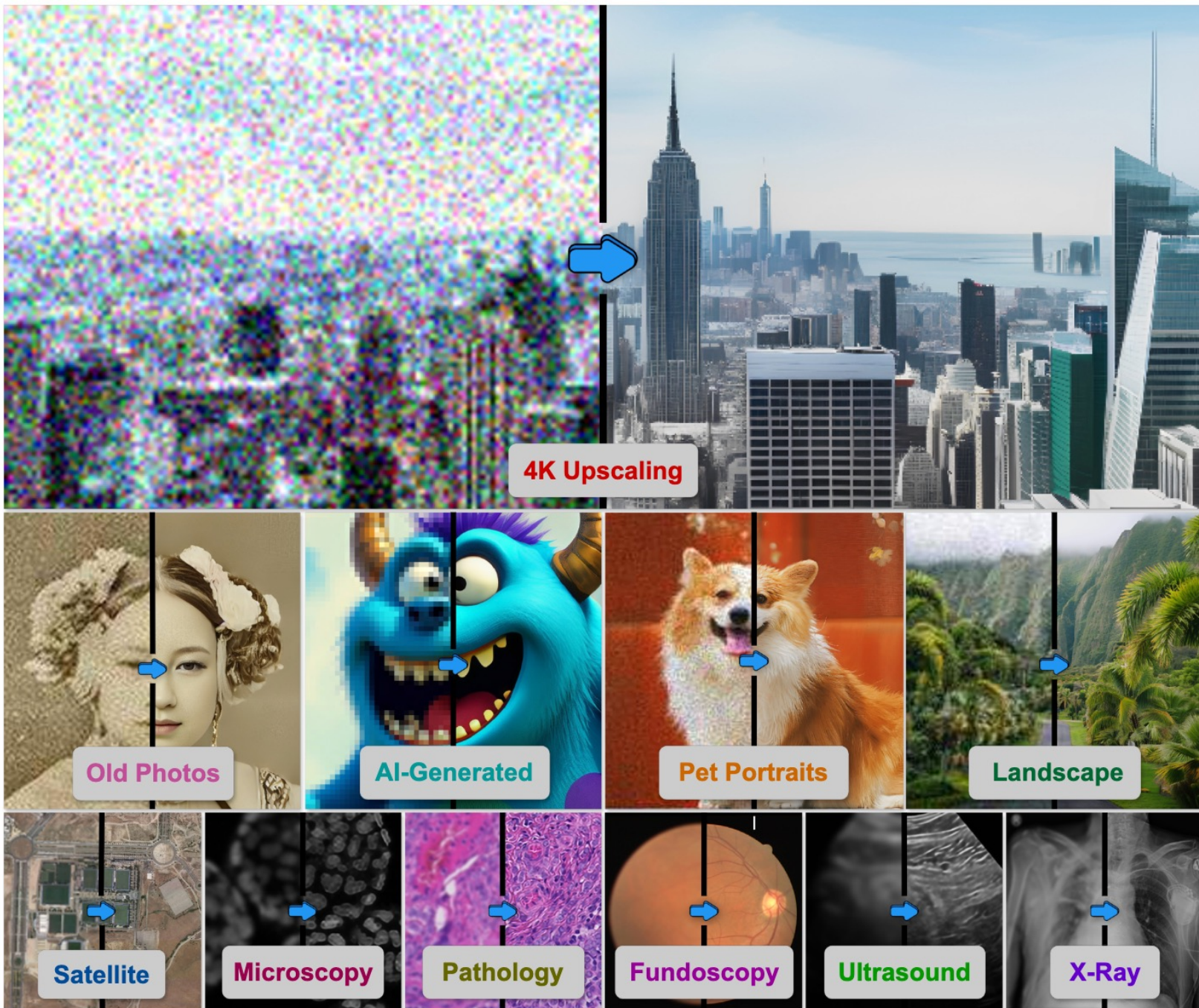
Agentic Any Image to 4K Super-Resolution

**Yushen Zuo¹, Qi Zheng^{1†}, Mingyang Wu^{1†}, Xinrui Jiang^{2†}, Renjie Li¹,
Jian Wang³, Yide Zhang⁴, Gengchen Mai⁵, Lihong V. Wang⁶, James Zou²,
Xiaoyu Wang⁷, Ming-Hsuan Yang⁸, Zhengzhong Tu^{1*}**

¹Texas A&M University ²Stanford University ³Snap Inc. ⁴CU Boulder
⁵UT Austin ⁶California Institute of Technology ⁷Topaz Labs ⁸UC Merced

