MAKE THINGS

Image Generation, in Practice

Reference images, inpainting, character LoRAs, upscaling, and the licence nobody read.

First edition, September 2026

Addaly

Series editor: Dr Mustafa Mun

Draft, open for academic review

MAKE THINGS

Image Generation, in Practice

Reference images, inpainting, character LoRAs, upscaling, and the licence nobody read.

Series editor: Dr Mustafa Mun

Addaly

First edition, September 2026 · Draft, open for academic review

Image Generation, in Practice

First edition, September 2026, build 62a26075.

Written by AI models to a written standard, with Dr Mustafa Mun as series editor. Exam keys checked blind by a second model: 3,665 of 3,666 agreed. Academic review invited.

The manuscript was drafted with the assistance of a large language model and was edited and published under his name. That is stated plainly here rather than left to be discovered, because a reader is entitled to know how a book they are trusting was made.

How to cite: *Mun, M. (2026). Image Generation, in Practice. Addaly. addaly.com/learn/ai-images.*

This book is the printed form of the course of the same name at addaly.com/learn/ai-images. The website is the living version and is corrected first. Where the two differ, the website is right.

Free to read, free to download, free to print, and free to teach from. There is no paid edition and there never will be.

Every diagram in it is drawn from data by code, not generated as a picture, so the numbers on the page are the numbers in the argument. The answers to every check, test and examination question are at the back.

7 modules · 69 lessons · 27 figures · 7 case studies · 55,057 words · about 9 hours

Brief contents

1

MODULE 1

The kit, and what each piece of it costs you

Assemble a working image-generation setup from the hardware you actually own — hosted service, local install or a free cloud GPU — and state for your chosen route the VRAM it needs, the seconds and rupees per image it costs, the licence it carries, and the exact files on disk that make up the model.

Lesson 1 · 9 min

The generator you pick is a licensing decision

THE ONE THING TO KEEP

The licence on a model's weights governs whether you may use it commercially, and it is a completely separate question from who owns the picture it made.

Everyone compares the pictures. Almost nobody compares the terms

Pick any two image generators and the sample galleries will look roughly as good as each other. That is not where they differ any more. They differ on three things that decide whether you can actually finish a job: what one image costs you, how much control you get when the first result is wrong, and what you are permitted to do with the file afterwards.

The third one is the one that ends careers-in-miniature — a designer who delivers 200 assets built on weights they were not licensed to use commercially.

If you have not read how these models actually work, read that first. This course assumes it and never repeats it.

The hosted tools, and what each is actually for

Midjourney still produces the most immediately attractive frame with the least effort. The cost of that is an opinion baked into every image — a house look you have to fight. It is subscription-only, with no free tier, and on the entry plan your generations are visible in the public gallery; private generation sits on the higher tiers. If you work under NDA, check that before your first render, not after.

Adobe Firefly is the conservative choice. It is trained on Adobe Stock and licensed or public-domain material, and Adobe offers indemnity against training-data claims on its business plans. It is less exciting and more

defensible. It has a free monthly credit allowance, and Generative Fill inside Photoshop is the same engine.

Google's Gemini and Imagen line is strong at following a long instruction and at editing an image you supply. Google stamps its generated images with SynthID, an invisible watermark. There is meaningful free use through the Gemini app and AI Studio.

Ideogram exists because text inside images was broken for years. It renders legible words far better than the general-purpose models. It is still not typesetting — more on that in the finishing lesson.

Recraft aims at design output rather than pictures: it will give you actual vector SVG, which matters if the thing you are making has to scale to a billboard. See vector and print for why that distinction is not cosmetic.

Higgsfield is a different species — a recipe-driven platform rather than a raw model. The next lesson is about that whole category.

Running the weights yourself

The open-weights line runs Stable Diffusion → SDXL → SD 3.x, alongside FLUX from Black Forest Labs and Qwen-Image from Alibaba. Running them yourself costs nothing per image, gives you every control that exists, and works offline.

The requirement is a GPU. SDXL is comfortable on 8 GB of VRAM. FLUX's larger variants want much more, though quantised versions run on 8-12 GB slowly. If you do not own a card: Google Colab's free tier, Kaggle's free weekly GPU hours, and Hugging Face Spaces all run these models at no cost. Running models yourself and Hugging Face cover the mechanics.

