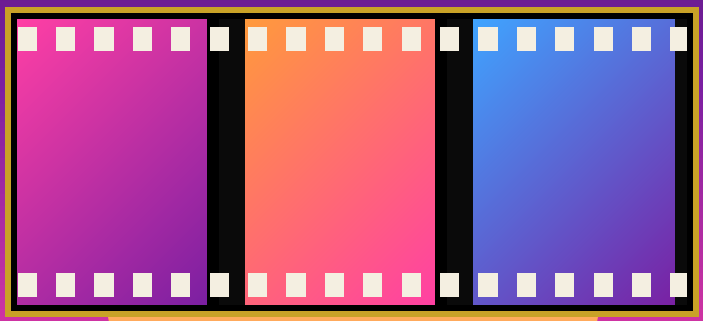
NODE PAK
INSTRUCTION
BOOKLET

IRONLEO78 MTV

MUSIC · TEXT · VIDEO



IRON
LEO
78

TESTED
BY 1,000+
CHECKS
(ALL GREEN)

YOUR SONG. YOUR CLIPS. ONE NODE.

For ComfyUI.
Read this booklet
before you press
Run.

Draft edition

CONTENTS

Get the models first	3
What it does	4
How many nodes?	5
Quick start	6
The MTV node	7
Casting and Direction	8
The Cutting Room	9
Post-Production	10
Undercover	11
Under the hood	12 to 13
Tested, and made with Claude	14
The man behind it	15
Troubleshooting	16
Dependencies	17

HOW TO READ Every node also works alone. Pages 7 to 11 show each node on its own and together with the others.

Turn your song into a music video, one clip at a time, without building a 50-node workflow.

GET THE MODELS FIRST

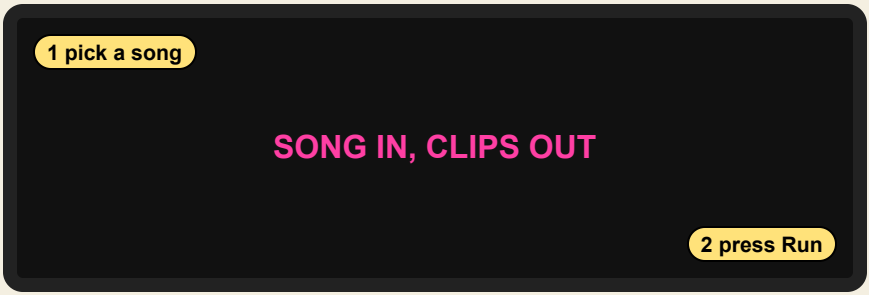
The main diffusion model is unique to this node pak. Easiest start: open ComfyUI's own **MiniMax H3 Ref2V** template and the **LTX 2.5** template once; they download the official files for you. The author's repo only holds the custom files that are not in those templates.

File	Folder	Where
IronLeo78_R2V4Vendetta_ (main model)	diffusion_models	Ironleo78/IronLeo78-MTV
h3_upscaler_lms_v0.1	latent_upscale_models	IronLeo78-MTV / Latent_Upscaler
Huihui-Qwen3-VL-4B (Casting LLM)	text_encoders	IronLeo78-MTV or ahmed22xa
minimax_h3_lynnreal_light_vae	vae	Kijai / IronLeo78-MTV (backup)
qwen3vl_32b_minimax_h3 (text encoder)	text_encoders	Comfy-Org/MiniMax-H3
minimax_h3_audio_vae_fp32	vae	Comfy-Org/MiniMax-H3
Turbo 4-step LoRA (ref2v)	loras	Comfy-Org/MiniMax-H3
yue2_3b_int8_convrot, sheetsage2_bf16	checkpoints, audio_encoders	Comfy-Org/YuE2 (Undercover)

huggingface.co/Ironleo78/IronLeo78-MTV. LTX 2.5 files: huggingface.co/Lightricks/LTX-2.5 (gated: request access).

TIP If a model is updated upstream, rename the new file to the exact name the node expects. A backup of the Lynn VAE is kept in the author's repo.

WHAT IT DOES



Pick a song, press Run. The node cuts the matching slice of your audio and generates the picture to follow the music. What you hear is your recording.

One clip per Run. A segment panel keeps your place; every song has its own counter. Go to any segment or time, redo a bad clip.

Continuity. Keep Rolling starts each clip from the last frame of the one before.

Second Unit Director extends the clip about 20 seconds with LTX 2.5.

Sliding modules. Casting and Direction on the left, the Cutting Room on the right, Post-Production docked beside it.

Four themes, saved prompts, 73 cover styles, 1,270+ checks.

HOW MANY NODES?

You see **3** nodes and work with **1**.

Node	What it is
IronLeo78 MTV	The one you work with: song, quality, buttons, segment panel.
IronLeo78-Post-Production	Refines and upscales a clip. Spawned beside the MTV.
IronLeo78 Undercover	Stand-alone song-cover maker.