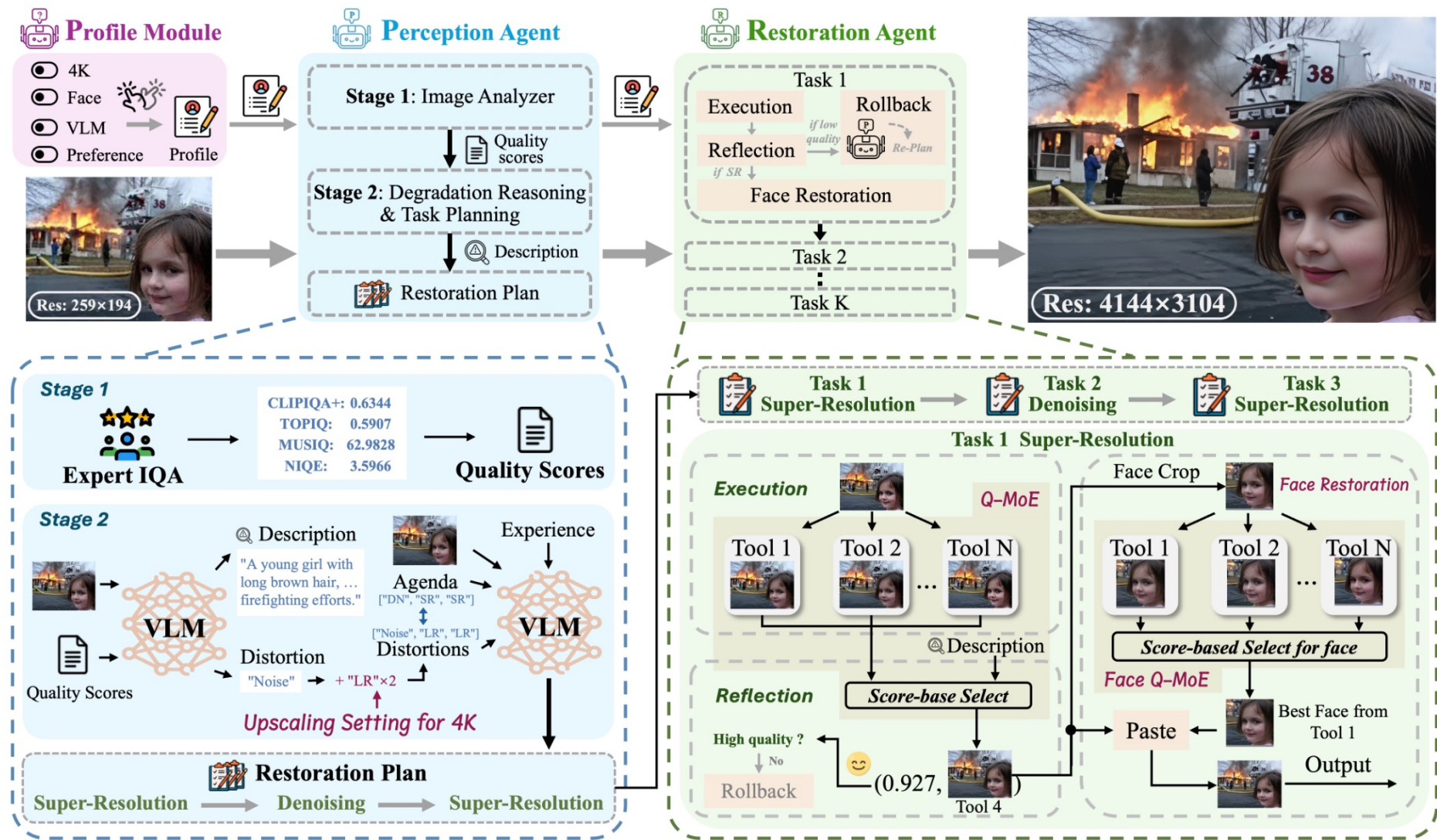
*Corresponding Author: tzz@tamu.edu. [†]Equal contributions.

Project Website: 4kagent.github.io



We present **4KAgent**, an agentic image super-resolution generalist designed to universally upscale any image to 4K, regardless of input type, degradation level, or domain. That is, 4KAgent effectively restores diverse imagery, spanning from natural scenes, severely degraded captures (e.g., old photos), human/pet portraits, AI-generated content (AIGC), as well as specialized scientific imaging domains, such as remote sensing, fluorescence microscopy, pathology, and various medical modalities like X-ray, ultrasound, and funduscopy—all **without** the need for any re-training or domain-specific adaptation.

