

# Images et IA

(page créée le 8 nov 2021)

“Text-to-image is the new Sudoku.”, [Mario Klingemann](#)

Listes de ressources

- <https://pharmapsychotic.com/tools.html>
- <https://github.com/Maks-s/sd-akashic>

## Génération d'images en ligne «Text-to-image»

### Stable Diffusion

Stable Diffusion est un modèle de génération d'images, créé par [Stability AI](#), le modèle a été entraîné sur 2,3 milliards d'images du dataset [LAION-5B](#)

#### mage.space

Service gratuit / réseau social permettant d'utiliser Stable Diffusion (NSFW) en ligne (version 1.5 le 28/11/2022). Permet aussi de partir d'une image existante pour en modifier une partie («img2img»).

- <https://mage.space>
- [https://twitter.com/MageSpace\\_](https://twitter.com/MageSpace_)

#### dreamstudio

(Accès gratuit limité) Service d'accès à Stable Diffusion par StabilityAI

- <https://beta.dreamstudio.ai/home>

#### PlaygroundAI

(?) Service d'accès à Stable Diffusion

- <https://playgroundai.com/login>

#### moteur de recherche

Dans des images créées par Stable Diffusion, images présentées avec les prompts :

- <https://lexica.art> + recherche inversée (voir aussi [stable-diffusion-prompts](#))
- <https://www.krea.ai>
- <https://arthub.ai/>
- <https://openart.ai/>

#### explorer le dataset

- dans 12 millions d'images : [https://laion-aesthetic.datasette.io/laion-aesthetic-6pls/images?\\_sort=rowid](https://laion-aesthetic.datasette.io/laion-aesthetic-6pls/images?_sort=rowid) (article sur le projet)
- <https://rom1504.github.io/clip-retrieval/>

- Vos images personnelles ont-elles été aspirés dans le dataset LAION-5B ? <https://haveibeen trained.com/>.
- LAION5B explained : <https://www.ivanzhou.me/blog/2022/12/4/paper-explained-laion-5b>

## aide à l'écriture de prompt

- <https://promptomania.com/stable-diffusion-prompt-builder/>
- Utiliser chatGPT pour écrire des prompts : [https://www.reddit.com/r/StableDiffusion/comments/11ezpcg/chat\\_gpt\\_stable\\_diffusion\\_prompt\\_generator/](https://www.reddit.com/r/StableDiffusion/comments/11ezpcg/chat_gpt_stable_diffusion_prompt_generator/)

## temps réel avec Streamdiffusion

- <https://github.com/cumulo-autumn/StreamDiffusion>
- <https://alltd.org/real-time-diffusion-in-touchdesigner-streamdiffusiontd-setup-install-settings/> intégration dans touchdesigner

## divers

- <https://www.reddit.com/r/StableDiffusion/>

## vocabulaire

Des principes/procédés/logiciels qu'on retrouve dans l'univers de Stable Diffusion :

**auto1111 / automatic1111** : interface web pour utiliser Stable Diffusion avec de nombreuses options, automatic1111 est l'auteur original qui a donné son nom au logiciel, mais le nom du projet est plutôt Stable Diffusion web UI

- <https://github.com/AUTOMATIC1111/stable-diffusion-webui>

## checkpoint / .ckpt

**ControlNet** : permet d'utiliser une image comme guide, avec beaucoup de subtilités! (utilisable dans automatic1111) par exemple : utiliser une image d'une pose corporelle comme guide appliquera cette pose au personnage créé par le prompt. Il existe différents parfums de ControlNet : ControlNet Canny, ControlNet openpose, ControlNet HED avec leurs spécificités

- <https://github.com/lllyasviel/ControlNet>
- <https://aituts.com/controlnet/>
- [https://www.reddit.com/r/StableDiffusion/comments/119o71b/a1111\\_controlnet\\_extension\\_explained\\_like\\_youre\\_5/](https://www.reddit.com/r/StableDiffusion/comments/119o71b/a1111_controlnet_extension_explained_like_youre_5/)
- [https://www.reddit.com/r/StableDiffusion/comments/11cwiv7/collected\\_notes\\_and\\_observations\\_on\\_controlnet/](https://www.reddit.com/r/StableDiffusion/comments/11cwiv7/collected_notes_and_observations_on_controlnet/) description des différentes méthodes
- **conteolnet 1.1 en détail** <https://github.com/lllyasviel/ControlNet-v1-1-nightly>

