

Free Midjourney



Előadás vázlata

Alap modell

Tanítóhalmazok és adatfeldolgozás

Háló optimalizálása

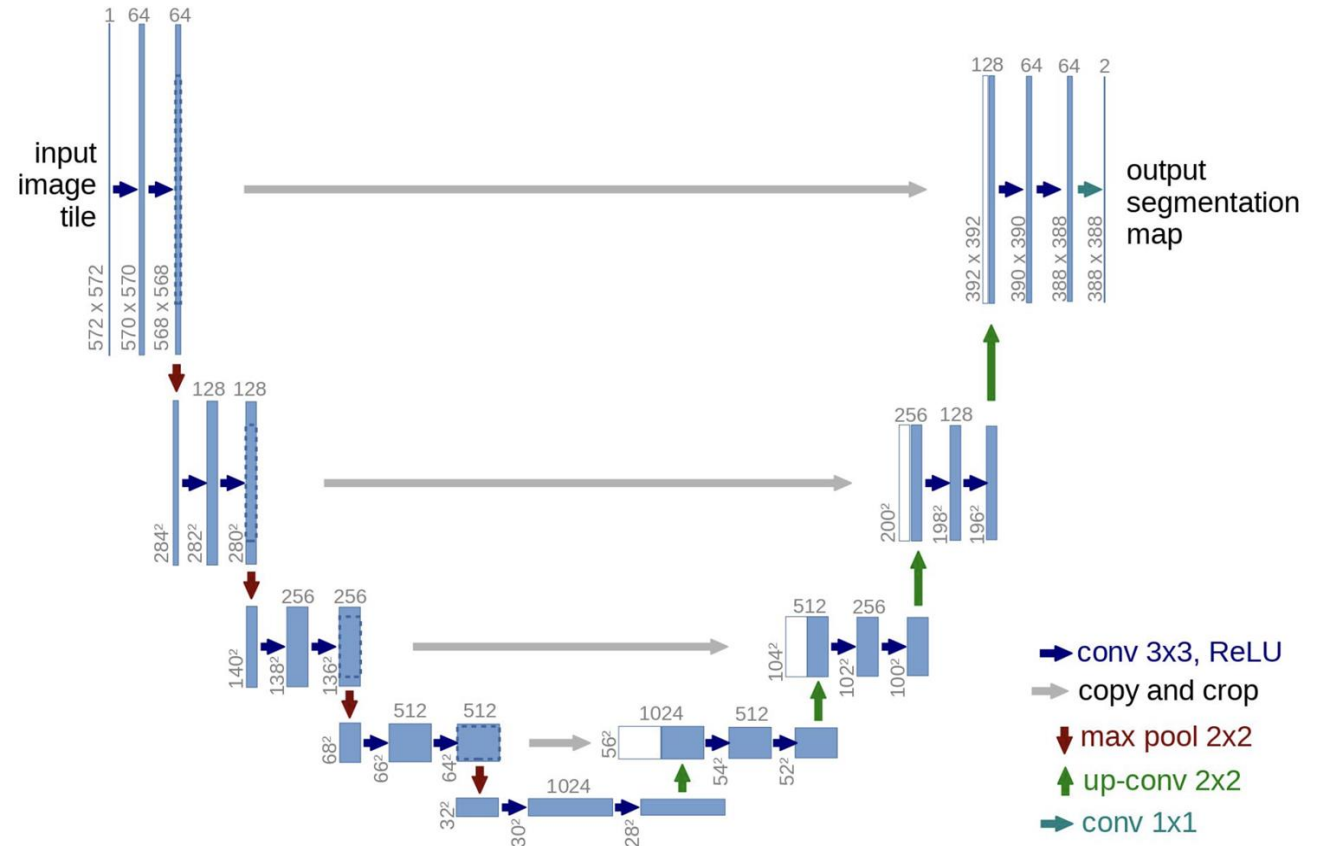
Kiértékelés és háló választása

Jelenlegi környezet és pipeline(demó)

Összefoglalás

U-Net

- Kódoló
 - Tulajdonságok összegyűjtése
- Dekódoló
 - Kép felépítése az kapott tulajdonságokból
- Kódoló rétegenként összekötve a Dekódolóval, hogy segítsen annak a rekonstruálásban.