Free front ends, all genuinely free: **ComfyUI** (a node graph, ugly, total control), **Fooocus** (Midjourney-like defaults, almost no settings), **InvokeAI**, and **Krita with the AI Diffusion plugin**, which is the best free inpainting experience there is.

The licence trap, stated plainly

A model's weights carry a licence, and it is a separate thing from what you may do with the pictures.

- **FLUX.1 [schnell]** is Apache 2.0. Do what you like.
- **FLUX.1 [dev]** is a non-commercial licence. Thousands of people have downloaded it, run it locally, and billed clients for the output. That is a breach of the model licence, regardless of any question about who owns the image.
- **Qwen-Image** was released Apache 2.0.
- **Stability's recent releases** use a community licence with a revenue threshold: free below it, paid above.
- **Hosted tools** replace all this with their terms of service, and free tiers frequently restrict commercial use even when paid tiers do not.

Those were the positions in 2025 and they move. Check the model card on the day you start the job. Lesson ten deals with output rights, which is a different question again.

Choosing, in about thirty seconds

- A single striking image, no budget for fiddling — Midjourney or the Gemini app.
- Corporate work where somebody will ask about training data — Firefly, or Getty's and Shutterstock's own generators.
- Words inside the image — Ideogram, then set real type over it anyway.
- Hundreds of near-identical variants — a preset platform.
- Precise control, a repeating character, or no money at all — open weights, locally or on free GPU hours.

Today: open the model card or terms page of whatever you are currently using and find the sentence about commercial use. Most people have never read it.

BEFORE YOU MOVE ON

A freelancer downloads FLUX.1 [dev], runs it on their own machine with no internet connection, and invoices a client for the resulting images. What is the actual problem?

A. Nothing is wrong. A licence attached to model weights controls redistributing the file, not what you make with it once it is on your disk.

B. The images are unlikely to attract copyright, so there was nothing there to sell in the first place.

C. The [dev] weights are released under a non-commercial licence, so billing a client breaches the model's terms whatever the copyright position on the images turns out to be.

D. Generating locally strips the Content Credentials the client needed for disclosure, which is what makes the delivery non-compliant.

The answer and why: p. 381

Lesson 2 · 8 min

A preset is somebody else's workflow, frozen

THE ONE THING TO KEEP

Recipe-driven tools trade control for speed, and the bill arrives at the first revision, when the thing the client wants changed is the thing the preset already decided.

Why the packaged tools feel like magic

Open a raw model and you face a blank prompt box, a sampler dropdown, a step count, a guidance scale, an aspect ratio and a seed. Most of those decisions have one sensible answer for the thing you are making, and you do not know what it is yet.

Open a platform like **Higgsfield** and you face a wall of named recipes instead — a product-shot look, a marketing card layout, a cinematic portrait, a set of camera moves borrowed from its video side. You click one, you drop in a photo or a line of text, and something usable comes out in under a minute.

That is not a cheap trick. A preset is a real workflow that somebody built and tested: a chosen model, a prompt scaffold you never see, probably a style adapter or a LoRA, a fixed aspect ratio, a lighting description, sometimes an upscale pass on the end. The speed is genuine. So is the price you pay for it.

What Higgsfield-shaped tools are good at

Three jobs, and they are good at all three.

Volume with a consistent look. Forty product cards a week that must belong to the same family. A raw model gives you forty siblings that do not match. A preset gives you forty that do, because the same hidden scaffold is under each one.

Character identity as a saved object. The interesting feature in this category is identity training — Higgsfield calls its version Soul ID. You supply a set of photographs of a person once, the platform trains an identity, and then that person can appear across many generations without you rebuilding the setup each time. Under the surface this is the same idea as a LoRA, which lesson six covers properly; the platform difference is that it is three clicks instead of a training script.

Camera language as a control. Because tools like this grew out of video, they expose motion and lens vocabulary as buttons — crash zoom, dolly, low angle, bullet time — rather than expecting you to write it into a prompt and hope. If you do not have a cinematography vocabulary yet, this teaches you one by letting you see each term applied to your own frame.

What you give up, and when you notice

You notice at the revision.

The client likes the bottle shot but wants the light coming from behind the glass instead of the front. That decision was made inside the preset. It is not in your prompt, so no amount of rewriting the prompt reaches it. The preset's opinion about lighting is applied after your words and overrides them. You cannot get to a point halfway between two presets, either — they are discrete choices, not sliders.