4KAgent: System Overview

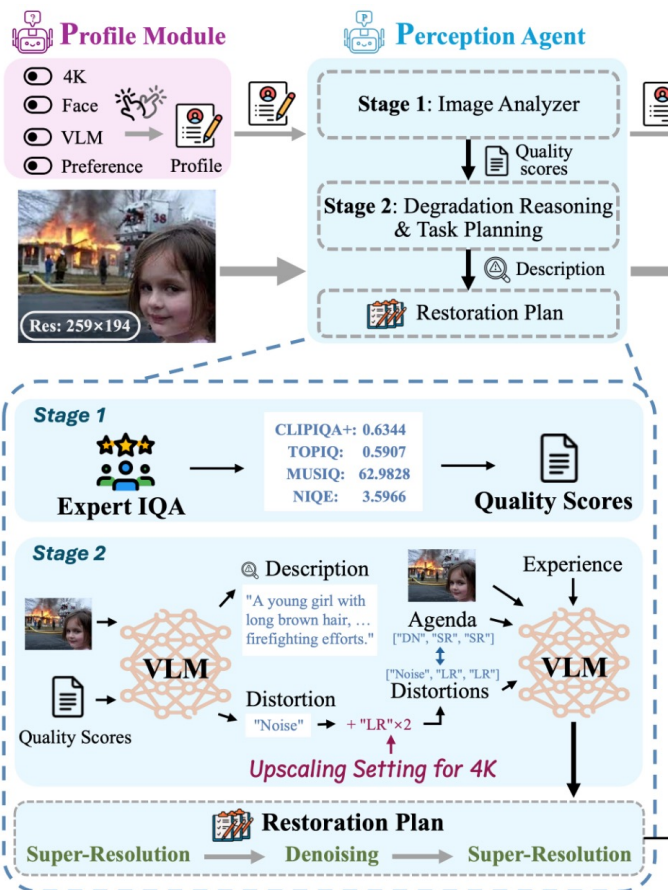


Perception Agent:
Analyze Image and propose a restoration plan.

Restoration Agent:
Execute restoration tasks based on restoration plan.

Profile Module:
Customize 4KAgent for different restoration / enhancement requirement.

4KAgent: Perception Agent



Input image



Resolution: 259 x 194

“Propose a restoration plan to restore this image and upscale it to 4K resolution.”

(1) Obtain IQA metrics based on the input image (Quality Scores)

CLIPQA+: 0.6344, TOPIQ_NR: 0.5907, MUSIQ: 62.9828, NIQE: 3.5966

(2) Input image & Quality scores to Reasoning Expert (VLM) for content & degradation analysis. (Description, Distortion -> initial restoration agenda)

Description: “A young girl with long brown hair, wearing a pink shirt, looks directly at the camera. The background shows a large fire burning outside a building marked '38'. A hose is visible, likely being used for firefighting efforts.”

Distortion: [‘noise’]

Initial Restoration Agenda: [‘Denoising’]

(3) Upscaling Factor Configuration: Calculate the minimum scale factor to upscale this image to 4K resolution. (The larger side ≥ 4000 pixels) Append super-resolution task to the initial restoration agenda.

$\max(259, 194) = 259$

$259 \times 16 = 4144 > 4000 \rightarrow \text{scale factor} = 16 \rightarrow \text{two ‘Super-resolution’ task with } 4\times$

Restoration Agenda: [‘Denoising’, ‘Super-resolution’, ‘Super-resolution’]

(4) Propose a restoration plan by Planning Expert (VLM / LLM) based on image description, distortions in the image, restoration agenda, restoration experience, and (optional) image itself.

“Plan”: [“Super-resolution”, “Denoising”, “Super-resolution”]



4KAgent: Restoration Agent - Task 1

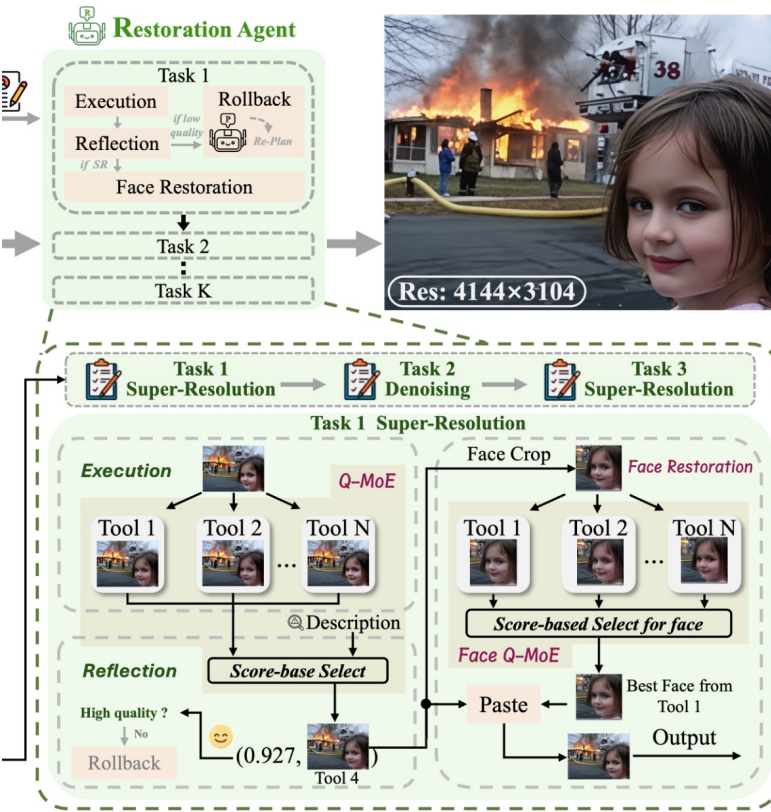


Image after this step:



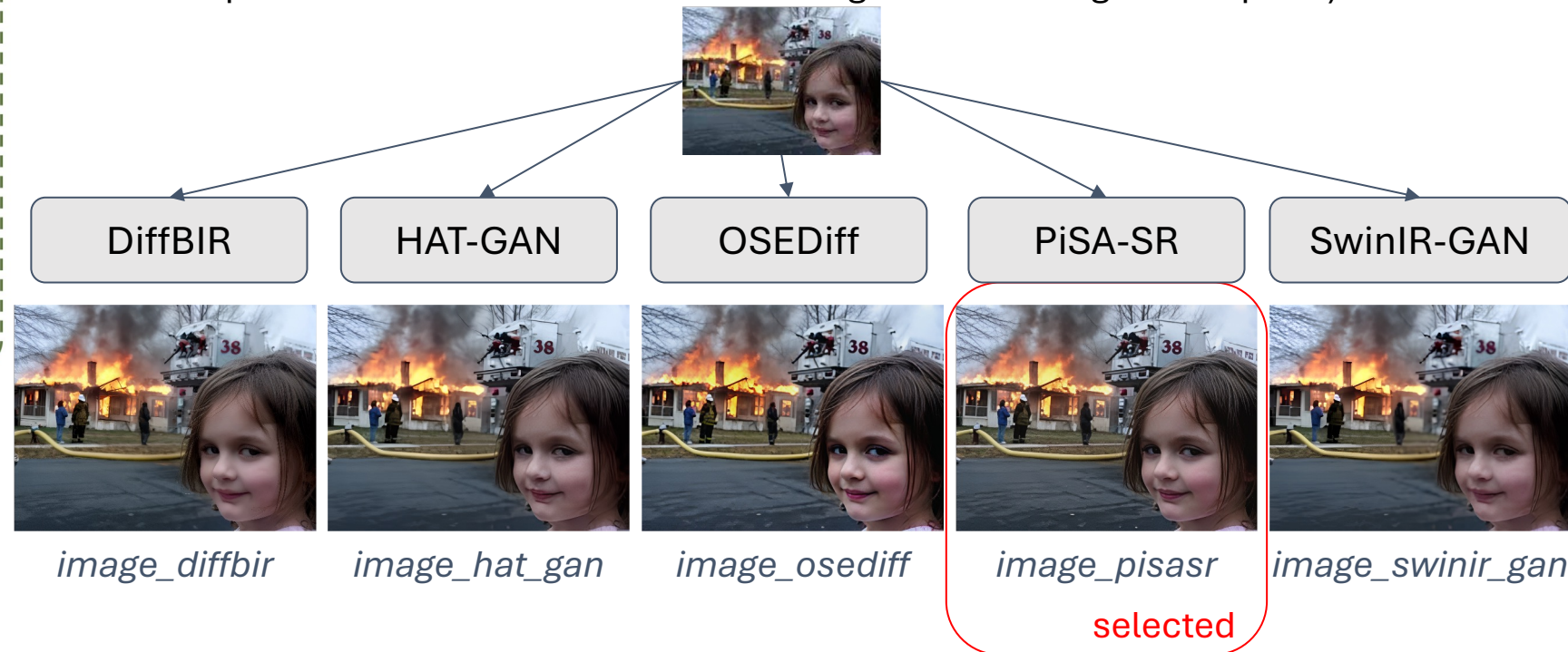
Resolution:
1036 x 776

"Plan": ["super-resolution", "denoising", "super-resolution"]

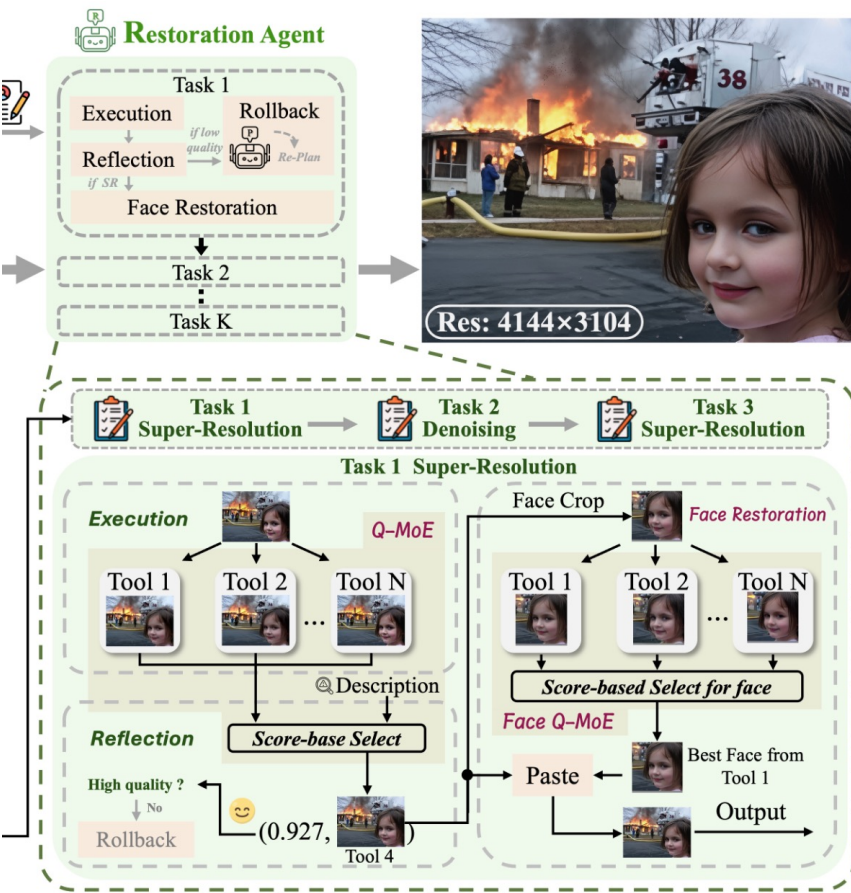
"Go through all tools and select the best."

