Guide: Image AI — generate and edit images with AI

DjMirror Pulse includes an **Image AI** panel (Enterprise edition) that improves logos and creates new images using either a **local diffusion engine** that runs on your graphics card, or the **Google Gemini API** in the cloud. It can enhance an existing logo in ten artistic styles, generate a brand-new image from a text prompt, and cut out the background so the result stays transparent over your visuals.

It is a full-screen panel like Global FX or the Layer Generators: open it from **Settings** → **Image AI**. The panel remembers your mode, style, prompt, model and last logo between sessions.

Related guides: stock search and the artist-dialog AI logo presets are in Media Search & AI; applying logos to artists is in Content Management; hardware sizing is in System Requirements.

---

1. At a glance

I want to...	Do this	Needs
Polish or restyle a logo instantly	**Image AI** → **Classic (instant)** → pick a style → **Enhance**	Nothing (works offline)
Restyle a logo with AI	**Image AI** → **Local AI** or **Gemini API** → pick a style → **Enhance**	Local engine + model, or a Gemini key
Make a brand-new image from text	**Image AI** → **Generate** → type a prompt → **Generate**	Local engine + model, or a Gemini key
Keep the result transparent	Turn on **Transparent background** (on by default)	—
Save the result	**Save** → choose a `.png` path	Write permission
Reuse a previous look	Re-enter the same **seed** shown as "Last seed used"	The same model

---

2. Edition and where to open it

Image AI is an **Enterprise** feature (`LicenseEdition::Feature::ImageAi`). Open it from the settings menu: **Settings** → **Image AI** (magenta entry, icon `Layer-Image.png`). The panel fills the space it is given and can be resized like the other controller windows.

What you need:

- **Enterprise license** (or trial) for the panel itself.

- **For the local engine:** a graphics card with enough VRAM for the chosen model, and an AVX2-capable CPU. The local engine is **not** included in the Mac App Store build; on macOS use the Gemini API instead.
- **For the Gemini API:** a Google Gemini API key (free or paid tier).

ⓘNote

The local engine (`sd-cli.exe`, stable-diffusion.cpp) is bundled next to the application on Windows when the build includes it (`engine/sd-cli.exe`). If it is missing, point **Engine (sd-cli.exe)** at a copy you downloaded.

---

3. The three engine modes

The top of the panel offers three mutually exclusive modes. Only one is active at a time.

Mode	What it does	Speed	Needs network?
Classic (instant)	Procedural enhancement of the loaded logo (sharpen, neon, chrome, glitch, …)	Instant	No
Local AI	Runs a diffusion model on your GPU to restyle the logo following one of the ten presets	Seconds to minutes	Only to download models
Generate	Creates a brand-new image from your prompt (text-to-image)	Seconds to minutes	Only to download models

Switching modes keeps your logo loaded; **Generate** hides the logo picker because it does not need a source image. **Classic** and **Local AI** both work on the logo you opened.

The **Provider** row appears only in the AI modes:

- **Local engine** — runs `sd-cli` with a downloaded model pack.
- **Gemini API** — sends the request to Google Gemini. Needs a key.

---

4. Working with a logo (enhance modes)

4.1 Open a logo

1. Make sure you are **not** in Generate mode.
2. Press the folder button (📁) next to **Logo** and choose a file (`*.png *.jpg *.jpeg *.bmp *.webp`).
3. The file name appears in the read-only field; the preview shows the original.

The last opened logo is remembered and reloaded on the next run if the file still exists and is readable. The logo you work on is never overwritten — results are written to the temporary work folder (§8).