Tanítóhalmazok és adatfeldolgozás

- DALL-E-Dogs, Flower102, CelebA
- A képek feldolgozására volt szükség
 - a háló tanulásának hatékonysága
 - a kutya dataset-et a flower dataset formájára kellett alakítani, hogy a pipeline ne akadjon meg
- a transform után reverse transform (képgenerálásánál)



```
transform = Compose([
    transforms.Resize(64),
    transforms.CenterCrop(64),
    transforms.RandomHorizontalFlip(),
    transforms.ToTensor(),
    transforms.Lambda(lambda t: (t * 2) - 1)
])
```



Tanítás és háló optimalizálása

- 100 epoch
- Legjobb modell kimentése
- Learning rate kitapasztalása (gradient spike)
- ADAM optimalizáló (Legjobban bevált diffúziós képgenerálásnál)
- U-Net rétegeinek a száma
- U-Net adott rétegeinek a nagysága

Kiértékelés és háló választása

- Több háló keletkezett az optimalizálás során
 - ezeknek a legfőbb különbsége a mélységük és a tanulandó feature-ok száma volt.
 - a tanítás után külön fájlokba mentettük a hálók súlyát
- A tanítóhalmaz és a generált képek közös méretre hozása után a KID és a FID score alapján választottuk ki a megfelelő hálót.
 - InceptionV3 által felismert feature-ok
 - Fréchet távolság
 - Kernel Inception Distance: maximum átlagos eltérés
 - Emberi szubjektivitás

	Model Name	FID Score	KID Score
0	samples_1_2_4_8	24.923375	0.334925
1	samples_1_2_3_6	27.126640	0.326628
2	samples_1_2_3	18.302253	0.396148
3	samples_1_2_4	24.489287	0.294032

Választott modell “szubjektív értékelése”

- KID és a FID számok beszédeseek, de az ember számára fontos, amit a saját szemével lát.
- BING chat kiértékelés
 - *“Az általad küldött kép alapján, az ábrázolt tárgy nagyon hasonlít egy virágra. A színek, textúrák és a közeli nézet mind hozzájárulnak ehhez a benyomáshoz.”*
 - *“Az általad küldött kép alapján, az ábrázolt tárgy egy rakás kő”*
- A rossz minőségű képek oka:
 - az erőforrás igény, google colab korlátai -> Komondor legközelebb
 - a képek feldolgozása szükséges volt, azonban problémákat okozott a sample-k generálásában.
 - crop -> lehet nem is a virág szerepelt a képen
 - resize -> minőség romlás

Jelenlegi környezet és pipeline

- Docker környezet korlátozásai: CPU
- Jupiter notebook
 - adatelőkészítés->tanítás->tesztelés
 - adatelőkészítés->tesztelés
- Hugging face: CPU, feltöltött súlyok
 - szám megadása
 - nehezítés: időkorlát

Demo: <https://huggingface.co/spaces/mgreg555/FreeMidjourney>



Összefoglalás

1. Docker környezet összerakása [Marci, Gergő]
2. Modell kiválasztása és tanítása [Marci]
3. Kiértékelő módszer választása (FID, KID) [Gergő, Dani]
4. Különböző modellek kiértékelése és kiválasztása [Gergő, Marci]
5. Tanítóhalmaz kezelése (DALL-E-Dogs, Flowers102, celebA) [Gergő, Marci]
6. Pipeline a docker környezetben [Marci, Gergő]
7. Demó összeállítása [Gergő]

Köszönjük a figyelmet!