Execution (based on the plan), Reflection (select the best result): (Q-MoE)

- (1) Execute every method in the corresponding **toolbox**. (super-resolution)
- (2) Select the best image based on a **quality score** (a combined score from no-reference image IQA metrics (NIQE, MUSIQ, MANIQA, CLIPQA), and HPSv2 score for human preference based on the restored image and the image description)



4KAgent: Restoration Agent - Task 1



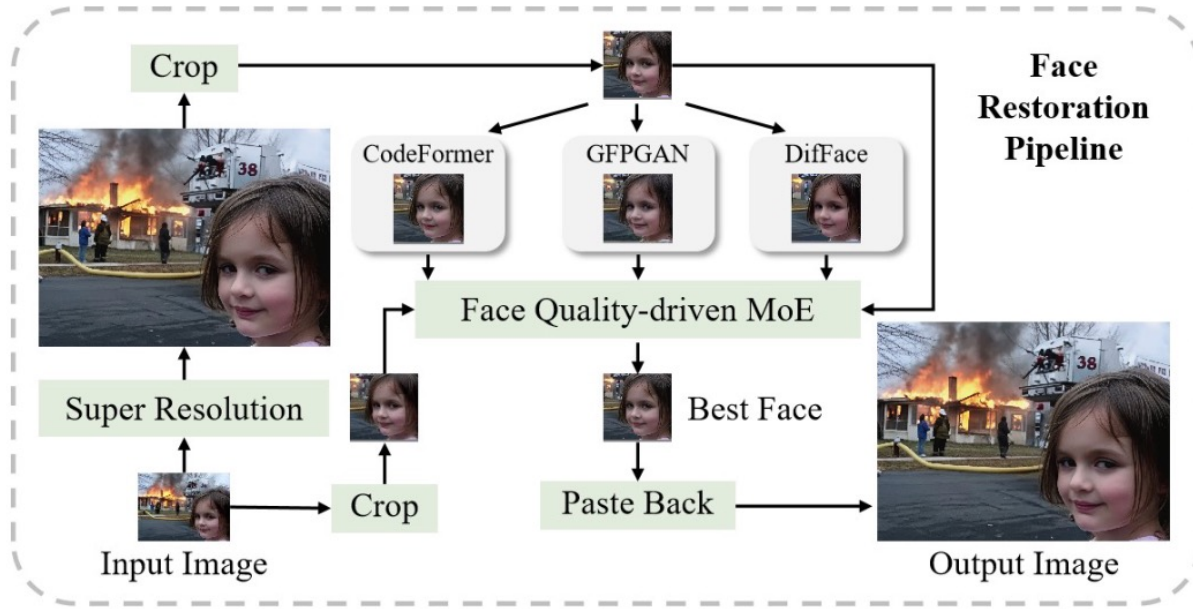
“Is the result of the ‘best’ image good enough?”

Rollback (adjust the plan)

- (1) If the score of the best image is lower than a **threshold**, then this step is a failure step, 4KAgent will generate a failure message.
- (2) Employ Planning Expert in Perception Agent to adjust the plan based on the failure message, image description, distortions in the image, restoration agenda, restoration experience, and (optional) image itself.

This image does not trigger rollback.