4.2 Pick an enhancement style

Ten styles are available, in the same order and with the same names as the artist-dialog AI presets:

#	Style	Look
1	4K Upscale & Polish ✨	Sharper, cleaner version of the same logo
2	Neon Club Glow 🎆	Vibrant purple and cyan neon edges
3	Gold Metallic 🪙	Polished gold with 3D reflections
4	Liquid Chrome 🧪	Molten silver finish
5	Cyber Glitch 👾	RGB split and scanlines
6	Holographic Shimmer 🌈	Iridescent rainbow finish
7	Techno Industrial ⛓️	Brushed steel, gritty edges
8	Retro Outrun 🌇	1980s synthwave gradients
9	Refractive Glass 💎	Thick crystal with light dispersion
10	Inferno Pyro 🔥	Fire and molten lava with glowing-ember edges

Click a style chip to select it. In AI modes the style's prompt is shown in the **Prompt** box and can be edited (§4.3). In Classic mode the style is applied instantly when you press **Enhance**.

4.3 Editing the prompt (AI modes)

The **Prompt** box holds the text sent to the model for the selected style. Two flavours exist:

- **Instruction prompt** — for instruction-following models (FLUX, Qwen). The model follows orders such as "upscale this logo…".
- **Descriptive prompt** — for descriptive models (SD 1.5, SDXL, HiDream). The same style written as a description of the wanted logo.

The flavour is chosen automatically from the selected model pack (`descriptivePrompts` flag) and shown as a hint under the prompt. The label reads *Instruction prompt* or *Descriptive prompt*.

- **Reset** (top right of the Prompt box) restores the style's default prompt.
- Prompts are remembered **per style and per flavour**, so editing one style or one model family does not clobber the others. Only edited prompts are stored.
- **Extra instructions** (the field below) are appended when the result must be transparent (§6). The default is *"Keep the background perfectly flat and uniform, without gradients, shadows or texture."*

Tip

In every prompt, the phrase "white background" is replaced at run time by the flat colour the logo is placed on (the background colour selected in §6). Write "white background" in your prompts and let the app adapt it.

4.4 Run the enhancement

Press **Enhance**. In Classic mode the result appears instantly; in AI modes a progress bar and a status line are shown, and a **Cancel** button replaces **Save** while the engine runs. The log panel (§9) streams the engine output.

