

2023 HCI KOREA Conference

#taggenerator: NovelAI를 이용한 이미지 생성 서비스의 텍스트 프롬프트 탐색 및 태그 생성 행태 연구

#taggenerator: How an Online Community Designs an Interface for Text Prompt Inputs
Using a NovelAI's Diffusion Model

김선욱 d02reams@unist.ac.kr
표현적컴퓨팅 연구실, UNIST

권유상 yk7244@unist.ac.kr
표현적컴퓨팅 연구실, UNIST

이경호 kyungho@unist.ac.kr
조교수, 표현적컴퓨팅 연구실, UNIST

NovelAI 사용자의 프롬프트 탐색 행동 범주화와 확산모델 기반 AI 이미지 생성 서비스 디자인 개선

NovelAI의 등장으로 활성화된 AI 그림 커뮤니티를 분석하여, 원하는 이미지를 생성하기 위해 사용자가 텍스트 프롬프트를 탐색하는 과정을 파악하였음. 주제 분석으로 정의한 사용자 행동 패턴 범주를 통해 확산모델 기반 AI 이미지 생성 서비스의 디자인 개선방안을 제시하고자 함.

사용자가 어떻게 텍스트를 입력해야 원하는 이미지가 생성될까?

1.
Introduction

2.
Research

3.
Discussion

4.
Conclusion

Introduction

Background

Related Work

Problems

Background

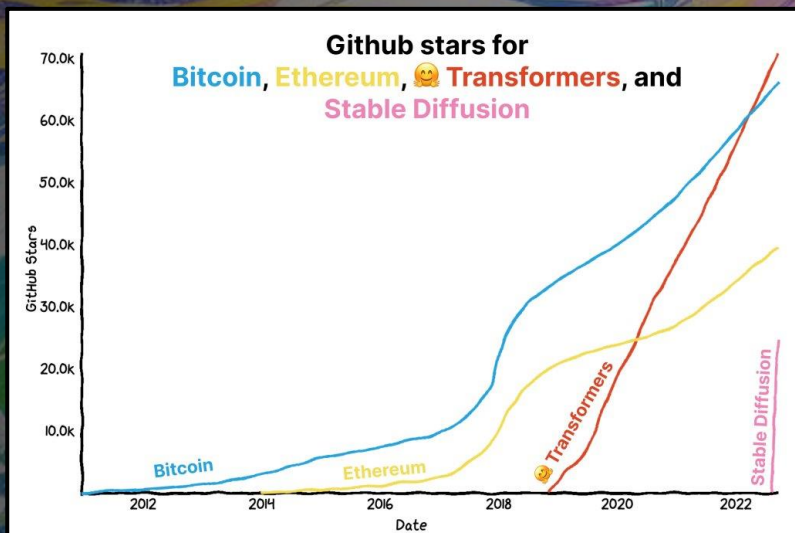
확산모델(diffusion model) 등장 후, 텍스트 프롬프트 입력으로 높은 수준의 이미지 생성이 가능해짐. 특히, Stable Diffusion 모델이 오픈소스로 공개되고, 이를 기반으로 새로운 데이터셋을 학습한 여러 모델이 나왔음. 이 중에서 높은 품질의 만화풍 이미지를 생성하는 NovelAI가 화제가 되면서 인터넷 커뮤니티에서 정보 공유와 논의가 활발하게 이루어지고 있음.

Prompt & Negative Prompt에
텍스트 태그를 나열하여 입력
(e.g. retro, sky blue theme, 1girl)