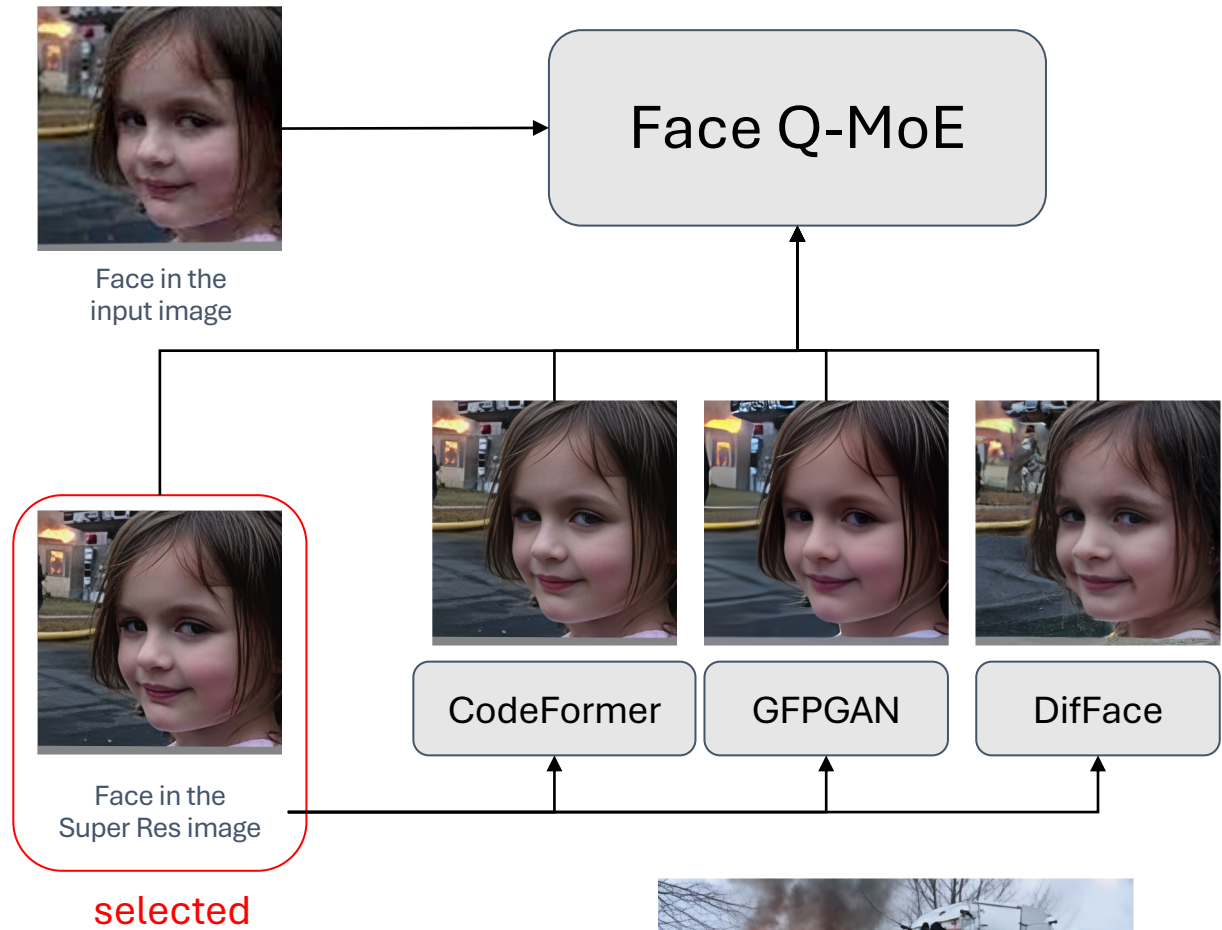
4KAgent: Restoration Agent - Task 1



“Crop face, Go through all tools and select the best.”

Face Restoration Pipeline

- (1) Enable after super-resolution.
- (2) Crop face(s) in the input image (Face set A) and the upscaled image (Face set B).
- (3) Restore / Enhance face in face set B using diverse face restoration method.
- (4) Q-MoE policy: select the best face and paste back.



4KAgent: Restoration Agent - Task 2

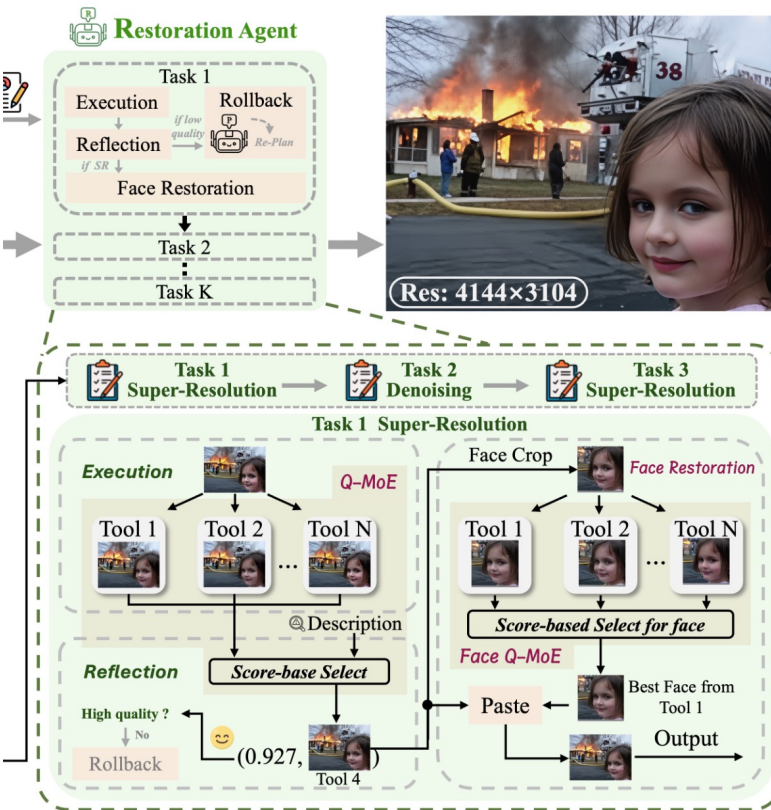


Image after this step:



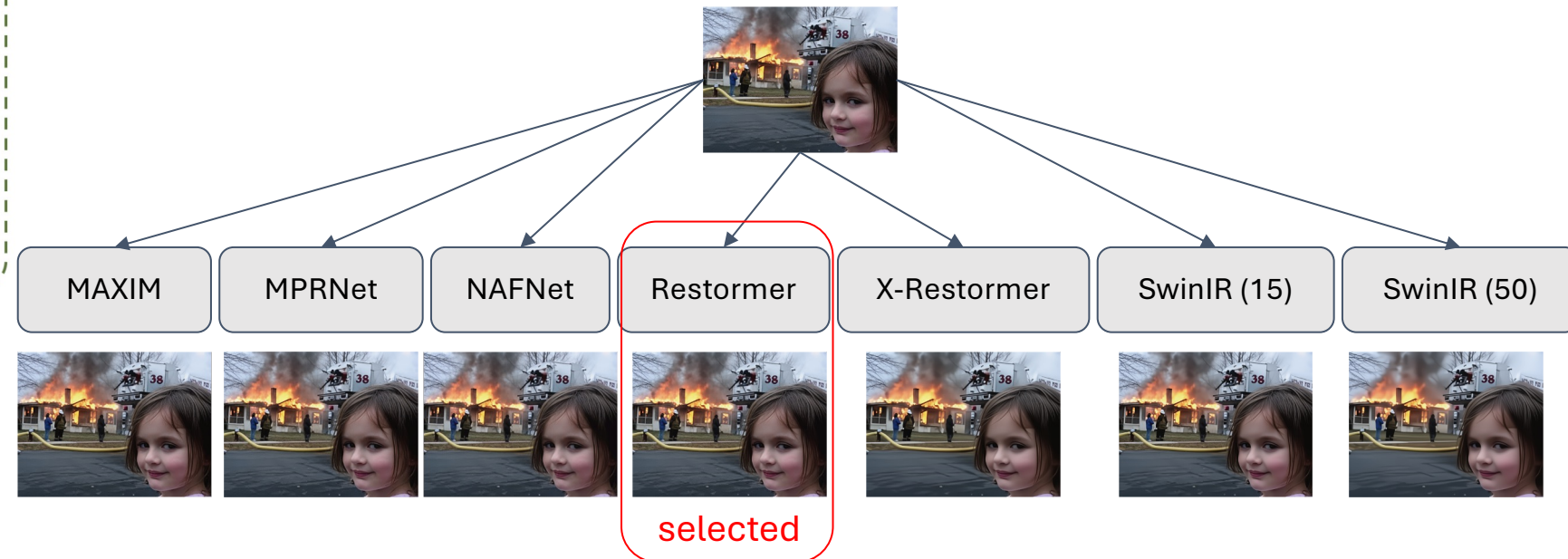
Resolution:
1036 x 776

"Plan": ["super-resolution", "**denoising**", "super-resolution"]

"Go through all tools and select the best."

Execution (based on the plan), Reflection (select the best result): (Q-MoE)

- (1) Execute every method in the corresponding toolbox.
- (2) Select the best image based on a quality score (a combined score from no-reference image IQA metrics (NIQE, MUSIQ, MANIQA, CLIPQA), and HPSv2 score for human preference of restored image and the image description)



4KAgent: Restoration Agent - Task 3

"Plan": ["super-resolution", "denoising", "super-resolution"]

“Go through all tools and select the best.”

Input

Output



Resolution: 259 x 194



Resolution: 4144 x 3104

4KAgent: Profile Module

“Customize 4KAgent for different restoration / enhancement requirement.”

1. **Perception Agent:** Specifies the choice of LLM / VLM employed by the Perception Agent. [Default: Llama-vision]
2. **Upscale to 4K:** Determines whether to upscale to 4K resolution. [Default: True]
3. **Scale Factor:** Explicitly defines the upscale factor for the entire pipeline. (Default: 4, Options: [\[2, 4, 8, 16\]](#)). This parameter overrides “Upscale to 4K” when specified.
4. **Restore Option:** Explicitly sets the restoration task(s) to be applied. If set to None, restoration task(s) are determined automatically by the Perception Agent. (Default: None)
5. **Face Restore:** Toggles activation of the dedicated face restore pipeline. (Default: True)
6. **Brightening:** Controls the activation of image brightening, which may cause color shifts in restored images. Provided as [Optional] to maintain image color fidelity. (Default: False)
7. **Restore Preference:** Defines whether to prioritize higher perceptual quality or higher fidelity in image restoration. (Options: [Perception, Fidelity], Default: Perception). Here we respect the perception-distortion tradeoff [\[4, 142\]](#), deeming models that optimize for distortion metrics (*e.g.*, PSNR, SSIM [\[107\]](#)) as Fidelity models while methods trained for perceptual quality (*e.g.*, NIQE [\[74\]](#), MUSIQ [\[45\]](#)) as Perception models.

