

🔴 *Alert! This guide is no longer actively updated, it is now a separate chapter in the [UNOFFICIAL MIDJOURNEY MANUAL](#) which you should go check out!* 🔴

IMAGE PROMPTING AND--

You? Again?

Look, it's fun as hell, OK?

Alright. What's the deal with image prompts?

Well, Jerry, image prompts give Midjourney image(s) to use as inspiration. Mechanically, what "inspiration" means is that the AI reverses its text-to-image neural doohickey and converts your image into a bunch of text prompts (in its own internal language, not English), crams all those text prompts into a dense boullion cube, and tosses it into the diffusion stock-pot at a default weight of 0.25 for your image vs 1 for your text.

MATH NOTE: 0.25 is a value, not a percentage. The image's default "percentage" with a single text prompt is 20%, since 0.25 is 20% of 1.25. Get it? Using `--iw 1` makes them 50/50. Returns on image weight diminish at high levels; I haven't found it useful above 5. Please support [the official documentation](#).



"Can you describe your assailant, his weapon, and a funny occupation to our improv sketch artist?"

Always keep in mind that the AI doesn't use the actual image when diffusing. This ain't metafilter, it's not modifying your image or using the pixel data in any way other than to derive its internal prompt description.

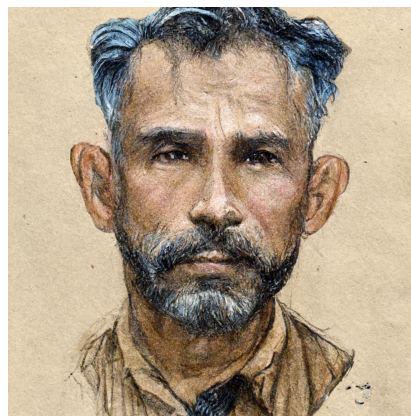
Image prompting is basically MJ letting one side draw what the other side describes to it over the phone. (Except since we don't know Midjourney's language, it can actually pass on some details more precisely than us prompting them by text would be!)

Can we skip to the fun stuff?

Face prompting! Find yourself constantly pleading with MJ for stuff like `--no extra face, extra eyes, just be normal please` or having to mimic celebrities to get consistent faces? Here's how to bootstrap your very own original character (OC) on MJ:

Step 1: Describe them and generate a headshot. My preferred starting medium is a color pencil sketch because MJ does pretty good faces with pencil. For your first prompt, don't describe anything from the neck down, or anything in the background. Stick to the basics: age, gender, ethnicity, hairstyle, and other facial details like earrings or cosmetics (including "no makeup" for women if applicable). No hats. If you have a personality in mind, something like "nervous" or "smug" can add flavor too. I would recommend `--s 1250` for this first sketch (default stylization currently makes faces too samey for my taste).

You can even prompt with a fake name to add some directed randomness. For instance, you can say something like "Osvaldo Cardenas, a thin, grizzled, 45-year-old Puerto Rican man, crew cut, grey stubble, portrait, color pencil sketch `--s 1250`" and it will add some flavor to your headshots, presumably based on the average of every man in its dataset whose pictures were labeled with "Osvaldo" and "Cardenas."

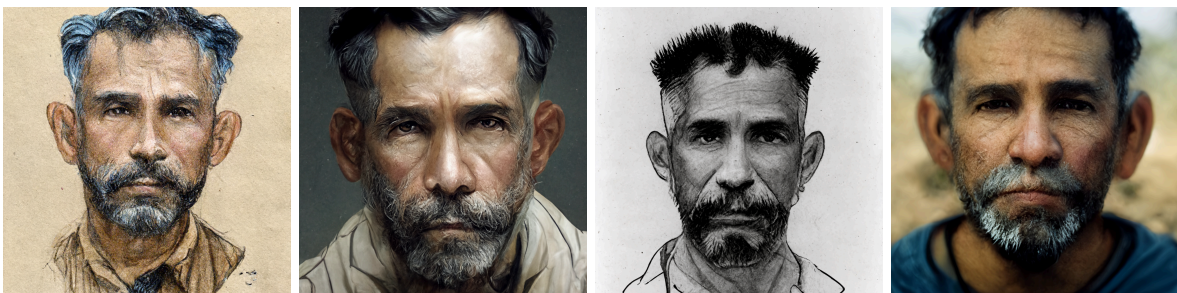


Didn't prompt the ears, but I like them. Gives him character!