NovelAI
Diffusion

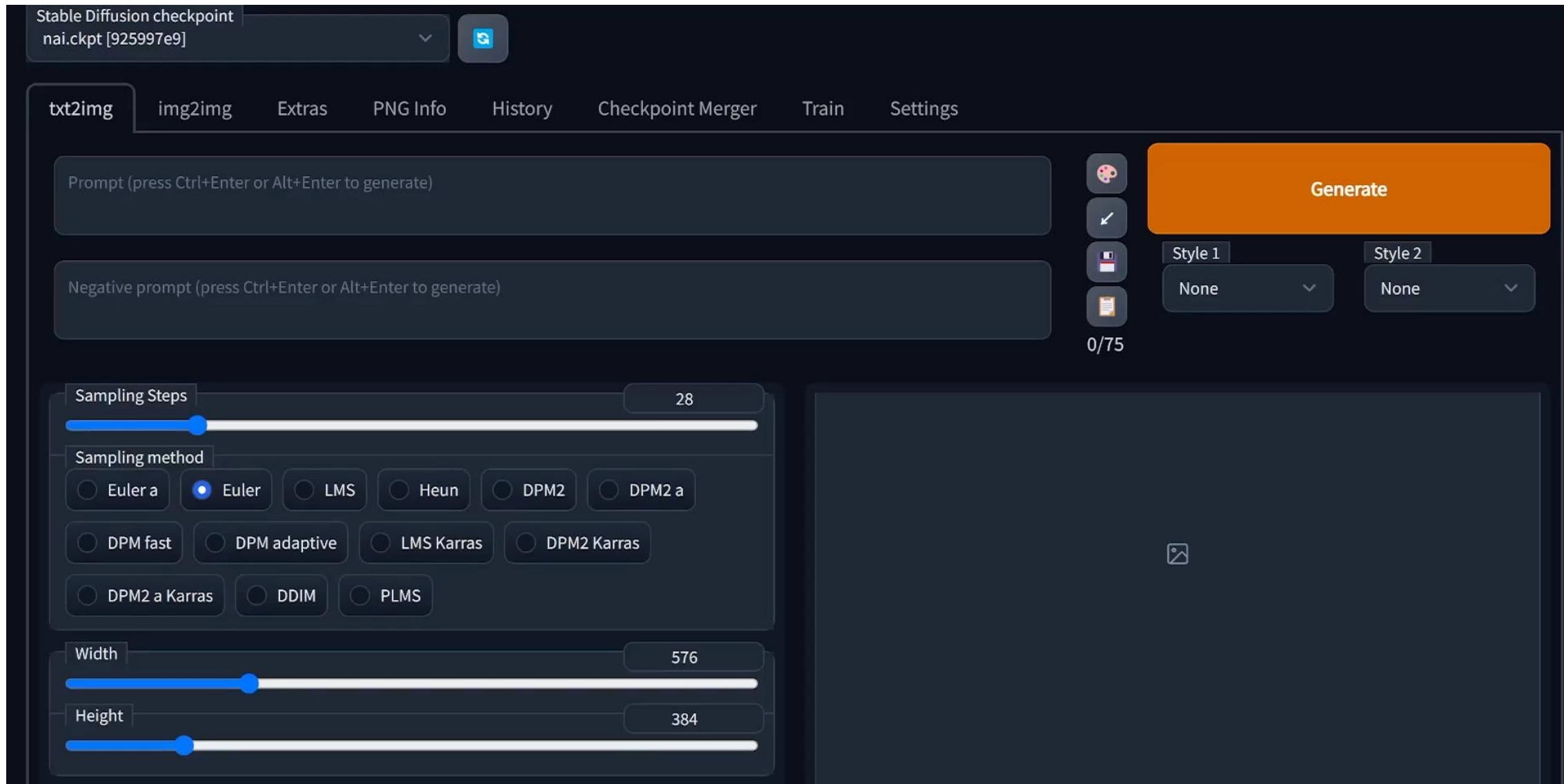
배경과 같은
만화풍 이미지 생성

기하급수적인 Github Stars 증가율을
보여주고 있는 Stable Diffusion



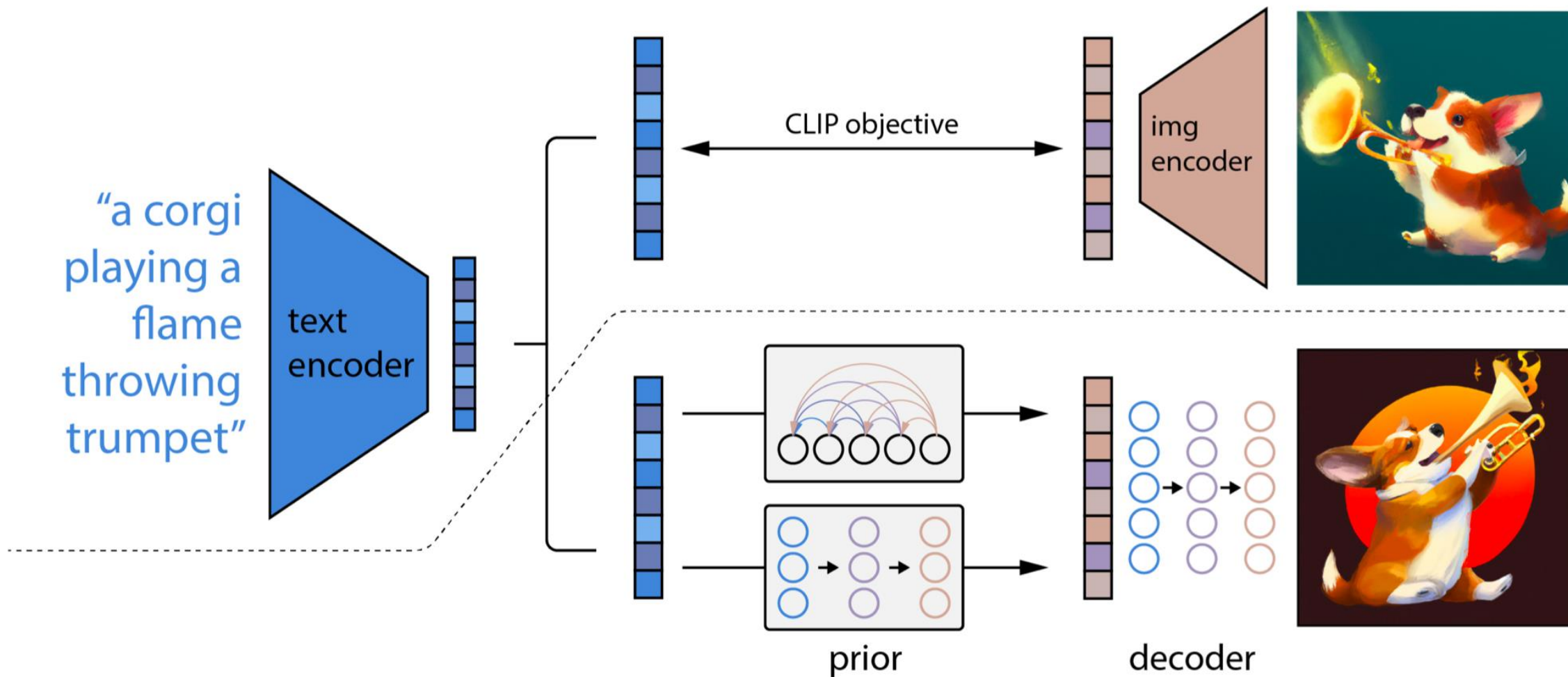
Background

아래는 NovelAI webui 구동 영상임. 텍스트 프롬프트에 따라 생성되는 그림이 달라짐을 확인할 수 있음.



Background

확산모델은 임의의 픽셀을 골라 왜곡된 이미지를 만들고, CLIP의 도움을 받으며 복원함. 복원하는 과정에서 의도적으로 텍스트를 단서처럼 입력해주면 영향을 받아 사용자의 의도가 반영된 완전히 새로운 이미지가 생성됨. 아래는 확산모델이 텍스트를 기반으로 이미지를 생성하는 과정에 대한 도식임.



Related Work

확산모델은 창작에 있어 좋은 재료가 될 가능성이 크지만, 보조하는 도구로 가공하기까지는 아직까지 어려움이 많음. 실제로 사용자는 원하는 이미지 생성을 위해 많은 시행착오를 겪어야 함. 확산모델과 HAI_{Human-AI Interaction} 기반 제품 및 서비스에 대한 논의는 아래와 같음.