This is the general rule and it is worth carrying beyond this one tool: **anything that packages a workflow buys you speed by removing exactly the controls you will need when the first answer is wrong.** It is a good trade for the ninety per cent of work that is routine, and a bad trade for the ten per cent that is the reason a client hired a person.

The credit arithmetic nobody checks

Platforms in this category run on credits, and within one platform the per-image credit cost across the available models can differ by nearly an order of magnitude.

A concrete measurement, from a real project in mid-2026: generating small avatar portraits on one such platform, a lightweight model cost 0.15 credits per image while two premium models cost 1.00 and 1.25. Seven to eight times the price. At the size the images were actually used — a 48-pixel avatar — no one could tell them apart.

The lesson generalises. Before you commit a batch, generate the same frame on the cheapest and the dearest model the platform offers, then look at both **at the size they will be seen**. Half the time the expensive one wins clearly and is worth it. The other half you have just found a way to make your allowance last eight times longer.

The free path

There is no free tier that gives you a preset platform. What there is, and what does the same job: **Fooocus**, which is a free local front end built on exactly this philosophy — good defaults, almost no visible settings, style presets in a dropdown — and runs on free Colab or Kaggle GPU hours if you have no card.

Canva's free tier packages generation into templates in a comparable way for social and print layouts, and works on a phone. Neither will give you trained character identity as a one-click object; you build that yourself in lesson six.

Use one honestly

A preset tool is not a lesser tool. It is a tool with a shape. Use it when the work is repetitive and the look is agreed. Reach for raw model controls when a specific person has a specific note.

Today: pick any preset you like the output of, then try to produce the *same* image with one deliberate change — a different light direction, a different lens. How far you get tells you exactly how much control that platform kept for itself.

BEFORE YOU MOVE ON

A team produces 40 product cards a week on a preset platform with no complaints. Then a client asks for one card where the light comes from behind the bottle rather than the front, and nobody on the team can produce it. What does this actually reveal?

A. The underlying model is too small to follow lighting instructions, so the team needs a platform running a larger model.

B. The preset has already fixed the lighting decision and applies it regardless of the prompt, so the fix is a tool that exposes lighting directly, or a manual relight in a raster editor.

C. The prompt needs to name the backlighting explicitly and in more detail than it currently does.

D. The preset needs retraining on backlit examples before it can produce them.

The answer and why: p. 382

CASE STUDY · MODULE 1

Two hundred assets, and a licence nobody had read

An illustrative case, built from documented patterns.

A two-person branding studio in Indore, eleven weeks into a retainer for a regional dairy co-operative.

Farhan and Divya run a studio out of a converted flat. Between them they had delivered two hundred and six images for the co-operative over eleven weeks: packaging mock-ups, market-stall backgrounds, a series of illustrated farmer portraits for the rebranded milk cartons, and a set of forty social cards. All of it had been generated locally, on a second-hand card Farhan had bought precisely so that per-image cost would stop being a line in their quotes.

The model was FLUX.1 [dev]. They had chosen it in week one because it held type better than anything else they could run, the samples on the model page were the best they had seen, and the download was free and took twenty minutes. Nobody read the licence file that came down with the weights. It was sitting in the folder the whole time.

In week eleven the co-operative's new marketing head asked a reasonable question: for the packaging artwork, which is going to a printer and onto half a million cartons, what are we licensed for? Divya opened the model card to copy a line into the email and found that FLUX.1 [dev] is released under a non-commercial licence. The weights may be used for research and personal work. Generating output that a client pays for is not that.

Two things were immediately clear and one was not. The first: this was a breach of the model licence, and it was entirely separate from any question about who owns the pictures. The second: it was not the client's fault and the client had no exposure they knew about. The third, the unclear one, was what to do.

The options were not equal.

Say nothing. Nobody had noticed in eleven weeks. The licence is not policed by anybody who would plausibly look at a dairy co-operative in Madhya Pradesh. The studio's reputation, their retainer and their next quarter all survive intact. The cost is that Divya would have to answer the marketing head's direct question with something that was not true, in writing, and then live with that sentence existing in an inbox for as long as the account did.

Rebuild everything. Re-generate two hundred and six images on Apache-licensed weights — FLUX.1 [schnell] or Qwen-Image, both of which permit commercial use — then repair, composite and grade them all again. Farhan estimated it at nine days of work they would not be paid for, and it would push two deliverables past dates the co-operative had already booked print slots against.