**deforum** : création d'animations à partir de prompts et d'indications sur l'animation (rotation, zoom, etc.)

- <https://deforum.github.io>
- <https://github.com/deforum-art/deforum-stable-diffusion>
- <https://docs.google.com/document/d/1RrQv7FntzOulG4ohjRZPVL7iptlyBhwwbcEYEW2Ofcl/edit>
- <https://github.com/deforum-art/sd-webui-deforum/wiki>
- <https://aituts.com/deforum/> bonne intro
- <https://stable-diffusion-art.com/deforum/> autre intro
- <https://reentry.org/AnimAnon-Deforum> tous les paramètres expliqués
- [https://docs.google.com/document/d/1pEobUknMFMkn8F5TMsv8qRzamXX\\_75BShMMXV8IFsll/edit](https://docs.google.com/document/d/1pEobUknMFMkn8F5TMsv8qRzamXX_75BShMMXV8IFsll/edit) guide pour Deforumv05 (pdf possible)
- <https://reentry.org/AnimAnon> excellent guide général sur l'animation
- [https://replicate.com/deforum/deforum\\_stable\\_diffusion/examples](https://replicate.com/deforum/deforum_stable_diffusion/examples)

**deforum/parseq** : parseq est un séquenceur de paramètres pour deforum

- <https://reentry.org/AnimAnon-Parseq>
- <https://github.com/rewbs/sd-parseq>
- un exemple : <https://www.youtube.com/shorts/fgiev3A93RU>

**diffusion** : procédé de création des images qui part d'une image bruitée pour la débruiter en la faisant correspondre au prompt

**dreambooth** : méthode pour compléter l'entraînement d'un modèle à partir de photos personnelles

- <https://dreambooth.github.io/>

### embeddings

**img2img** : image to image = modifier une image de base par un prompt

**inpainting** : modifier uniquement une partie d'une image préexistante

- inpainting avec automatic1111 : <https://onceuponanalgorithm.org/using-inpaint-in-stable-diffusion-tutorial/>

**instruct pix2pix** ; une forme particulière d'img2img qui permet de modifier des images existantes par une instruction tout en gardant leur structure, ex : "swap sunflowers with roses", "turn it into a still from a western"

- <https://github.com/timothybrooks/instruct-pix2pix>

### invisible watermark

- <https://medium.com/@steinsfu/stable-diffusion-the-invisible-watermark-in-generated-images-2d68e2ab1241#504e>

**invokeAI** : logiciel web pour générer des images (inpainting, outpainting, etc.)

- [https://github.com/invoke-ai/InvokeAI/blob/main/docs/features/UNIFIED\\_CANVAS.md](https://github.com/invoke-ai/InvokeAI/blob/main/docs/features/UNIFIED_CANVAS.md)

**Latent Couple** : composition contrôlée, segmentation, plusieurs prompts

- [https://www.reddit.com/r/StableDiffusion/comments/11jmtel/basic\\_guide\\_7\\_using\\_latent\\_couple\\_controlnet\\_to/](https://www.reddit.com/r/StableDiffusion/comments/11jmtel/basic_guide_7_using_latent_couple_controlnet_to/)
- <https://github.com/opparco/stable-diffusion-webui-two-shot>

### latent diffusion

#### latent space

**LoRA** (Low-Resource Adaptation) : entraînement complémentaire d'un modèle, "fine-tune" / réglages fins (voir aussi dreambooth, textual inversion, qu'il y a des méthodes alternatives)

- <https://www.shruggingface.com/blog/self-portraits-with-stable-diffusion-and-lora>
- <https://www.shruggingface.com/blog/blending-artist-styles-together-with-stable-diffusion-and-lora>

### model

- <https://reentry.org/sdmodels>

### modifier

**outpainting** : "agrandir" une image en la complétant avec un prompt

**paint-with-words** : segmentation d'images, multiples prompts

- <https://github.com/cloneofsimon/paint-with-words-sd>

**prompt** : énoncé texte fourni pour produire l'image, selon les logiciels cet énoncé peut être renforcé par des signes de ponctuation (parenthèses, crochets, etc.) pour donner plus ou moins de poids à certains mots ou expressions