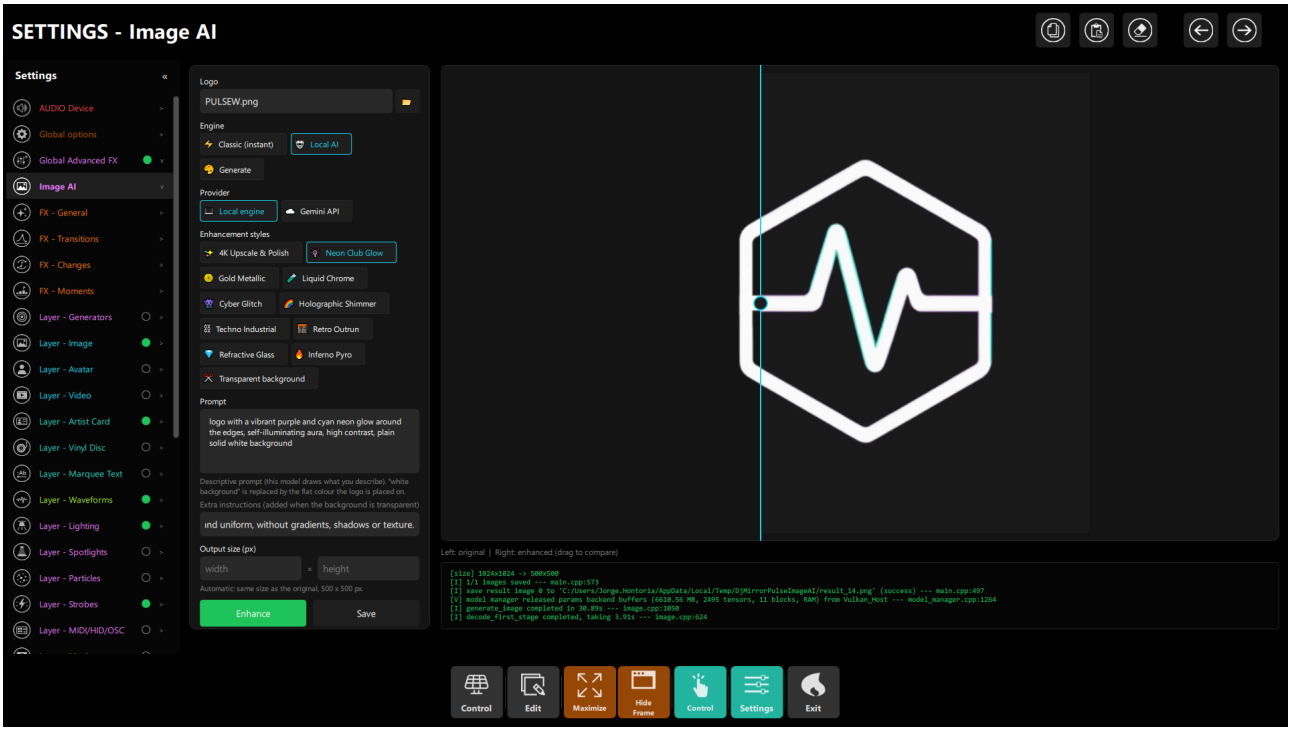

Figure 1: Image AI — Enhance mode with the Neon Club Glow style. The split preview shows original (left) vs enhanced (right).

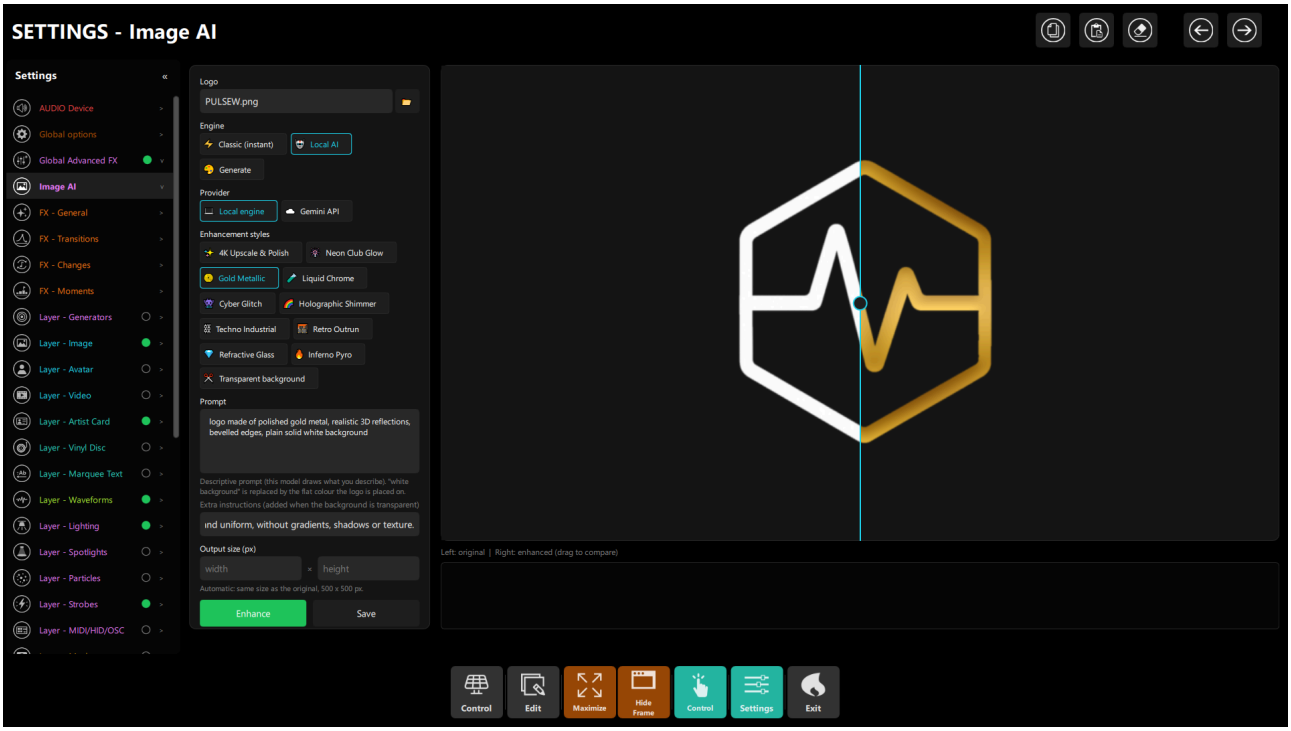

Figure 2: Image AI — Enhance mode with the Gold Metallic style.