Rebuild only what is exposed. The packaging artwork and anything going to print is the part with real exposure and volume behind it. The social cards are ephemeral. That is about thirty images rather than two hundred, and it is defensible if you can say where the line was drawn and why.

Buy the licence. Black Forest Labs sells a commercial licence for the dev weights. The studio had never priced it because they had never thought they needed to.

Divya's position was that the third option was a rationalisation dressed as a plan: the licence does not distinguish between a carton and a social card, and drawing a line by commercial risk is deciding which breaches matter. Farhan's position was that nine unpaid days would take the studio's cash below the point where they could pay the rent in April, and that a rule you cannot afford to follow completely is still worth following partly.

The argument took a day. It was not really about the licence by the end of it. It was about whether a two-person studio that does not check terms before it starts is a studio a client should hire, and whether either of them wanted to keep working in a way that depended on nobody asking.

YOUR ANSWERS

1. Why did buying the commercial licence not, on its own, settle the question?

2. Farhan proposed rebuilding only the print work. What is the strongest argument for that position, and why did it lose?

3. What is the difference between the question the marketing head asked and the question about who owns the images?

STOP HERE

Write your answers before you turn the page. How it played out starts on p. 62.

This page is blank so that how it played out stays out of view until you turn the page.

CASE STUDY · MODULE 1

How it played out

They priced the commercial licence first, and it was affordable — less than one of the nine unpaid days would have cost them. They bought it, which made the past eleven weeks compliant going forward but did not retroactively license what had already been made, so Divya wrote to the marketing head the same afternoon, said plainly what had happened, said it was now licensed, and offered to regenerate the packaging artwork at no charge if the co-operative's own lawyers wanted a clean provenance. The co-operative asked for the packaging artwork to be regenerated, and nothing else. It took three days. The studio kept the account and added one step to their project template: a line in the brief naming the model, its licence and the date the licence page was read, filled in before the first render. Two months later that line stopped a different job from starting on a community merge whose ancestry nobody could establish.

The questions, discussed

1. Why did buying the commercial licence not, on its own, settle the question?

A licence bought in week eleven governs use from week eleven. The two hundred images already delivered were generated while the studio was not licensed to use the weights commercially, and no later purchase changes what happened. That is why the disclosure still had to be made: the fix was for the future, and the client was the one carrying the artwork.

2. Farhan proposed rebuilding only the print work. What is the strongest argument for that position, and why did it lose?

The strongest argument is proportionality: the packaging carries volume, permanence and a printer's audit trail, so it is where a claim would actually land, and spending nine unpaid days on ephemeral social cards buys almost no reduction in real exposure. It lost because the licence does not draw that line, so choosing it means deciding privately which breaches count — which is a position the studio could not have explained to the client if asked.

3. What is the difference between the question the marketing head asked and the question about who owns the images?

They are separate. Being licensed to use a model is a term you can look up on the model card in two minutes and which has a definite answer. Whether the output is protectable, and by whom, is unsettled and differs by country — India deems the person who made the arrangements to be the author, the United States requires human authorship. The studio failed the easy question, not the hard one.

MODULE 1 TEST

Check what stuck

6 questions, one right answer each. This one is for you, not for a record. The answers, and the reason behind each, are in the key on p. 380. If two go wrong, re-read the module before the exam.

1 You download FLUX.1 [dev], run it on your own card, and bill a client for the images. Which statement is accurate?

A. It breaches the model licence, whatever the position on who owns the pictures.

B. It is fine as long as you tell the client the images were made with AI.

C. It depends on whether your country grants copyright in computer-generated works to the arranger.

D. It is fine, because the weights were free to download and you generated the images yourself.

2 A client approves a bottle shot made on a preset platform, then asks for the light to come from behind the glass instead of the front. Why does rewriting the prompt not reach it?

A. Lighting terms only work on open weights, not on hosted services.

B. The prompt would have to exceed the model's token window to carry that instruction.

C. The platform charges more credits for a revision than for a first generation.

D. The lighting decision lives inside the preset and is applied after your words.

3 Your first generation after switching checkpoints takes four minutes; every later one on the same model takes fifteen seconds. What is the likeliest cause?