### sampler / sampling method

- (SD v1.4)  
<https://proximacentaurib.notion.site/SD-Steps-vs-CFG-vs-Sampling-Method-e8765704d8a6457ca3f66058466fe43a>

**T2I adapter** : transfert de style, composition contrôlée

- <https://github.com/TencentARC/T2I-Adapter>
- [https://www.reddit.com/r/StableDiffusion/comments/11j07iu/style\\_t2i\\_adapter\\_model\\_mikubills\\_controlnet/](https://www.reddit.com/r/StableDiffusion/comments/11j07iu/style_t2i_adapter_model_mikubills_controlnet/)

## textual inversion

**token** : plus petit élément quand un prompt est décomposé (une sorte d'“atome” de langage, syllabe mais pas vraiment, etc.)

## tokenizer

**txt2img** : text to image = générer une image depuis un prompt

## upscale

- <https://github.com/AUTOMATIC1111/stable-diffusion-webui/discussions/6509?ref=learn.rundiffusion.com>

## Dall-E mini

Création d'images à partir d'un énoncé («*prompt*») basé sur le modèle Dall-E Mini. Rapide et facile à utiliser

<https://www.crayon.com/>

## Dall-E 2

(Accès gratuit limité) Création d'images à partir d'un énoncé («*prompt*»)

- <https://openai.com/dall-e-2/>
- aide de l'éditeur : <https://help.openai.com/en/articles/6516417-dall-e-editor-guide>

## MidJourney

(Accès gratuit limité) Création d'images à partir d'un énoncé («*prompt*»). Accessible sur Discord

- <https://www.midjourney.com/home/>

## Ressources

- styles dans midjourney : <https://ckovalev.com/midjourney-ai/guide/how-to-apply-any-style-to-any-prompt-in-midjourney-ai>

## NijiJourney

= MidJourney x Spellbrush (Accès gratuit limité). Modèle entraîné et spécialisé sur les images d'anime)

- <https://nijijourney.com/>

## VQGAN+CLIP

Création d'images à partir d'un énoncé («*prompt*»)

- [https://huggingface.co/spaces/akhaliq/VQGAN\\_CLIP](https://huggingface.co/spaces/akhaliq/VQGAN_CLIP) (temps d'attente variable)
- (fév 2022) [https://huggingface.co/spaces/EleutherAI/VQGAN\\_CLIP](https://huggingface.co/spaces/EleutherAI/VQGAN_CLIP)
- [https://colab.research.google.com/drive/1\\_4Jl0a7WlJeqy5LTjPjFzOwMZopG5C-W?usp=sharing#scrollTo=g7EDme5RYCrt](https://colab.research.google.com/drive/1_4Jl0a7WlJeqy5LTjPjFzOwMZopG5C-W?usp=sharing#scrollTo=g7EDme5RYCrt) (colab / google nécessite un compte google)

Explication du principe : <https://lvmiranda921.github.io/notebook/2021/08/08/clip-vqgan/>

Des exemples sur reddit

- [https://www.reddit.com/r/artificial/comments/n5jyet/statues\\_photo\\_input\\_with\\_clip\\_vqgan\\_image/](https://www.reddit.com/r/artificial/comments/n5jyet/statues_photo_input_with_clip_vqgan_image/)
- [https://www.reddit.com/r/deepdream/comments/q1o6l5/what\\_prompts\\_do\\_i\\_feed\\_to\\_get\\_such\\_images/](https://www.reddit.com/r/deepdream/comments/q1o6l5/what_prompts_do_i_feed_to_get_such_images/)

Une plateforme de publication : <https://creator.nightcafe.studio/top>

Une analyse des artistes/bédéistes/illustrateur·ice·s dans VQGAN : <https://remidurant.com/artists/#> (déjà la préhistoire en

2023)