Six node classes are registered. Casting and Direction and the Cutting Room are real nodes too, kept hidden and wired by the MTV, and shown as slide-out panels.



The MTV alone stands in for **30+** stock nodes. Post-Production replaces **20+**, Undercover **10+**. About 60 nodes and all their wiring become the ones you touch.

QUICK START

1. Put the folder **ironleo78_mtv** in **ComfyUI/custom_nodes/**.

2. Get the models (page 3). Restart ComfyUI, press Ctrl+F5. Use the Nodes 2.0 front end.

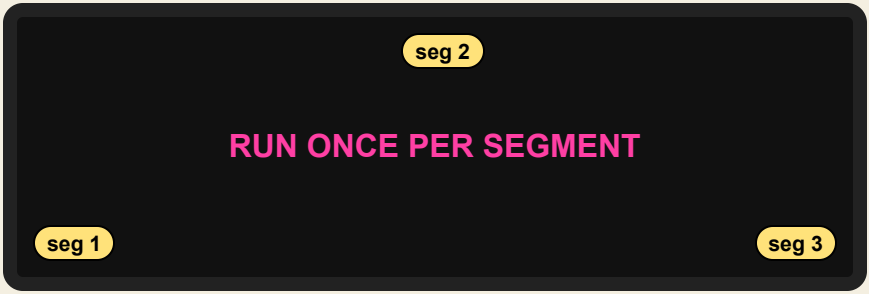
3. Add **IronLeo78 MTV** (IronLeo78 > MTV).

4. Custom Audio File: pick your song.

5. Press **Casting and Direction**: load your singer in slot 1, write a short prompt.

6. Pick **Test** (fast) or **Prod**, keep **Rolling!**, press **Run**.

7. The clip plays in the **Cutting Room**. Like it? Switch **Remastering** on and Run Post-Production.



NOTE After the last segment the node stops quietly. Use Go, Latest or Reset in the panel.

Control	What it does
Test / Prod	0.2 MP, 4 steps / 0.4 MP, 6 steps.
Custom Audio File	Your song. Its name names the output folder.
Choose Audio latent	Custom Audio, MiniMax Audio, or Audio From Video.
duration	Seconds per clip (default 5).
Keep Rolling	Rolling! next segment every Run. Cut! same segment again.
Second Unit Director	LTX 2.5 extends the clip about 20 s.
Remastering / Rough Cut	Mirror of the Post-Production switch.
Seed Random / Fixed	New seed per clip, or the same one.
Segment panel	Go, Go time (1:23), Latest, Reset.

Outputs: video (to Save Video), filename, status.

Alone: a song and a picture are enough. **Together:** the hidden modules are already wired to it.

CASTING AND DIRECTION

Slides out of the **left** side of the MTV.

Pictures: a 2 x 2 grid. Click a cell to pick or upload, drop files, red x to remove.

Mode: v2v swaps things in a video you load; i2v starts from pictures.

Preset: plain .md files in **presets/**. Edit them or add your own.

Prompt enhancer: Qwentin Tarantino's Script = the LLM writes the prompt. Your Script = your text goes through as typed.

Saved prompts: a list above the prompt box. Save prompt as... keeps one as a .txt in **prompt_presets/**.

Themes: four big swatches at the bottom: Rose Delight, 80s, Corporate, Pioneer.

ALONE It is an ordinary node too: wire its Pictures and MTV_Prompt outputs to the MTV yourself.

THE CUTTING ROOM

Slides out of the **right** side. Nothing to wire.

Video Preview: plays every clip with its sound. Runs list (newest first), frame stepping with , and . keys, a Seam check (2 seconds each side of a join), Save, Delete, rating 1 to 10.

Previews: the live sampler picture with step, speed, time left and a replayable filmstrip.



TIP Set ComfyUI's Preview Method to taesd or latent2rgb to see the live picture. Clips live in output/IronLeo78-Video-Preview.

POST-PRODUCTION

Refines and upscales one saved clip, x2.00 on each side. Docks to the **right** of the Cutting Room and follows the MTV.

1. Make clips with the MTV (latents are always saved).

2. Pick the **latent** (newest first; press R). The conditioning is found by itself. Upload brings clips from another machine.

3. Switch **Remastering** on and Run. Link video to Save Video. It then flips back to **Rough Cut**, so a new scene never starts a refine.

4. Delete all latents clears every clip file. Press twice. Videos stay.

It refines with the clip's own seed, with no LoRa, at euler and 3 steps.

ALONE Works from any saved clip files, no MTV needed.

UNDERCOVER

Always stand-alone. Load a song, write your lyrics, pick a style: you get a **YuE2** cover saved as FLAC or MP3 320k in ComfyUI's input folder.

SheetSage2 reads the melody. Key and BPM can change: the tune is transposed, not replaced.

73 style presets (ported from SongScribe by TheLocalLab, MIT) with Vocals, Era, Texture, Mood Shift, Language and Blend.

Analyse only reads key, BPM and meter without making a cover.