A. The VAE decodes at full resolution on the first pass and at half on later ones.

B. System RAM is short, so loading the checkpoint is swapping to disk.

C. The sampler is ancestral and injects fresh noise only on the first run.

D. The card is thermally throttling and recovers once the fans have spun up properly.

4 Generation reaches one hundred per cent and then fails with CUDA out of memory. Which fix addresses the actual peak?

A. Raise the batch size so the fixed setup cost is shared across more images.

B. Lower the step count, since every step holds its own set of activations in memory.

C. Turn on tiled VAE, because the decode is the peak rather than the sampling.

D. Switch to a non-ancestral sampler so memory is not reallocated on every step.

5 A character LoRA that was exact on fp16 is a near-miss on a Q4 model. What is happening, and what should you not do about it?

A. The checkpoint shipped a pruned VAE, so swap the VAE before touching anything else.

B. The LoRA was built for a different base checkpoint, so it should be retrained from scratch against the quantised file.

C. Q4 drops the text encoder, so the fix is to raise the LoRA's text-encoder weight.

D. Rounded weights land the LoRA's correction slightly off; do not simply raise its strength.

6 A client approves an image in March and asks for a small change in September. The same prompt and seed give a different picture. Which record would have prevented it?

A. The model hash, the interface commit, and every LoRA with its strength.

B. The negative prompt, which is the setting people most often forget to write down.

C. The model's name and the date you downloaded it.

D. A JPEG of the approved image, kept in the delivery folder.

wanted. Outpainting is covered later in the course, and this is the commonest reason to reach for it.

For a very wide or very tall final format — a roll-up banner at 1:2.35, a web skyscraper ad — do not try to generate it in one go. Generate the interesting part at a native ratio and extend outwards in steps. The result is coherent; the single-shot attempt is a stack of duplicated torsos.

Figure 2.1

Where SDXL composes, and where the work is delivered

Generate here — the trained buckets

- 1024 × 1024, square
- 1152 × 896, about 9:7
- 1216 × 832, about 3:2
- 1344 × 768, about 7:4
- 1536 × 640, 12:5 and the widest that holds

Deliver here — crop or extend afterwards

- 1080 × 1350 feed portrait: upscale, crop 39 px
- 1200 × 630 link card: outpaint the sides
- 1280 × 720 thumbnail: crop from 1344 × 768
- 2480 × 3508 A4 at 300 dpi: enlarge separately
- 5020 × 11811 roll-up: extend outwards in steps

Every bucket on the left is about 1,048,576 pixels. Outside them the network's receptive field covers more canvas than it ever saw, two regions each decide they are the centre of a portrait, and you get two torsos.

Checking a model rather than trusting a table

Bucket lists differ between model families and the table above is only for SDXL. A four-minute test tells you any model's real range: fix the seed and prompt, generate a simple single-subject portrait at 1:1, 4:5, 3:2, 16:9 and 2:1, and look at which ones hold. Write the answer next to the model in your notes. This is the sort of thing everybody re-learns from forum posts every few months and nobody measures once.

What this does not fix

Bucketing explains duplicated subjects at extreme ratios. It does not explain a duplicated subject at 1:1, which has a different cause — usually a prompt that describes a scene wide enough that the model reasonably renders several people, or a composition control pushing in two directions.

And there is an honest limit to the "generate native, crop later" rule. Cropping changes the composition the model chose. A frame that was well balanced at 3:2 can lose its subject's breathing room when taken to 4:5, and the fix is not a better crop but generating with the final crop in mind — which is what the next lesson is about. Ratio arithmetic gets you a frame with no duplicated limbs. It does not get you a frame that still works after the crop.

Resolution is a separate question from ratio

People conflate the two and then generate a 2048 × 2048 image expecting more detail. Past the trained pixel count you do not get a more detailed picture; you get the same failure as an extreme ratio, because the canvas is again larger than anything the model composed for. A 2048-square SDXL render very often contains two horizons, two suns, or a subject assembled from two overlapping attempts.

The route to a large image is therefore always the same: **generate at native size, then enlarge in a separate step**. An upscaler — an ESRGAN-family model, which is free, tiny and needs no GPU worth speaking of — takes a good

EXAMINATION

Image Generation, in Practice — final examination

This paper covers the whole course:

- assembling a working setup and knowing what it costs you
- turning a request into a written specification
- exploring to a frame worth keeping
- steering with images rather than adjectives
- repairing and extending and compositing
- holding a character and a product across a set
- finishing a file for the medium it is going to

56 questions, the whole pool. The site draws 25 for a sitting, pass mark 70%. Work any 25 under your own conditions. Answers from page 425.