## ruDALL-E

Autres modèles, russes cette fois : <https://rudalle.ru/en/>

## Le prompt qui va bien

Des ressources : <https://pharmapsychotic.com/tools.html#sec-1f7f>

### How-to, books

- exemple de *modifiers* : <https://promptomania.com/stable-diffusion-prompt-builder/>
- stable diffusion 2.1 prompt book : <https://app.usp.ai/static/Stable%20Diffusion%202.1%20Prompt%20Book%20by%20USP.ai.pdf>
- guide prompts pour le modèle RPG v4 (mais informatif et général) : <https://huggingface.co/Anashel/rpg/resolve/main/RPG-V4-Model-Download/RPG-Guide-v4.pdf>
- un article : <https://medium.com/@silkworm/high-quality-art-generation-with-the-prompt-suggestion-or-search-dbd59fb0f538>
- <https://strikingloo.github.io/stable-diffusion-vs-dalle-2>
- <https://photogpedia.com/ai-art-generators-prompt-building/>
- <https://docs.google.com/document/d/17VPu3U2qXthOpt2zWczFvf-AH6z37hxUbvEe1rjTsEc/edit>
- <https://github.com/joelparkerhenderson/stable-diffusion-image-prompt-gallery>
- <https://serokell.io/blog/stable-diffusion>
- <https://www.unlimiteddreamco.xyz/2022/03/16/writing-good-prompts-part-1.html>
- <https://matthewmcateer.me/blog/clip-prompt-engineering/>

### image to prompt

Plusieurs services en ligne proposent des prompts à partir d'une image qu'on leur envoie

- pour Stable Diffusion v1.x : <https://huggingface.co/spaces/pharma/CLIP-Interrogator>
- pour Stable Diffusion v2 : <https://huggingface.co/spaces/fffiloni/CLIP-Interrogator-2>
- <https://huggingface.co/spaces/pharma/sd-prism>
- <https://replicate.com/pharmapsychotic/clip-interrogator>

Comment ça marche ? : <https://medium.com/@silkworm/diversify-photo-database-with-clip-interrogator-5dd1833be9f5>

### Conversion de prompts

De stable Diffusion v1.x à SD v2.0 : <https://huggingface.co/spaces/fffiloni/prompt-converter>

### Vocabulaire

- listes de mots [https://github.com/pharmapsychotic/clip-interrogator/tree/main/clip\\_interrogator/data](https://github.com/pharmapsychotic/clip-interrogator/tree/main/clip_interrogator/data)
- modifier studies : <https://proximacentaurib.notion.site/2b07d3195d5948c6a7e5836f9d535592?v=e8062bc85bfd43a99fb6e187e9bac926>
- artist style studies : <https://proximacentaurib.notion.site/e28a4f8d97724f14a784a538b8589e7d?v=42948fd8f45c4d47a0edfc4b78937474>
- <https://www.the-ai-art.com/modifiers>

## Animation d'image fixe

### AdaMPI

- <https://yxuhan.github.io/AdaMPI/>
- <https://replicate.com/pollinations/adampi>

## 3D Photo inpainting

- <https://shihmengli.github.io/3D-Photo-Inpainting/>
- <https://replicate.com/pollinations/3d-photo-inpainting>

## Services privés en ligne

Plusieurs applications et services en ligne proposent d'animer des images fixes, par exemple : pour animer le visage d'un portrait, d'une photo ancienne, etc.

**cutout.pro AI photo animer** <https://www.cutout.pro/photo-animer-gif-emoji> email + gratuit pour le téléchargement basse définition avec watermark, payant pour une meilleure qualité

**deep nostalgia** <https://www.myheritage.fr/deep-nostalgia> (attention à la société qui propose ce service!)

**leai.pix** <https://convert.leiapix.com/> transformation d'images fixes en images 3D lightfield (utilise [depthy](#) et [BoostingMonocularDepth](#))

## Augmentation de la définition d'une image

### Restauration d'images

swinir (usage gratuit sur Replicate) : <https://replicate.com/jingyunliang/swinir>

D'autres possibilités sur replicate (gratuits ou payants)

- <https://replicate.com/collections/super-resolution>
- <https://replicate.com/collections/image-restoration>

### Real ESRGAN

Augmentation de la définition d'une image

- <https://replicate.com/nightmareai/real-esrgan>
- <https://huggingface.co/spaces/akhaliq/Real-ESRGAN>