TOGETHER Wire its YuE2 output to the MTV's song input to cover and film in one queue.

UNDER THE HOOD (1)

One engine at a time. The MTV and Post-Production share one cached H3 base model. The Casting LLM, LTX and YuE2 give memory back before H3 loads.

The MTV Run:

1 wires resolved, 2 saved seed applied, 3 models load (cached), 4 conditioning made with ComfyUI's own MiniMax node and cached, 5 sample (BasicGuider, res_multistep), 6 decode and create the video (24 fps, h264), 7 latents, sidecar and manifest saved, 8 Seed Random picks the next seed, the counter moves on.

Segments: segment 1 starts at second 0; each next one a real clip length later (frames / 24). A clip is frames % 17 == 5 long: 5 s = 124 frames.

Why audio is frozen: your song's latent gets a zero noise mask, so sampling never touches it. The picture follows your recording.

UNDER THE HOOD (2)

Files a clip leaves: .mp4, _video.latent, _audio.latent, .conditioning, .json sidecar, plus the song's manifest CSV. The latent header holds the seed.

Post-Production: check, load, upscale x2.00 (trained node if installed, else built-in resize), build model, 3 euler steps (0.9035, 0.6316, 0.3158, 0), decode, create video. Rough Cut returns blockers and does nothing.

The UI: plain ES modules. Slide-out panels are children of the node's own element. Hidden ghost nodes carry the real widgets. Values are restored by name. Colours come from one contrast-tested theme table.

Hardcoded on purpose: diffusion model, VAEs, text encoder, attention, upscaler, refine schedule, Casting LLM, LoRa (off), 24 fps. Fewer knobs, fewer bad clips.

TESTED, AND MADE WITH CLAUDE

27 test files, 1,270+ checks: Python against stand-ins for ComfyUI (no GPU), JavaScript in jsdom, a preset audit, a builder check, ruff, and a syntax check of every web file. Every release must be ALL GREEN before it is committed, archived and installed.

RUN IT `python tools/run_all.py`

A reviewer (Gemini) once called the code "industrial grade solid" and robust. Take that lightly, lol. The tests back it up, and so would Claude.

Built with Claude Code, and written to be changed with it: ARCHITECTURE.md is the map, CHANGELOG.md lists every release, and **every release ships handoff notes** so a fresh session knows where things stand.

Want another button name, theme or preset? Open the folder in Claude Code and ask.

THE MAN BEHIND IT

IRONLEO78

IronLeo78 spent **25 years in metal music**, and has digital video editing experience on top: cutting picture to sound is the whole point of this node.

A deep passion for music, and for directors like **Terry Gilliam** and **Quentin Tarantino** (to name a few), is why the UI talks like a film set: Casting and Direction, the Cutting Room, Post-Production, Keep Rolling, Second Unit Director, Remastering, Rough Cut.

More than 55 releases, about a hundred commits, many iterations and a lot of "no, not like that". The design you see is what survived.

LICENSE MIT. Shared in that spirit.

TROUBLESHOOTING

Problem	What to do
Nodes missing	Folder directly in custom_nodes with __init__.py. Restart, Ctrl+F5, check the console.
Import error (MiniMax H3)	Update ComfyUI.
Model not found	Exact file name in the right models folder, press R.
Panels or buttons look wrong	Use Nodes 2.0, Ctrl+F5.
Segments: unknown	Pick a Custom Audio File or Video File.
Node stops and does nothing	Last segment passed: Go, Latest or Reset.
Out of memory	Close GPU programs, use Test, retry.
Post-Production list empty	Make a clip, press R, or Upload.
"audio latent missing"	The three clip files must share a name and folder.
Post-Production does nothing	Remastering is OFF, or it already ran once.
Phone video on its side	Turned upright by itself; mirrored video is not un-mirrored.

DEPENDENCIES

ComfyUI recent enough to include the MiniMax H3 nodes, LTXVSeparateAVLatent / LTXVConcatAVLatent, CreateVideo, SaveVideo, ModelAttentionBackend (optional) and the core samplers and loaders.

Python packages: none to install. torch, torchaudio, numpy, Pillow, PyAV, safetensors and aiohttp ship with ComfyUI. No ffmpeg.

Optional: the Comfyui_Minimax_h3_latent_Upscaler node (MIT, LBH-123-AI) for the trained upscaler:

```
git clone https://github.com/LBH-123-AI/Comfyui_Minimax_h3_latent_Upscaler.git
```

Hardware: an Nvidia GPU (the upscaler runs on CUDA). H3 and LTX are never in RAM together: about 79 GB peak with Second Unit Director.

Gated: LTX-2.5 needs access on Hugging Face. Qwen3-VL 4B / 8B encoders do not work as the H3 text encoder.

NODE PAK

THE END

THANKS FOR READING



IRON
LEO
78

GITHUB.COM/IRONLEO78-
SYS/IRONLEO78_MTV

MIT licensed
by IronLeo78