1 Module 1 · The kit, and what each piece of it costs you

A 16 GB M2 MacBook against an 8 GB Nvidia desktop card, for SDXL work. What is the honest comparison?

A. The Mac is three to five times slower but can load models the 8 GB card cannot.

B. The Mac is faster, because unified memory removes the transfer over the PCIe bus entirely.

C. They are equivalent, since Metal and CUDA compile to the same kernels.

D. The Mac cannot run these models at all, because they require CUDA.

2 Module 1 · The kit, and what each piece of it costs you

What is the catch shared by Colab's free tier, Kaggle notebooks and Hugging Face Spaces for image work?

A. They cap output resolution at 512 pixels on the free tier.

B. They watermark every image that is generated on their free hardware.

C. They block diffusion models outright under their published acceptable-use policies.

D. Nothing persists, so you reinstall and re-download the model every session.

3 Module 1 · The kit, and what each piece of it costs you

Why does the course insist on .safetensors rather than .ckpt?

A. safetensors files are smaller, so they load faster from disk.

B. A .ckpt is a Python pickle, so loading it executes code from the file.

C. The .ckpt format cannot store fp8 or quantised weights.

D. Only safetensors files carry the model hash needed for reproducibility.

4 Module 1 · The kit, and what each piece of it costs you

You will run the same eleven-step process over image after image for a paying client. Which interface does the course recommend, and why?

A. Fooocus, because good defaults remove the eleven steps entirely.

B. Forge, because a settings panel is the fastest to learn and the tutorials match.

C. ComfyUI, because a saved graph does not decay the way notes do.

D. InvokeAI, because a unified canvas makes every repair one gesture.

5 Module 1 · The kit, and what each piece of it costs you

On Windows with an Nvidia card, speed collapses after you load a second model and nothing errors. What is it?

A. The second model overwrote the first and is being reloaded per step.

B. Windows paged the model out because the file was sitting on a network drive.

C. The scheduler silently fell back to a CPU implementation of attention.

D. The driver is spilling VRAM into system RAM instead of raising an error.

6 Module 1 · The kit, and what each piece of it costs you

On an 8 GB card, in what order should you attack the four consumers of memory?

A. Offload the text encoder, tile the decode, cut the batch, then quantise.

B. Reduce resolution first, because memory scales with pixel count.

C. Enable sequential CPU offload immediately, because it makes almost anything fit.

D. Quantise the weights first, since the network is the bulk of it.

Module 1 · The kit, and what each piece of it costs you

The module starts on p. 9.

MODULE 1 TEST

Module 1 test, Q1, p. 64
You download FLUX.1 [dev], run it on your own card, and bill a client for the images. Which statement is accurate?
Answer A: It breaches the model licence, whatever the position on who owns the pictures.
Why: The dev weights carry a non-commercial licence, and running them locally does not change that. Whether you own the output is a separate, unsettled question that differs by country; whether you are licensed to use the model has a definite answer you can read on the model card in two minutes.

Module 1 test, Q2, p. 64
A client approves a bottle shot made on a preset platform, then asks for the light to come from behind the glass instead of the…
Answer D: The lighting decision lives inside the preset and is applied after your words.
Why: A preset is a real workflow somebody froze: a chosen model, a prompt scaffold you never see, probably an adapter, a fixed ratio and a lighting description. Its opinion is applied after your words and overrides them, and there is no point halfway between two presets because they are discrete choices.

Module 1 test, Q3, p. 65
Your first generation after switching checkpoints takes four minutes; every later one on the same model takes fifteen seconds. What is the likeliest cause?
Answer B: System RAM is short, so loading the checkpoint is swapping to disk.
Why: Loading a large checkpoint usually means reading it into system memory first. A machine with 8 GB of RAM swaps to disk and takes minutes to do what a 32 GB machine does in twenty seconds. If the first generation after a model switch is glacial and later ones are fine, it is loading rather than inference.

Module 1 test, Q4, p. 65
Generation reaches one hundred per cent and then fails with CUDA out of memory. Which fix addresses the actual peak?
Answer C: Turn on tiled VAE, because the decode is the peak rather than the sampling.
Why: A failure that arrives after the progress bar completes is the decoder, not the sampler. Tiled VAE splits the decode into overlapping tiles so the final step stops being the memory peak, and it fixes this without touching resolution or batch size.