---

5. Generating a new image (Generate mode)

Switch to **Generate** (chip). The logo picker disappears and a prompt editor appears.

1. Type a description in **Prompt**, for example "*a neon logo of a skull wearing headphones*".
2. Choose an **Aspect** (1:1, 4:3, 3:4, 16:9, 9:16, 3:2, 2:3).
3. Optionally set a **Seed** (leave empty for a random seed each run). The last seed used is shown so you can type it again to reproduce an image.
4. Optionally set **Steps** (0 = the model's own default, up to 60 in the UI).
5. Optionally turn on **Transparent background** and pick a **Background** colour (§6).
6. Optionally set an **Output size** (§7).
7. Press **Generate**.

The result is shown on the right. Save it with **Save**.

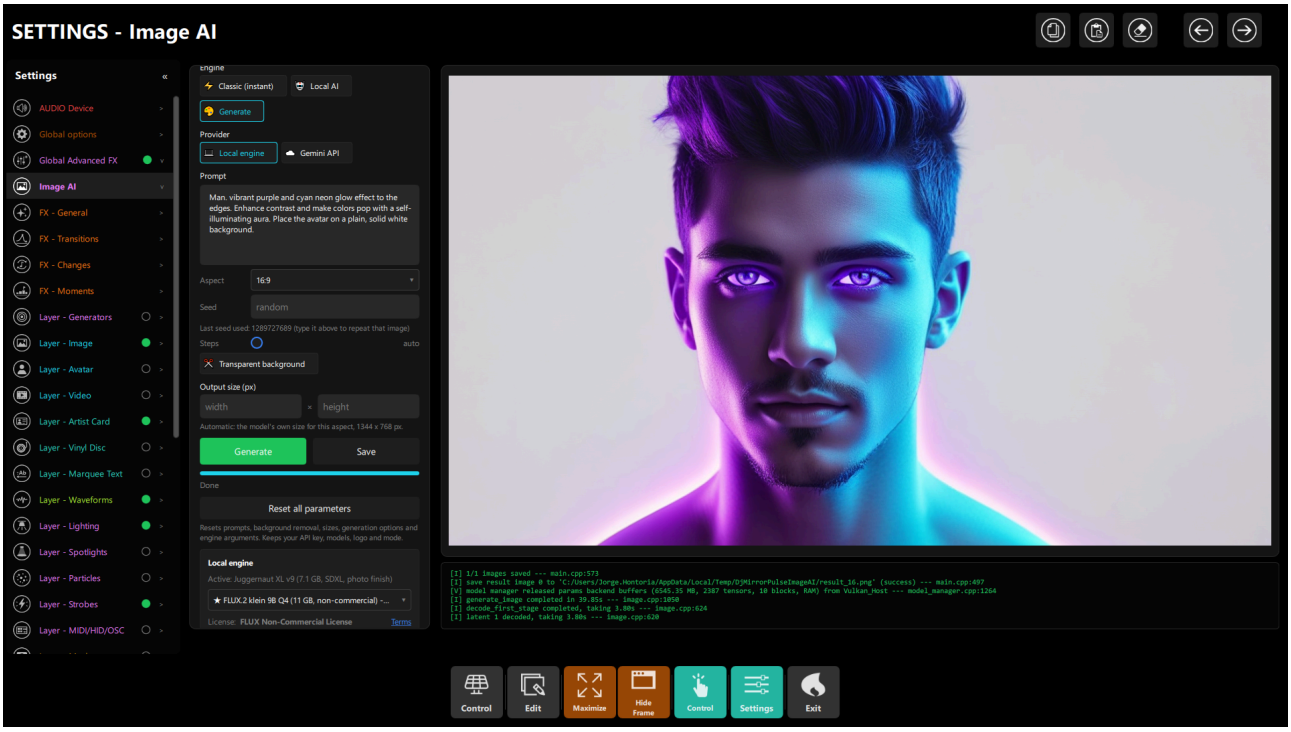

Figure 3: Image AI — Generate mode with a neon portrait created by the local engine (FLUX.2 klein 9B Q4).

