

Difuzní modely (Diffusion Models)

Difuzní model je typ generativního modelu, který vytváří nová data (nejčastěji obrázky) pomocí procesu postupného odstraňování šumu. Princip je inspirován termodynamikou – konkrétně šířením (difuzí) plynu. Model se učí, jak z čistého obrazu udělat chaos a, co je nejdůležitější, jak tento proces obrátit.

1. Jak to funguje: Od chaosu k řádu

Proces funguje ve dvou hlavních fázích:

A. Dopředná difuze (Forward Diffusion)

Do čistého trénovacího obrázku se postupně v mnoha krocích přidává náhodný šum (Gaussovský šum), dokud se obrázek nezmění v naprostý chaos, kde není poznat původní obsah.

B. Zpětná difuze (Reverse Diffusion)

Toto je fáze, kde probíhá „kouzlo“. Neuronová síť (obvykle architektury **U-Net**) se učí předpovídat, kolik šumu bylo v daném kroku přidáno, a tento šum odečíst.

- Při generování pak model začne s **čistým náhodným šumem**.
- Krok za krokem šum odstraňuje a postupně „vyřezává“ z chaosu konkrétní tvary a detaily.

2. Text-to-Image: Jak AI rozumí zadání?

Aby model věděl, co má z šumu vytvořit (např. „kočku na měsíci“), používá se mechanismus **Conditioning** (podmiňování):

1. Textový prompt je převeden na číselný vektor pomocí modelu **CLIP** (od OpenAI).
2. Tento vektor je "vstříknut" do procesu zpětné difuze pomocí mechanismu `[[it_encyklopedie:attention_mechanism|Attention]]`.
3. Model pak při odstraňování šumu preferuje ty tvary, které statisticky odpovídají vašemu popisu.

3. Latentní difuze (Stable Diffusion)

Generování obrázků ve vysokém rozlišení je výpočetně extrémně drahé. Model **Stable Diffusion** přinesl revoluci tím, že difuzní proces neprovádí přímo na pixelech, ale v tzv. **latentním prostoru** (komprimované verzi obrázku vytvořené pomocí [autoencoderu](#)).

- Výpočty probíhají na malé „mapě“ (64 × 64 místo 512 × 512).
- Až finální výsledek je dekodérem zvětšen do plného rozlišení.
- To umožnilo spouštět tyto modely i na běžných domácích grafických kartách.

4. Srovnání s ostatními modely

Vlastnost	GAN	Difuzní modely
Kvalita obrazu	Velmi vysoká, ale někdy trpí artefakty.	Špičková, velmi detailní a realistická.
Diverzita	Nízká (často generuje podobné věci).	Vysoká (dokáže pokrýt celý rozsah dat).
Rychlost	Velmi rychlé (jeden průchod sítí).	Pomalejší (vyžaduje desítky kroků čištění).
Stabilita učení	Špatná (sítě se mohou zhroutit).	Výborná (trénování je velmi stabilní).

5. Využití v praxi

- **Generování obrazu:** Tvorba grafiky, konceptů pro hry a filmy.
- **Inpainting & Outpainting:** Oprava chybějících částí fotky nebo „dokreslení“ krajiny za okraj původního snímku.
- **Image-to-Image:** Převod skici na realistickou fotografii.
- **Video a Audio:** Nejnovější modely (např. Sora nebo Stable Audio) aplikují stejný princip difuze na časovou osu videa nebo zvukové vlny.

Zajímavost: Proces difuze je v podstatě jako pozorování mraků. Vidíte náhodné tvary, ale vaše mysl (v tomto případě neuronová síť podpořená textovým zadáním) v nich začne hledat a zvýrazňovat známé objekty, dokud se tam skutečně neobjeví.

[Zpět na AI rozcestník](#)

From:
<https://serviceit.cz/> - IT ENCYKLOPEDIE

Permanent link:
https://serviceit.cz/doku.php?id=difuzni_modely

Last update: **2025/12/31 14:29**