Input

chicken and pizza are on the table

Output



1. 확산모델의 가능성, 그리고 확산모델의 성능 및 결과물의 품질 향상을 위한 방향성을 제시하였지만,

(Ho, et al., 2020), (Radford, et al., 2021), (Rombach, et al., 2022)

2. Human-AI Interaction 기반의 제품 및 서비스 과정에서 AI는 그 이용에 있어 완전히 새로운 패러다임 전환이 필요하고, 동작 방식에 대한 배경지식과 경험을 요구한다는 점에서 UX 디자인에서 사용하기 어려운 새로운 재료라고 주장함.

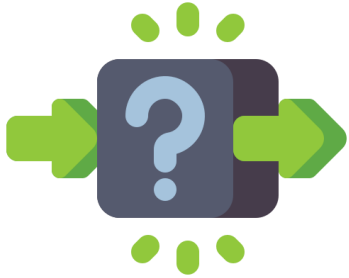
(Yang, et al., 2020)



실제로 많은 사용자가 본인의 의도와 바램을 담아 특정한 주제와 형식을 갖춘 이미지를 만드는데 어려움을 겪고 있음.

Problems & Research Questions

사용자의 의도대로 이미지를 생성하지 못하는 이유를 구체화하였으며, AI 이미지 생성 서비스 사용성을 개선하기 위해 아래와 같은 세 가지 연구질문을 선정하였음.



Problems

1. 입출력과 생성 과정에 중요한 역할을 하는 확산모델의 잠재 공간_{latent space} 속 정확한 동작기전에 대해 밝혀진 바가 크지 않음.
2. 기존의 이미지 생성 서비스와는 전혀 다른 언어 태그 형태의 텍스트 프롬프트를 입력으로 요구함.
3. 이미지 생성의 세부사항을 조절하는 파라미터_{parameter} 입력 방식 등이 직관적이지 않음.

Research Questions

1. AI 이미지 생성 서비스 이용 과정에서 텍스트 프롬프트를 도출하기 위해 사용자는 어떤 행동 패턴을 보이는가?
2. 각 행동 패턴은 어떤 관계가 있고 어떤 일련의 과정을 거치는가?
3. 특정 행동 패턴을 촉진하거나 불필요한 절차를 줄이기 위해 개발된 보조 서비스는 무엇이며, 어떤 개선된 서비스가 필요한가?

Research

Subjects

Thematic Analysis

Case Study

Subjects

아카라이브 AI 그림 채널은 2022년 9월 8일에 첫 게시글이 올라온 이후, 11만개가 넘는 게시글과 2만명이 넘는 구독자를 가진 대형 인터넷 커뮤니티임. 2022년 11월 14일을 기준으로 추천 10개 이상을 받은 게시글 1068개 연구대상으로 선정했고, 선정된 게시글을 읽고 만화풍 이미지 생성 AI 서비스 사용자에게 대한 사례연구를 진행하였음.

구독 채널 ▾ 주요 채널 ▾ 채널 검색

AI그림 채널
구독자 27451명 | 알림수신 474명 | @NO_NSFW
인공지능으로 제작한 그림을 자랑하고 정보를 공유하는 채널

채널위키 | 알림 | **구독 중**

전체글 | ★ **개념글** | 등록순 ▾ | 추천컷 ▾ | 글쓰기

패인알성 RPG가 모바일로 돌아왔다! 사전예약 개시

최근 방문 채널: 울로라이브 × 이터널 리전 × 리스트오리진 × 시반실사그림 × 종합속보 × 원신 × 백묘 × AI그림 학습 × 유머 × 텍스트게임 × 프롬프트 공유 ×

전체 | 일반 | **정보** | 질문 | 그림 | **그림** | 콘살(일반+) | **도커** | 대회 | 공지 | 운영

번호	제목	작성자	작성일	조회수	추천
공지	채널 알림 수신 기능 안내	★음★	2022.10.06	990075	
공지	아카라이브 모바일 앱 이용 안내 (iOS/Android)	★음★	2020.08.18	1530...	
공지	⚠️⚠️⚠️ AI그림 채널 공지 (2023-01-13) ⚠️⚠️⚠️	NO_NSFW	2022.10.23	46649	
공지	필독! AI그림 정보글 모음(★★★★질문 전 필독!★★★★)	노덴스	2022.11.29	84161	
공지	!!!! AI그림 채널 처음 들어왔으면 이것부터 보라(정보, 채널 규칙) !!!!	AIPs	2022.11.09	38554	
공지	(진행중) 대회 관련 공지 / 현재 진행중인 대회	NO_NSFW	2022.10.27	11092	