Step 2: Now let's use that picture to bank a few more headshots. Add the pencil sketch you picked in Step 1 as an image prompt, but in the text prompt, change `colored pencil sketch` to `digital matte painting`. Then lower to `--s 625` at the end (to minimize the effect of MJ "house style" for this step) and add `--iw 0.75`. What we're trying to do here is get a fancy matte painting of the same character we just did in pencil. Not necessarily their specific pose or angle, remember it's not tracing, just something that keeps the character intact. At this step you should feel free to also add a bunch of the usual "make look cool" stuff like `trending on artstation`, `cg society`, `artgerm` etc.

You can dial up the image weight if you want it to stay more "on model" or dial it down to keep the option open of tweaking details via text prompt. I took "crew cut" out at one point here because reinforcing it with image AND text was making his sides totally shaven, while the image alone was enough to push it through. (Which details can be pushed through via image and which need repeating in text depends heavily on things like how big the rest of your prompt is, how far out the face is framed, your `--iw` and `--s` values, etc. Experiment!)

In my original tweet I said you could use the pencil headshots right off, and you can, but I've since realized that MJ seeing that it's a "pencil sketch" can smudge your character downstream since that information gets passed along with the facial features. So my current practice is to have a bank of images for each character: one "color pencil sketch", one "digital matte painting", one "ink outline drawing", and one "Zeiss 35mm photograph". These are arbitrary choices; feel free to have stuff like "watercolor" or "anime-style" in your own OC bank. The important thing is a couple of recognizably similar faces in different formats so no one medium dominates.



You may need a lower `--iw` for the "photograph" headshot. Feel free to be picky here with a few re-rolls, you'll be using these a lot.

Author's note 8/31 - The "remaster" feature currently does not support multiple image prompts. So for things that you intend to remaster in the higher-level engines, just use one headshot, whichever best fits your medium. Multiple faces are still recommended for V3 creations (and so you have choices for different mediums to "remaster").

Step 3: Armed with a couple of headshots, we can start doing fun stuff! Stack all four (or however many) as image prompts, swap portrait for medium shot or full shot, pick out a costume, some scenery, and an aspect ratio. You can tune your stylizing and image weight as much as you feel like, depending on how accurate you want or need the face prompting to be. Even if you don't necessarily need a recurring character for storytelling purposes, I've found them to improve the average face result in a given prompt, because `--no multiple faces` and other such things are now implied.



<https://s.mj.run/wwAWhIECiiA> https://s.mj.run/1FDVUw-8d_8 <https://s.mj.run/MbcH5-OK5ZU>
<https://s.mj.run/xJaJiLPDUB0> osvaldo cardenas is an elizabethan actor rehearsing shakespeare at the globe theatre to empty seats, oil painting, medium shot --ar 3:2 --s 1250 --iw 0.75

Okay, nice. This was the first result for that prompt, no re-rolls, and it's non-optimized natural language that I actually made a little hard on MJ by taking out all of the face details and passing the name instead (effectively a chaos prompt at this point), but there are still three faces here that seem to have the right idea. Let's tighten things up a bit more by reinforcing an abridged description instead of his name, and buff the image weight a bit so we can get a little more crazy.



<https://s.mj.run/wwAWhIECiiA> https://s.mj.run/1FDVUw-8d_8 <https://s.mj.run/MbcH5-OK5ZU>
<https://s.mj.run/xJaJiLPDUB0> "THE OTHER O.C." :: movie poster for "THE OTHER O.C." featuring a middle-aged puerto rican man with short hair, cinematic lighting, clean typography --ar 2:3 --s 1250 --iw 1

Did I re-roll this to flex? Yes. Did I re-roll it an excessive amount to flex? Nah, read "Multiprompting and You" on Discord. Do, uh, three-and-a-half of the four faces up there seem legibly similar? Hopefully, yes. (I also got four good faces but I love my text memes.)



<https://s.mj.run/wwAWhIECiiA> https://s.mj.run/1FDVUw-8d_8 <https://s.mj.run/MbcH5-OK5ZU>
<https://s.mj.run/xJaJiLPDUB0> "TEXT IS PLAYED OUT" :: "TEXT IS PLAYED OUT" in a speech bubble from a middle-aged puerto rican man with short hair --ar 3:2 --iw 1 --s 1250

I didn't ask you, Ozzie. Get back to wardrobe!



<https://s.mj.run/wwAWhIECiiA> https://s.mj.run/1FDVUw-8d_8 <https://s.mj.run/MbcH5-OK5ZU>
<https://s.mj.run/xJaJiLPDUB0> a thin man wearing a fedora and trenchcoat, medium shot, noir illustration, dramatic street lighting, trending on artstation --ar 16:9 --iw 3 --s 625

Ultra-minimal "thin man" text reinforcement, and a widescreen face is pushing it even with --iw 3 and medium shot only partially moving the camera out. But still! Faces!



Here's some more messing around, I'm too lazy to c/p all 4 prompts.