Note

With the Gemini API, the seed and steps controls are hidden — Gemini chooses them. The aspect is still sent as part of the prompt.

---

6. Transparent backgrounds and background removal

AI models return opaque pictures. When **Transparent background** is on (default), the panel asks the model for a flat, uniform background colour and then removes it, producing a PNG with an alpha channel.

- **Background** (only when transparent is on): White (safest for logos), Green screen, Black, or Magenta. Use a colour that does not appear in the subject — *a model may tint the subject with a*

green or magenta screen.

- **Background removal** parameters (shown in the grey box when transparent is on):
  - **Tolerance** (0–120): colour distance from the background still treated as background.
  - **Softness** (1–120): width of the soft edge (1 = hard cut).
  - **Margin %** (0–8): how far from the original logo's shape pixels may stay.
  - **Limit to the logo's shape** (enhance modes only): forces everything farther than the margin from the original logo's shape to transparent, so textured or shaded backgrounds the model invented disappear too.
- Changes to these sliders **re-cut the last result instantly** — the AI is not run again. A hint under the sliders reads "*Changes re-cut the last result instantly, the AI is not run again.*" when a raw result exists.
- If the picture already has real transparency (more than a stray pixel), no cut-out is performed.

The resulting PNG works well over dark visuals. If the automatic removal cannot separate the logo (for example a white logo on a white background), try another preset or start from a logo with a plain, high-contrast background.

---

7. Output size

By default the size is automatic:

- **Enhancing a logo:** the original photo's size.
- **Generating:** the model's own size (about 1024 px for Gemini).

Set **Output size (px)** (width × height) to force an exact size. Leave a field empty (or press **Auto**) to go back to automatic. Limits: 1 to 4096 px per side. When only one side is given, the other follows the aspect ratio of the mode (the chosen aspect for generation, the logo's ratio for editing).

---

8. Where files are stored

- **Work files** (source copy, raw AI output, cut-out result) are written to the system temp folder `DjMirrorPulseImageAI/` as `input.png`, `result_N.png`, etc. They are overwritten on the next run.
- **Nothing is saved permanently until you press Save.** Use **Save** to write a `.png` wherever you want.
- **Downloaded models** are stored in the application data folder under `models/` (§10).
- **Settings** (mode, style, prompts, cut-out parameters, sizes, engine paths, Gemini model) are stored in the application's Image AI settings file (`*-ImageAI`), separate from the main DjMirror Pulse settings. The Gemini API key is read from the shared `apiKeys/imageEnhancement/gemini` setting (Settings → API Keys) or from the `GEMINI_API_KEY` / `GOOGLE_API_KEY` environment variable.

⚠Warning

Work files live in the temp folder and can be cleaned by the OS. Always **Save** the results you want to keep.

---

9. The log panel

In AI modes a read-only console at the bottom of the right-hand pane streams:

- The full JSON request/response flow for Gemini (via `ImageEnhancementService`).
- The `sd-cli` command line and its progress for the local engine.

It is green-on-black and scrollable. Use it to diagnose failures (missing model, out of memory, rejected prompt).

---

10. Local engine setup

Open the **Local engine** box (shown when an AI mode is active and **Local engine** is selected).

10.1 Engine executable

- **Engine (sd-cli.exe)**: path to `sd-cli`. The build bundles it as `engine/sd-cli.exe` next to the application on Windows; if it is missing, use the `...` button to point at a copy.
- **Model file**: optional path to a hand-picked model (`.safetensors` / `.gguf`). Leave empty to use a downloaded pack.