4KAgent: DIV4K-50



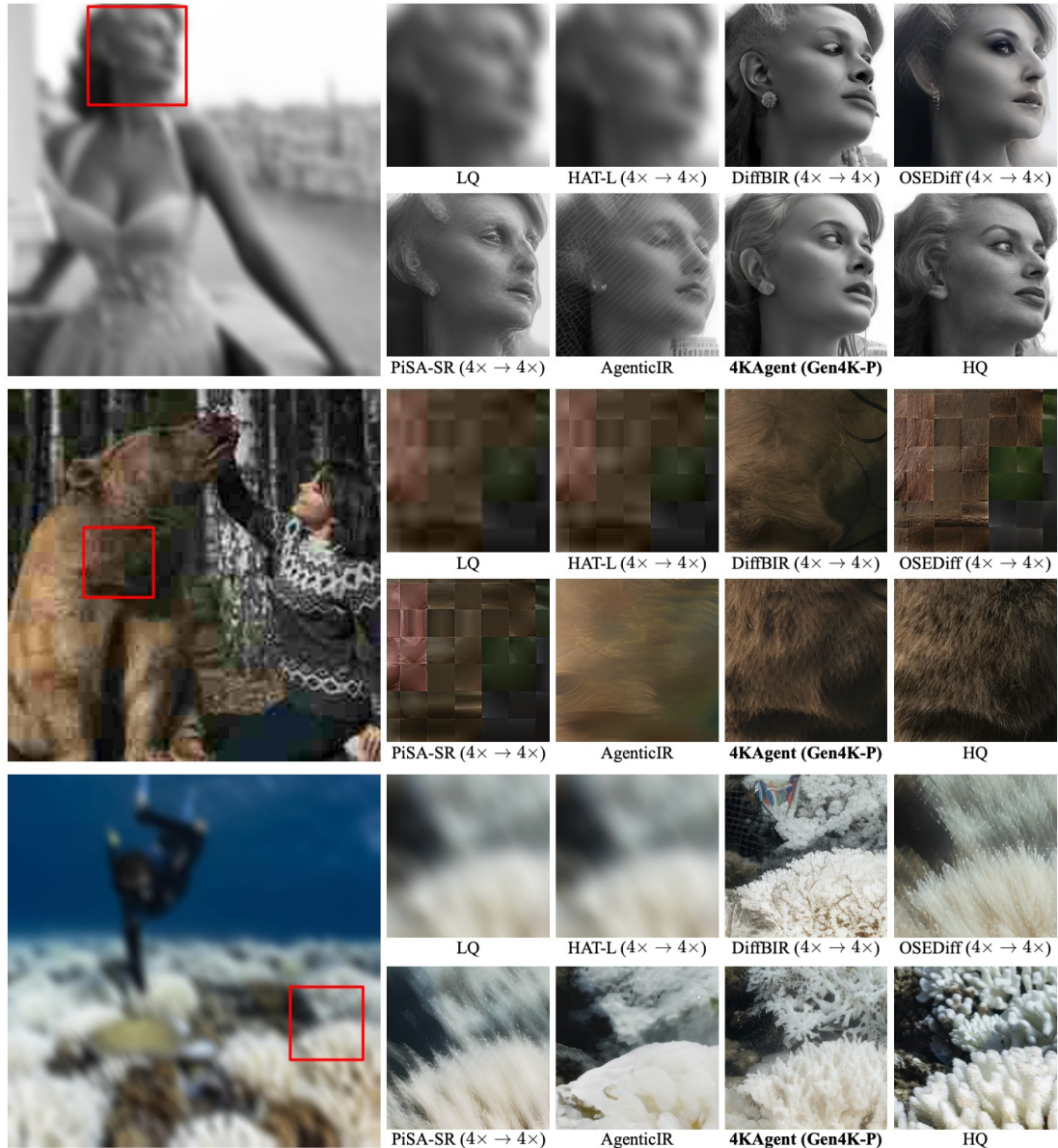
LQ image



HQ image

1. Constructed from the Aesthetic-4K dataset to test end-to-end restoration and ultra-high-scale SR.
2. Specifically, we select 50 images from the Aesthetic-4K dataset, then center-crop each image to 4096×4096 as the high-quality (HQ) ground truth image.
3. Then we downsample HQ images to 256×256 and randomly apply combinations of defocus blur, motion blur, additive Gaussian noise, and JPEG compression to generate the corresponding low-quality (LQ) image.

4KAgent: Experiment Results on DIV4K-50



4KAgent consistently reconstructs finer, sharper and more natural details. Including the facial features in the top row, the bear's fur in the middle row, and the intricate coral textures in the bottom row, highlighting the superiority of 4KAgent.

For more experimental results, please refer to the paper.

Thanks