Module 1 test, Q5, p. 65

A character LoRA that was exact on fp16 is a near-miss on a Q4 model. What is happening, and what should you not do about…

Answer D: Rounded weights land the LoRA's correction slightly off; do not simply raise its strength.

Why: A LoRA encodes a correction to specific weights. When those weights have been rounded to a coarser grid the correction lands slightly off, so the effect is weaker or subtly wrong. Raising the strength over-applies the style while still missing the identity, so test once at a higher precision before spending an hour tuning.

Module 1 test, Q6, p. 66

A client approves an image in March and asks for a small change in September. The same prompt and seed give a different picture. Which…

Answer A: The model hash, the interface commit, and every LoRA with its strength.

Why: Model names are reused constantly and interfaces change defaults between versions. The facts that make an image reproducible are the model hash rather than its name, the interface version as a commit, the sampler and scheduler with steps, CFG and seed, every LoRA with its filename and strength, and the resolution and any upscaler.

THE LESSON CHECKS

Lesson 1, p. 14

A freelancer downloads FLUX.1 [dev], runs it on their own machine with no internet connection, and invoices a client for the resulting images. What is…

Answer C: The [dev] weights are released under a non-commercial licence, so billing a client breaches the model's terms whatever the copyright position on the images turns out to be.

Why: The licence on a model's weights governs whether you may use it commercially, and it is a completely separate question from who owns the picture it made.

A Nothing is wrong. A licence attached…: Reads a model licence as governing distribution only, when it also governs what the outputs may be used for.

B The images are unlikely to attract…: Confuses whether the output is protectable with whether the freelancer was permitted to run the model at all.

D Generating locally strips the Content Credentials…: Answers a provenance question when the question was about permission to use the model.

Lesson 2, p. 17

A team produces 40 product cards a week on a preset platform with no complaints. Then a client asks for one card where the light…

Answer B: The preset has already fixed the lighting decision and applies it regardless of the prompt, so the fix is a tool that exposes lighting directly, or a manual relight in a raster editor.

Why: Recipe-driven tools trade control for speed, and the bill arrives at the first revision, when the thing the client wants changed is the thing the preset already decided.

A The underlying model is too small…: Blames model capacity for a control the tool never exposed in the first place.

C The prompt needs to name the…: Assumes prompt wording can override conditioning that the preset applies after the prompt.

D The preset needs retraining on backlit…: Treats a packaged workflow as a model the user can retrain, rather than a fixed set of settings.

Lesson 3, p. 22

Somebody generates about 150 images a month and keeps hitting a hosted service's refusal filter on ordinary historical war imagery for a school project. Which…

Answer A: Control, not cost: at this volume hosting is a few dollars a month, so the real question is whose policy decides what they may make.

Why: The choice between a hosted service, your own card and a free cloud notebook is decided by how many images you make per month and how much control you need, and the crossover is usually somewhere around a few thousand images.

B Cost, because 150 images a month…: Treats the decision as volume-driven when 150 images at a few cents each is a few dollars a month; a card would take many years to repay at that rate.

C They should move to a free…: Half right about the filter but ignores persistence: nothing survives the session, so every run re-downloads the checkpoint and the setup, which is a poor fit for steady low-volume work.

D They should rewrite the prompts, since…: Assumes filters are precise. They combine a keyword layer with an output classifier and are deliberately over-inclusive, so ordinary historical language trips them regardless of intent.

MAKE THINGS

Image Generation, in Practice

Reference images, inpainting, character LoRAs, upscaling, and the licence nobody read.

WHAT IT TEACHES	The craft of getting a specific, finished image out of a model on a deadline. Choosing between hosted tools and open weights on cost, control and licensing.
WHAT IS INSIDE	Seven modules and 69 lessons, 27 figures drawn from data, seven case studies, seven module tests, a 56-question examination, and every answer with the reason behind it.
LICENCE	CC BY-NC-SA 4.0. Copy, print, share and adapt it for any non-commercial purpose, with credit, under the same licence. There is no paid edition.
THE COURSE	Free to read, with the lessons, the films and the examination, at **addaly.com/learn/ai-images**

Scan to open the course

Addaly

First edition, September 2026 · build 62a26075