10.2 Downloading a model pack

The combo lists the catalog of model packs (FLUX, Qwen, SDXL, Stable Diffusion 1.5, HiDream, ...). For the selected pack the panel shows:

- **License**: name, kind (permissive / restricted / other) and a link to the terms, plus a one-line note.
- **Requirements**: GPU, RAM, disk, Mac (Apple Silicon) and speed, marked *measured* or *estimate*.

Press **Download** to fetch the pack (resumable, SHA-256 verified) into the application data `models/` folder, or **Use this model** if it is already downloaded. A progress bar and status line track the download; **Cancel download** stops it and keeps the partial file for resuming later.

Each pack declares what it is good for (`recommendedFor`: `logo`, `generate`, or both) and whether it uses descriptive or instruction prompts.

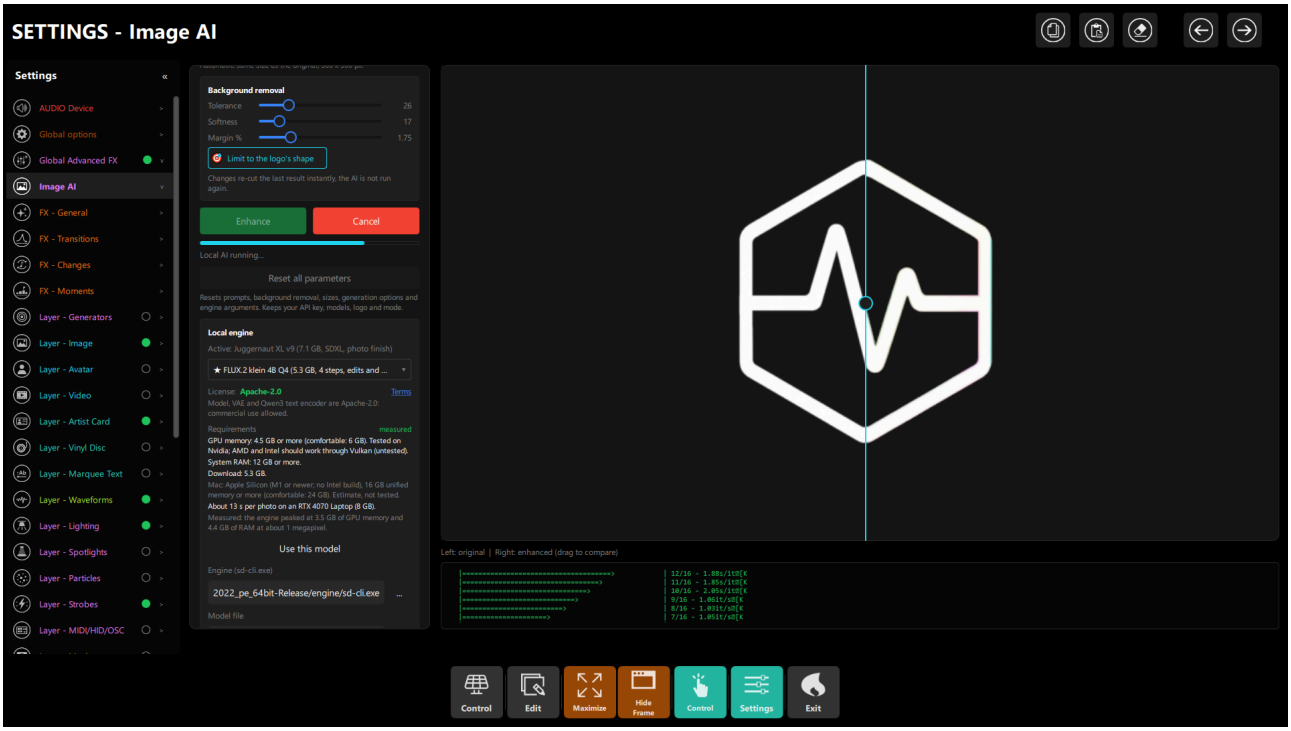

Figure 4: Image AI — the local engine section showing the FLUX.2 klein 4B Q4 model pack with its licence and hardware requirements.