숨겨진 공지 펼치기(1개) 📄

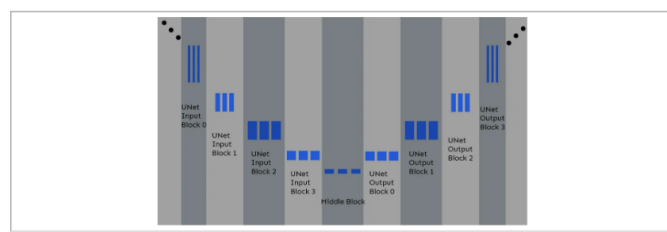
814	정보 ★ 초간단 손 타을 올리기 [28]	초광속입자	00:53	1123	37
813	정보 ★ U-net 블록 병합에 대해 대충 알아보자 (작성 중) [24]	애용이도독	23:17	698	29
812	정보 ★ 애니딴은 높은 확률로 병합 모델 + 우리의 프롬프트는 찜하고 있었다 [79]	애용이도독	19:53	1776	33
811	정보 ★ [작작] 학습용 태그 에디터 [4]	틀딱분쟁기	19:35	606	27
810	정보 ★ 작성중) AI 그림 채널 용어 정리 [16]	물큰	17:00	1665	45
809	정보 ★ 키타산 블록(말뭉) LoRA [9]	ㅇㅇ	2023.01.26	1378	20
808	정보 ★ 토라 두개 주워옴 [5]	바코드닉	2023.01.26	1239	14
807	정보 ★ 아, 너두 제로 베이스에서부터 학습시킬 수 있어! [10]	ㅇㅇ	2023.01.26	1005	15
806	정보 ★ InstructPix2Pix : 그냥 단순히 강력한 i2i 툴 [26]	qwer	2023.01.26	1516	33
805	정보 ★ 그래서 그 모델 어딴응? 종합모음집. [33]	Marionette	2023.01.25	2923	101

정보 ★ U-net 블록 병합에 대해 대충 알아보자 (작성 중)

애용이도독 | 추천 29 | 비추천 0 | 댓글 24 | 조회수 698 | 작성일 2023-01-26 23:17:29 | 수정일 2023-01-26 23:31:47
<https://arca.live/b/aiart/68377943>

모든 이미지 및 설명 출처:
<https://reentry.org/BlockMergeExplained>

이 글은 과거에 제가 썼던 **질문 비스무리한 정보 글**의 셀프 답변과도 같은 그 무엇이기도 함...



전부 읽기 싫다. 나는 끝만 읽고 싶다!

- 중앙(MID)에서 연 레이어는 질감, 선의 두께 같은 디테일을 표현한다 (화중, 그림제 등)
- 중앙과 가까운 레이어는 디테일한 특성을 표현한다 (옷의 종류, 얼굴 배치 및 간격 등)
- 어떤 레이어가 무슨 특성을 담당하는지는... 직접 노가다 뛰는 수 밖에 없다...

정보 ★ 화풍이 독특한 pastelmix

divdkn | 추천 37 | 비추천 0 | 댓글 10 | 조회수 2196 | 작성일 2023-01-25 05:31:13 | 수정일 2023-01-25 06:36:43
<https://arca.live/b/aiart/68251998>

<https://huggingface.co/andite/pastel-mix>

lora 섞어서 어떻게 잘 병합하든?
제작자가 올린 예시

masterpiece, best quality, upper body, 1girl, looking at viewer, red hair, medium hair, purple eyes, demon horns, black coat, indoors, dimly lit



Thematic Analysis

사례연구를 기반으로 원하는 이미지 생성을 위해 사용자가 입력할 프롬프트를 찾는 과정을 파악하고, 주제 분석^{thematic analysis}을 통해 사용자 행동 패턴의 범주^{category}와 유형^{theme}을 정의하였음.

