

**UNIVERSIDADE FEDERAL DE UBERLÂNDIA**

**GUILHERME BORGES PAGANO**

**Uma Aplicação de Aprendizagem de Máquina: Stable Diffusion**

**UBERLÂNDIA**

**2022**

## O que é o Stable Diffusion?

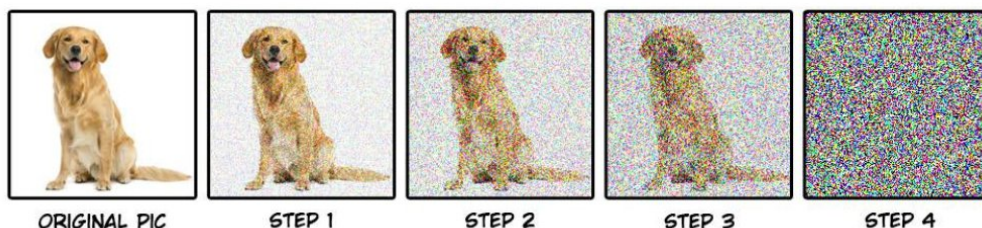
O Stable Diffusion é uma software de aprendizagem de máquina, com rede neurais que gera imagens através de inteligência artificial. Originalmente foi desenvolvido para gerar imagens a partir de um texto inserido pelo usuário, mas acabou sendo atualizado com novas funções relacionadas.

A aplicação possui código aberto e é mantida pela comunidade “*open source*”, atualmente estando na sua versão 2.0. Em muitos casos a imagem produzida pode ser confundida com uma imagem real, ou até mesmo com uma imagem feita por um artista renomado.

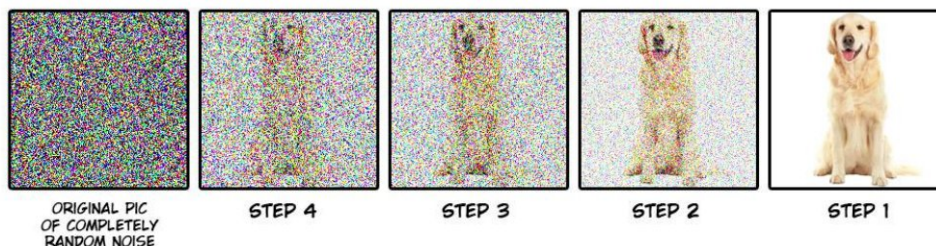
## E como o software gera imagens do “nada”?

O Stable Diffusion é um algoritmo de redução de ruído, o que significa que ele tenta remover o ruído das imagens.

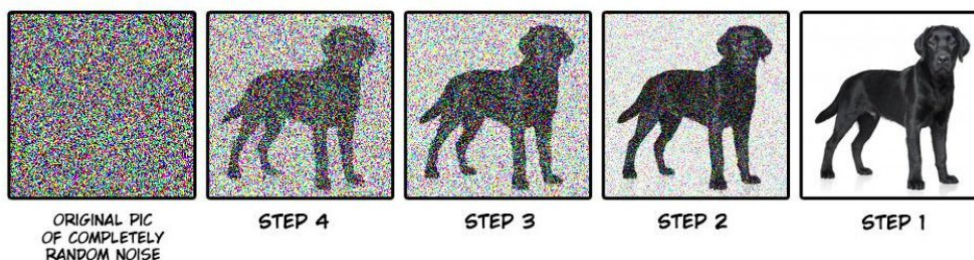
Para isso, de maneira simples, podemos dizer que primeiro pegamos uma foto de uma coisa, digamos, um cachorro, e “entregamos” para o computador: “ei computador, por favor, gradualmente transforme esta imagem em ruído, e memorize cada passo que realizar”.



Agora, pegamos uma foto de um ruído aleatório e “pedimos” pro computador: “efetue os passos anteriores, porém agora de forma reversa!”. Dessa forma, o algoritmo vai gerar uma imagem de um cachorro, não o mesmo cachorro, visto que começamos de um ruído aleatório.





A partir disso conseguimos ensinar para o computador outras palavras, como por exemplo a cor preta, e pedimos um cachorro preto, fazendo uma junção das duas palavras.



Basicamente, o algoritmo é calibrado mostrando imagens parciais cobertas por ruído artificial e vendo como ele adivinha qual é o ruído a ser removido. As calibrações dão um pequeno empurrão dependendo de quão errado é cada palpite, não o suficiente para fazer diferença em qualquer imagem. Eventualmente, surge uma solução geral para redução de ruído de imagem.

O algoritmo nunca salva as imagens de treinamento. O tamanho do arquivo permanece o mesmo, seja em 1 imagem ou em 1 milhão de imagens

Name	Size
 sd_v1-4.ckpt	4,165,411 KB
 sd_v1-5.ckpt	4,165,411 KB

As palavras são mapeadas para pesos exclusivos que são adicionados ao algoritmo de redução de ruído e, portanto, a calibração precisa funcionar em equilíbrio com o impacto das palavras. Por exemplo, “foto” ou “desenho” pesaria algumas escolhas de redução de ruído de maneira diferente.

Eventualmente, o Stable Diffusion fica tão bom que pode gerar uma imagem totalmente nova apenas com ruído, executando o algoritmo de denoiser várias vezes seguidas para continuar melhorando a imagem.



## Considerações

O Stable Diffusion é um ótimo exemplo de ferramenta inovadora criada com Aprendizagem de Máquina, e ainda é uma tecnologia que ainda tem muito pra evoluir. Exemplo disso é um artigo que li alguns dias atrás, realizado por dois biocientistas da Universidade de Osaka no Japão, em que foi possível utilizar algoritmos de difusão para reconstruir imagens a partir da atividade cerebral humana.

Além disso já é uma ferramenta muito útil nos dias atuais, podendo ser utilizada para auxiliar artistas gráficos, ou apenas entusiastas que desejam utilizar alguma imagem sem direito autoral ou fazer alguma montagem divertida.

## REFERÊNCIAS

**Reddit.** Disponível em:

[https://www.reddit.com/r/StableDiffusion/comments/zbi8zl/my\\_attempt\\_to\\_explain\\_how\\_stable\\_diffusion\\_works/](https://www.reddit.com/r/StableDiffusion/comments/zbi8zl/my_attempt_to_explain_how_stable_diffusion_works/). Acesso em 07 mar. 2023

**Reddit.** Disponível em:

[https://www.reddit.com/r/StableDiffusion/comments/zp2zq6/ive\\_made\\_an\\_eli5\\_in\\_case\\_youre\\_tired\\_of\\_arguing/](https://www.reddit.com/r/StableDiffusion/comments/zp2zq6/ive_made_an_eli5_in_case_youre_tired_of_arguing/). Acesso em 07 mar. 2023

**Diolinux – Stable Diffusion: Como a tecnologia está mudando a Arte.** Disponível em:

<https://diolinux.com.br/editorial/stable-diffusion-tecnologia-arte.html>. Acesso em 07 mar. 2023

**Medium – Stable Diffusion Clearly Explained!.** Disponível em:

<https://medium.com/@steinsfu/stable-diffusion-clearly-explained-ed008044e07e>. Acesso em 07 mar. 2023

**Medium – Stable Diffusion: From Description to Visualization.** Disponível em:

<https://medium.com/mlearning-ai/stable-diffusion-from-description-to-visualization-381113890b7>. Acesso em 07 mar. 2023

**High-resolution image reconstruction with latent diffusion models from human brain activity.** Disponível em:

<https://www.biorxiv.org/content/10.1101/2022.11.18.517004v2.full.pdf>. Acesso em 07 mar. 2023