10.3 Engine parameters

- **Strength** (0.10–1.00): how strongly the AI changes the source logo (editing only).
- **Arguments**: the template passed to `sd-cli` (advanced). Defaults come from the selected pack.
- **Output size** and **transparent background** are shared with §6 and §7.

`engineProblem` (shown in red when non-empty) explains why the AI cannot run right now (missing executable, missing model, etc.).

---

11. Gemini API setup

Open the **Gemini API** box (shown when an AI mode is active and **Gemini API** is selected).

- **API key**: same key as Settings → API Keys → Image Enhancement (Gemini). Paste it here or set `GEMINI_API_KEY` / `GOOGLE_API_KEY`. Masked with dots. The key is only sent to Google, in a header.
- **Get a key** opens `https://aistudio.google.com/apikey`.
- **Model**: default `gemini-2.5-flash-image`; alternative `gemini-3-pro-image-preview`. Type a name or pick a chip.
- **Test key and list image models**: checks the key and lists the image models it can use. Status is shown below (green when "accepted").
- `aiProblem` (red) explains any blocker.

ⓘNote

Gemini quotas: free tier is 15 requests/minute and 1 million tokens/month. Implement backoff on HTTP 429.

---

12. Resetting parameters

Reset all parameters is a two-click button (first click arms it, second click confirms; it disarms after 4 s). It resets prompts, background removal, sizes, generation options (aspect, seed, steps, background colour), the Gemini model name, and the engine arguments and strength of the active model.

It **keeps**: the API key, downloaded models, the engine path, the loaded logo, the text of the generation prompt, the mode, the provider and the selected style.

---

13. Pre-show checklist

1. Enterprise license is active on the show machine.
2. If using the local engine: model pack is downloaded and `engineProblem` is empty; run one test enhancement.
3. If using Gemini: key tested ("accepted") and quota available.
4. AI results are **Saved** to a known folder, not left in the temp work folder.
5. Transparent backgrounds look right over your dark visuals (check edges).
6. Heavy generation was done in advance — not during doors.

---

14. Troubleshooting

Symptom	Likely cause	Fix
Panel is missing or locked	Not Enterprise edition	Activate an Enterprise license/trial
`engineProblem` non-empty	Missing `sd-cli` or model	Install/bundle the engine, download a pack, or set paths
"No candidates in response" (Gemini)	Prompt refused or quota	Adjust the prompt, avoid restricted terms, check quota
Result still has a background	Cut-out could not separate the subject	Try another preset/background colour, or start from a high-contrast logo
Subject tinted green/magenta	Background colour leaked into the subject	Pick a different **Background** colour
Edges too hard / too soft	Tolerance or softness off	Adjust **Tolerance** / **Softness**, then it re-cuts instantly
Generation too slow	Heavy model or GPU congestion	Use a lighter pack, reduce steps, simplify the prompt
Saved file missing	Left in temp folder	Use **Save** to a permanent path
Out of memory	Model too large for VRAM	Use a smaller pack or enable CPU offload via args

Symptom	Likely cause	Fix
Settings lost on a second laptop	Settings are per machine	Repeat setup; keys are per computer

More: Troubleshooting.

---

15. Quick recipe card

- **Open:** Settings → Image AI (Enterprise).
- **Modes:** Classic (instant) · Local AI · Generate. **Providers:** Local engine · Gemini API.
- **Enhance a logo:** open logo → pick one of ten styles → (AI) edit the prompt → **Enhance** → adjust background removal → **Save**.
- **Generate: Generate** mode → prompt + aspect + seed/steps + transparent background → **Generate** → **Save**.
- **Transparent:** ask for a flat background colour, then it is keyed out; sliders re-cut without re-running the AI.
- **Local engine:** bundle or point at `sd-cli.exe`, download a pack (license + requirements shown).
- **Gemini:** paste/test a key, pick `gemini-2.5-flash-image` or `gemini-3-pro-image-preview`.
- **Save everything you keep** — work files live in the temp folder.