게시글 내용 기반 범주 도출

내용 (contents)	범주 (categories)
프롬프트와 네거티브를 완벽히 동일하게 하면 무언가 그려지는 이미지가 나오고 거기에 Undesired Content가 선택되어 있을 시 예상외로 좋은 퀄리티의 그림이 등장한다	이중부정, 약화된 태그 부정으로 이미지를 생성하는 술어논리적 접근
즉 기본상태와 자기부정 상태의 중간모습은 네거티브의 가중치를 조절하는 것으로 원하는 만큼 조절할 수 있다는 점이다	
헤어 컬러색갈, 캐릭터의 나이 상상도 못한 것들이 세세하게 조절 가능할 것으로 보인다.	
비슷한 맥락에서 얼굴 쪽에 좀 더 초점을 두고 싶다! 할 때도 focus on face를 쓰기에 앞서 halo, earring 등 얼굴 근처에 포인트를 주는 약세를 추가해보는 것도 나쁘지 않음. 특히 halo의 경우에는 자체적으로 광원이 되기 때문에 얼굴 조명 효과도 좋아서 가장 애용하는 프롬 중에 하나.	의도한 요소가 나오게 하기 위해 비슷한 위치에 있거나 특징을 공유하는 태그로 대체적 접근
관심부위를 직접 명시하기보다는 그 부위가 추론되기 쉬운 동작, 시점, 상황을 명시하고 시행횟수를 늘리는 게 좋아 보임	
프롬프트 탭엔 텍스트, 숫자 외에 이모지도 넣을 수 있고 실험 결과 꽤나 유의미한 차이가 나타남. 더욱이 'smiling_face' 같은 단어조합을 '😊' '한칸만'으로 요약이 가능하니 프롬프트를 짧게 가져갈 수 있다는 장점도 있음.	이모지, 이모티콘 등을 태그로 활용하는 비언어적 접근

범주 기반 주제 분석Thematic Analysis

유형 (themes)		범주 (categories)
소비	의미적 접근	다른 사용자의 프롬프트를 모방
		다른 사용자에게 질문
		커뮤니티 내부 정보 검색
탐색	의미적 접근	정식 명칭 탐색 등 의미적 접근
		태그 간의 공통점과 특징을 바탕으로 태그 간의 궁합 판단
		의도한 요소가 나오도록 비슷한 위치나 특징을 공유하는 태그로 대체적 접근
	수치적 접근	세분화한 요소를 태그로 사용해 예측 가능한 범위의 이미지 생성
		이중부정, 약화된 태그 부정으로 이미지를 생성하는 술어논리적 접근
		반복 입력을 통한 태그 강화
생산		가중치 설정 기반 태그 강화 및 약화
		다른 사용자와 프롬프트를 공유
		보편적 품질 향상을 위한 태그 제시
정리		커뮤니티 외부 정보 검색 및 공급
		프롬프트 관련 정보 정리
		그림의 주요 요소를 분류하여 프롬프트 내의 태그 형식과 순서 정형화
역설계		임베딩을 통한 태그 정의
		모델 재학습, 서비스 개발 등 프롬프트 외적 개척 활동으로 프롬프트 탐색을 위한 인사이트 획득 및 제공
		AI 작동 과정을 인간이 문장을 처리하는 사고 과정에 대입하여 접근
가정		AI 모델 특징 탐구 및 학습 데이터 등 기반 프롬프트 제시
		커뮤니티에서 제기된 가설을 검증
		여러 태그를 조합한 프롬프트를 생성된 이미지를 보고 경험적으로 접근하여 태그 간의 궁합 판단
실험		대조실험을 통한 프롬프트 선별
		시행착오를 통한 휴리스틱 탐색
		그림의 주요 요소 구조를 재구성
혁신		이모지, 이모티콘, 유니코드 문자 등을 태그로 활용하는 비언어적 접근

Case Study

인터넷 커뮤니티 사용자들 사이에서 어떤 용어가 통용되고 있으며, 일부 사용자가 편의를 위해 자발적으로 만든 보조 서비스는 무엇이 있는지 정리함. 도출된 유형과 각 서비스의 장단점을 바탕으로 추가 분석하여 사용자가 원하는 그림을 생성하는 경험을 극대화할 수 있는 방안을 논의.