Some final considerations: the fact that your seed images are all portraits, while great for face definition, does have side effects, most noticeably that it gets harder to “pull the camera out.” Asking for a full shot will often result in a medium shot, etc. So you may need to do something like multi-prompt a guy, full shot :: the rest of your stuff and weight that up in order to really guarantee a full shot. Or do the usual stuff like prompt footwear, vertical --ar, etc.

Also, don't underestimate the power of low (or rather, modestly buffed) image weighting when you really want to try something wild with the character. I used to use --iw 2 as a standard but these days I mostly do --iw 0.75 and even lower sometimes. The whole point of face prompting is to exploit the AI's ability to interpret prompts in its own language so often you don't need a super-high weight to “fill in the gaps.”

Alright, alright, I'm tired of faces. What else is there?

Style transfer! There are lots of artists too obscure to reliably prompt by name, and this can lead to undesired behaviors like MJ attempting to write out the name, or interpreting a Chinese artist as “oh this name sounds vaguely Chinese, you must want a portrait of a random Chinese guy.”

While face prompting depends on collecting the same subject in different styles, style transfer is more or less the opposite. What you want to do here is grab a couple of your favorite pieces from a particular artist with differing subjects, which will hopefully cancel each other out during MJ's image prompt parsing and leave the remainder of the focus on the unifying style.



Not Midjourney™ images. Obviously.

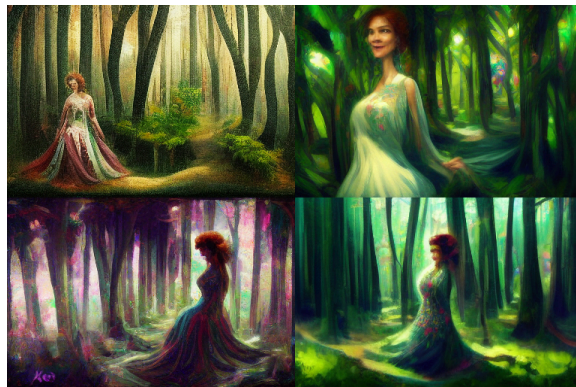
Let's take the example of Kuno Veeber, a negative entry I picked randomly from [the MJ artist reference spreadsheet](#) because the name sounds funny. Wikipedia informs me that he's an Estonian oil painter whose work is in "cubism," "constructivism," and "expressionism." Above, you'll see four paintings of his that I pulled off of Wikimedia. I didn't really pick them out carefully and the only adjustments I made were cropping out frames when they were part of the image (since I don't want it to trigger drawing frames on the outside for the purposes of this test).

Now let's run a quick test to confirm that his name isn't strong enough to trigger a distinct style:



a woman in an enchanted forest, in the style of kuno veeber --ar 3:2 --seed 12345

Pink skies, digital detail, the default girl in V2. Sure looks like Midjourney house style to me, and definitely doesn't look like a cubist oil painting from the 1920s. (I'm locking --seed 12345 here to ensure like-for-like comparisons going forward.) Just for kicks here's one with --s tuned back to minimum:



a woman in an enchanted forest, in the style of kuno veeber --ar 3:2 --s 625 --seed 12345

Possibly a little more oil-painting-ish? But we're not getting any of the broad shapes and blank faces in his work (if you zoom in, a face is being attempted in V1, the resolution just isn't there). Lots of artists are in the dataset a little bit but it takes a certain threshold of frequency and consistency in the menchiefs for MJ to be capable of their style by name.

So now let's load up those four paintings as image prompts on top of his name and see what happens.



<https://s.mj.run/lxCw-7sO6AM> <https://s.mj.run/8qG35i6xeQU> <https://s.mj.run/7Theuom6gQA>
<https://s.mj.run/6R1pgTeMujY> a woman in an enchanted forest, in the style of kuno veeber --ar
3:2 --s 625 --seed 12345

Even at default `--iw`, you can see an immediate and clear difference in most of these with the prompts added. The shapes become broader, the human figure is much less defined, and MJ has much more of an oil painting vibe than it did before. This is definitely closer than with just the name. And unlike most of the face transfers, we've done nothing via text or increased `--iw` to help it along. This is intentional; high image weights here will start crowding out your subjects and simply telling it "cubist, oil painting" would be cheating. But hell, let's cheat a bit just to flesh out the example:



<https://s.mj.run/lxCw-7sO6AM> <https://s.mj.run/8qG35i6xeQU> <https://s.mj.run/7Theuom6gQA>
<https://s.mj.run/6R1pgTeMujY> a woman in an enchanted forest, in the style of kuno veeber --ar
3:2 --s 625 -iw 0.75 --seed 12345