Article : <https://github.com/xinntao/Real-ESRGAN#Portable-executable-files>

## Interpolation

### FILM

- <https://film-net.github.io/>
- <https://replicate.com/google-research/frame-interpolation>

### RIFE

- <https://replicate.com/pollinations/rife-video-interpolation>

### RealBasicVSR

«Real-World Video Super-Resolution»

- <https://github.com/ckkelvinchan/RealBasicVSR>
- <https://arxiv.org/pdf/2111.12704.pdf>

## Autres principes

<https://toonify.photos/>

# Recherche

## DragGAN

Edition d'images basée sur l'apprentissage automatique

- <https://github.com/XingangPan/DragGAN>
- <https://vcai.mpi-inf.mpg.de/projects/DragGAN/>

## DAIN (Depth-Aware Video Frame Interpolation)

Interpolation d'images : à partir d'une vidéo en stop motion, recréer la fluidité d'un tournage classique

- <https://github.com/baowenbo/DAIN>

## Dreambooth

Génération fine d'images à partir d'une image source avec changements de contexte, de propriétés, etc.

- <https://medium.com/@silkworm/dreambooth-how-google-hacks-diffusion-model-to-generate-personalized-photos-b4721763f0f8>
- <https://dreambooth.github.io/>

## CLIP (Contrastive Language-Image Pre-Training)

- <https://github.com/openai/CLIP>

## Mitsua

«This model is trained from scratch using only public domain/CC0 or copyright images with permission for use, with using a fixed pretrained text encoder (OpenCLIP ViT-H/14, MIT License).»

- <https://huggingface.co/Mitsua/mitsua-diffusion-one>

# A installer

## Stable Diffusion

- <https://stability.ai/blog/stable-diffusion-public-release>
- <https://github.com/CompVis/stable-diffusion>
- <https://huggingface.co/spaces/stabilityai/stable-diffusion>
- <https://www.reddit.com/r/StableDiffusion/>
- tuto <https://blog.paperspace.com/generating-images-with-stable-diffusion/>

## Stable Diffusion Web UI

- <https://github.com/AUTOMATIC1111/stable-diffusion-webui>
- tuto : [https://gigazine.net/gsc\\_news/en/20220912-automatic1111-stable-diffusion-webui-prompt-interrogate/](https://gigazine.net/gsc_news/en/20220912-automatic1111-stable-diffusion-webui-prompt-interrogate/)
- exemples : <https://github.com/AUTOMATIC1111/stable-diffusion-webui-feature-showcase>

## Art numérique et IA génératives

**Paragraphica** is a context-to-image camera that uses location data and artificial intelligence to visualize a “photo” of a specific place and moment. The camera exists both as a physical prototype and a virtual camera that you can try. ( <https://bjoernkarmann.dk/project/paragraphica> )

**90 miles** controverse sur l'utilisation de l'IA générative pour créer des images "documentaires" d'événements passés : <https://www.blind-magazine.com/fr/stories/comment-limagerie-ia-ebanle-le-photojournalisme/>

## Ressources diverses

- <https://learn.adafruit.com/generating-ai-art-with-vqgan-clip/piloting-the-weird>
- <https://github.com/CompVis/taming-transformers>
- <https://www.kaggle.com/basu369victor/playing-with-vqgan-clip>
- <https://www.aiweirdness.com/tag/clip-vqgan/>
- <https://pollinations.ai/c/Anything>
- <https://lvmiranda921.github.io/notebook/2021/08/11/vqgan-list/>
- <https://ml.berkeley.edu/blog/posts/clip-art/>
- <https://hotpot.ai/art-maker>
- <https://aiartists.org/ai-generated-art-tools>
- <https://www.artbreeder.com/>
- <https://www.aiweirdness.com/tag/clip-vqgan/>

Article extrait de : <http://www.lesporteslogiques.net/wiki/> - **WIKI Les Portes Logiques**

Adresse : [http://www.lesporteslogiques.net/wiki/ressource/code/images\\_et\\_ia/start](http://www.lesporteslogiques.net/wiki/ressource/code/images_et_ia/start)

Article mis à jour: **2024/01/31 16:33**