통용되는 Stable Diffusion 용어

용어	설명
Model	Stable Diffusion 기반 AI 모델(.ckpt)
Hypernetwork	[그림 2]에서 U-Net의 각 Step에서 Query, Key, Value 값 중 텍스트와 연관된 Key, Value 가중치(.pt)
Prompt	유저가 AI에 입력하는 텍스트 전체. Main prompt에 입력된 텍스트 정보를 기반으로, Negative prompt에 입력된 텍스트 정보를 배제하는 방향으로 이미지가 생성
Tag	Prompt에 입력한 각각의 단어 하나. 쉼표(.)로 구분
Embedding	사용자가 지정하는 특정 태그에 이미지 모음의 특징을 학습. AI 모델에서 특정 태그와 새로운 개념을 매칭시켜 특징에 맞는 그림 생성이 가능케 함
Sampling steps	이미지를 생성하는 과정의 수. 보통 step이 높을수록 그림 품질이 높아지지만 시간이 많이 걸리고, 일정 수치 이상일 때는 품질이 오히려 떨어짐
Sampling method	이미지를 생성하는 알고리즘. 주로 Euler, Euler a, DDIM를 사용
Width/Height	출력 이미지 크기(폭/높이)
CFG Scale	사용자가 넣은 프롬프트 입력을 얼마나 따를지 판단하는 척도. 높을수록 프롬프트 입력을 지키면서 그림을 생성하고, 낮을수록 AI가 해석의 자유를 가지고 그림을 생성
Seed	최초의 노이즈 이미지 생성에 사용되는 값(-1일 경우 난수 자동 입력)
Batch count*	작업할 배치(batch)의 수
Batch size*	단일 배치 내 동시 작업할 이미지 수

