

Raport Projektowy

1. Opis Koncepcji (1 pkt)

Posiadam w domu rzeźbę po Cioci o imieniu Licia chciałbym dodać ją do modelu "sd-legacy/stable-diffusion-v1-5".



2. Przygotowanie Danych (5 pkt)

{/* Wypełnij tę sekcję tylko, jeśli używasz własnego zbioru danych */}

Opis procesu przygotowania danych:

1. **Pozyskiwanie:** Zrobiłem zdjęcia moim telefonem, w różnych konfiguracjach, z różnych kierunków, w różnym oświetleniu i z różnym tłem.
 2. **Preprocessing:** Zdjęcia przetworzyłem do formatu 1024x1024 z użyciem webowego narzędzia <https://www.birme.net/>
 3. **Opis:** Każde zdjęcie opisałem podkręślając elementy które nie są samą licją aby nauczyć model odróżniać nowy koncept od tła.
-

3. Metody uczenia (9 pkt)

3.1. Textual Inversion

Ta technika pozostawia bardzo wiele do życzenia najlepszy obrazek jaki byłem w stanie z niej wyciągnąć to:



w dodatku czas uczenia był olbrzymi w porównaniu z DreamBooth i o drobne dłuższy niż LoRA

3.2. DreamBooth

Ta technika okazała się przynosić bardzo zadowalające rezultaty:



czas uczenia nie był bardzo długi a wygenerowany obrazek zawiera cechy charakterystyczne dla licji jak opaska we włosach czy szara toga

3.3. LoRA

To jedyna technika dla której zastosowałem opisy zdjęć i niestety wyniki nie są szczególnie dobre:



powyższe zdjęcie to najlepsze co udało mi się wygenerować

4. Eksperymenty (15 pkt)

Experyment 1:

Jako eksperiment postanowiłem spróbować techniki dreambooth lora, która obiecuje nawet bardziej przyspieszyć proces uczenia się nowego konceptu.

Rezultaty okazały się mieszane, owszem przy takiej samej liczbie kroków co w przypadku każdej innej techniki (uważałem 1000 kroków) była niesamowicie szybka i skończyła trening w około 10 min, jednak wygenerowane obrazki można tylko z dużym wysiłkiem zakwalifikować jako licie:



Experiment 2:

Ze względu na obiecującą prędkość uczenia z użyciem techniki dreambooth lora postanowiłem spróbować zastosować tą technikę z większym batch size: 2 (oryginalnie we wszystkich treningach używałem 1) oraz większą ilością kroków: 2000

Rezultaty okazały się przypominać licje bardziej, ale są dziwnie "halucynogenne" w dodatku wydaje się że model nauczył się elementów tła z mojego data setu pomimo wielu próbek które ich nie zawierają:





Notatki końcowe:

W tym repozytorium znajduje się większość mojej pracy w pliku [code.ipynb](#) jest cały kod którego użyłem. w katalogu [Licia](#) są moje dane treningowe. I są te wagi dla technik dla których umiałem je załączyć niestety niektóre były tak duże (dreambooth) ze postanowiłem ich nie załączać mam je zapisane na [dysku google](#) i można je z tamtąd pobrać.