<https://s.mj.run/lxCw-7sO6AM> <https://s.mj.run/8qG35i6xeQU> <https://s.mj.run/7Theuom6gQA>
<https://s.mj.run/6R1pgTeMujY> a woman in an enchanted forest, in the style of kuno veeber, oil painting, cubist, expressionist --ar 3:2 --s 625 -iw 0.5 --seed 12345

Tripling the image weight (to 0.75) gives us even more broad shapes, but sort of kicked us out of the forest a little bit and left more confused/undiffused space as MJ is now importing more subject cues from the images (which we selected because of their clashing subjects). Doubling it and adding “expressionist, cubist, oil painting” brings in way too much Picasso, who is just too popular and synonymous with this genre to avoid as the shapes get more angular and dense than what we can see in Veeber’s work. Default to slight-increase with minimal text reinforcement is the way to go here. (Unless you’re also doing multiprompt weighting; remember 2 prompts = you need to double your image weight to keep pace with the text.)

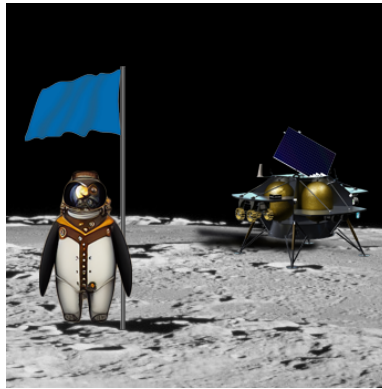
That’s cool and all, but what if I’m not an artist and just want funny meme pictures?

Image mockups! MJ is notoriously bad at taking instruction for complex scenes involving multiple subjects, particularized colors, or posing certain things above/below/beside each other. This, again, is because MJ’s internal language only sort-of overlaps with what we can specify in a text prompt. One way you can nudge your odds of interpretation up is by making a mockup of your mise-en-scene as a visual image.

I want to emphasize that these can be incredibly crude. For one thing, MJ reads and clips your image prompts at a maximum resolution of 256x256, so quality is mostly wasted on it. We’re just trying to visually get across subjects/composition. You can use separate MJ generations, creative-commons clip art, even MS Paint can help a bit.

Say you want a penguin putting a flag on the moon next to a lunar lander, as someone asked about in prompt-craft. I had MJ give me a penguin in a (steampunk) spacesuit

separately, cut it out in Photoshop, then layered it with a clip art flag and two NASA images of a lunar lander concept and the moon's horizon:



Not a Midjourney™ image. Get your eyes checked.



(L) <https://s.mj.run/93Fr29HHXS8> two subjects: a spacesuited penguin placing a flag next to a landing module with mechanical legs, the gray surface of the lunar horizon, matte painting, trending on artstation, professional digital art, hd --iw 1 --ar 5:3
(R) <https://s.mj.run/93Fr29HHXS8> two subjects :: a penguin wearing a spacesuit, planting a flag on the moon :: a penguin astronaut next to a lunar lander :: a lunar lander on the horizon --iw 4 --ar 3:2 --s 1250

And here's two results, one with a single prompt I did a few weeks ago, one I tried just now with a multi-prompt (which is why the image weight is so much higher on the right; four text prompts means 4x image weight to keep them "equal"). No one will confuse these with Dalle2 output, but given how much skinnier MJ's model is, it's pretty damn acceptable.

Note that unlike multi-prompted text converging over time, image prompting is a one-shot deal. After the first grid your prompted image won't be used, so this is more about reducing the number of re-rolls to get a promising initial diffusion you can then refine with v-rolls. Very useful for something like two subjects with two different faces in your result.



(L) Not a Midjourney™ image, again, please. This is *inspiration*.

(R) <https://s.mj.run/uz4Eum6vqJw> three subjects, shaded pencil sketch :: dr. house and dr. wilson looking at a wombat :: a wombat shown to two grumpy doctors --no color --iw 4

Still not Dalle2 (forgot to draw Wilson's nose), but still getting an impressive amount of detail with both text and visual arguments, and also, still very funny.

Okay, just dump your last paragraph here and we'll finish this sucker up later.

CHAOS! One neat thing about image prompting is it makes your result highly indeterminate. With two images and a healthy amount of text you can get wildly differing results even with the same seed. If you're not much for prompt engineering and prefer a more experimental approach, just start kit-bashing with wild combinations of images, prompts, and weights, and let MJ take what it feels like from all of them. In all likelihood you will stumble on stuff even cooler than what I've listed above.

Davey Bossman himself has said in multiple office hours that he thinks image prompts are an extremely under-utilized tool in creating wild images. So please treat this FAQ as a guide to cinematography written in 1896 by a guy passing on second-hand rumors about how cameras work.