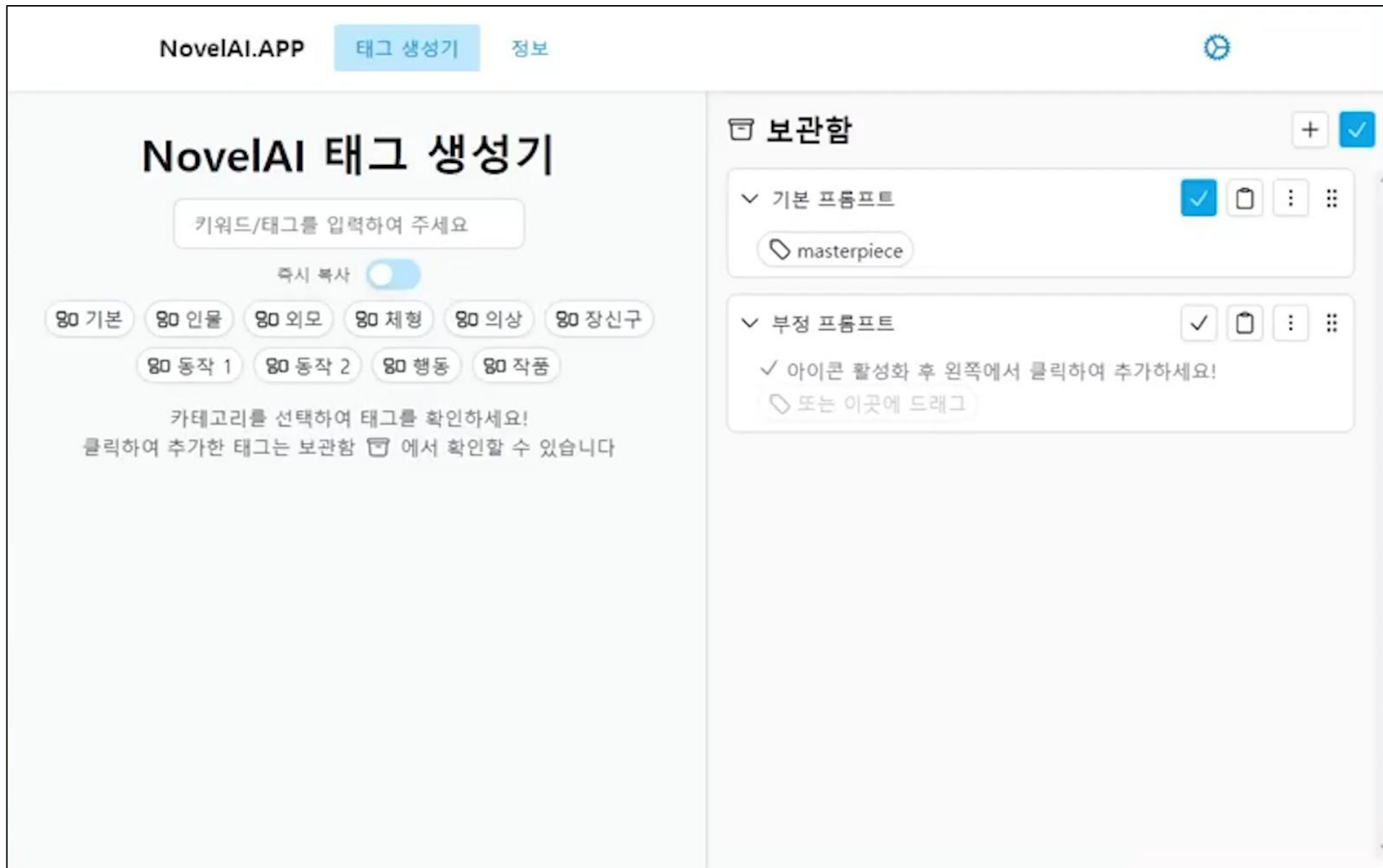
* 출력 이미지 수 = Batch count × Batch size

제작된 AI 보조 서비스

서비스	설명
Snowshell	이미지의 화질을 높여주는 업스케일러
Illuminarty	이미지가 AI가 그린 그림인지 판단
AI 이미지 EXIF 뷰어	이미지 EXIF 정보 기반 프롬프트 추출 및 딥러닝 기반 제안
NovelAI 태그 생성기	프롬프트에 활용할 태그 분류, 프롬프트 저장, 복제 보조
WebUI 태그 자동완성	태그 자동완성 및 Danbooru 태그 수를 기준으로 정렬하여 추천
AIBooru	AI 생성 그림 및 프롬프트 정보 공유
Prompt Search	AI 생성 그림 및 프롬프트 정보 공유, 추천 기준별 정렬, EXIF 뷰어, 입력한 프롬프트 기준으로 공유된 그림 중 가장 유사한 그림 제시
NovelAI Tag Experiments	태그 분류 및 세부 태그 정리, 이미지 및 프롬프트 공유
Quick NAI	Web UI 문법 지원, 맞춤형 태그 추천 시스템, 즐겨찾기
태그 가중치 변환기	Web UI 및 NovelAI Image Generator 문법 상호 변환
Anime-Face-Detector	만화풍 이미지에서 캐릭터 얼굴 부분만 자동 인식 및 수집 보조

Case Study

아래는 NovelAI 태그 생성기 사용 영상으로, 프롬프트를 쉽게 구성하기 위해 커뮤니티 사용자가 자발적으로 제작한 서비스 중 하나임.



Case Study

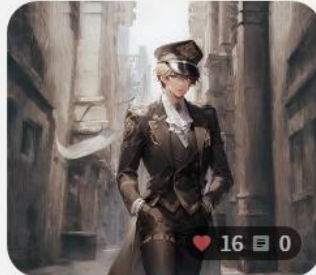
아래는 인터넷 커뮤니티 사용자가 자발적으로 제작한 AI 보조 서비스 Prompt Search임. AI로 생성된 이미지 뿐 아니라 재현할 수 있도록 각종 파라미터를 공유하고, 입력한 프롬프트 기준으로 공유된 그림 중 가장 유사한 그림 제시하는 등의 서비스를 제공함.

Prompt Search

Generate **ALPHA** Board Exif Viewer
Weekly Best Random Prompts Login SignUp

2022-10-17 2022-10-24 2022-10-31 2022-11-07 2022-11-14 2022-11-21 2022-11-28 2022-12-05
2022-12-12 2022-12-19 2022-12-26 2023-01-02 2023-01-09 2023-01-16 2023-01-23 2023-01-30

Weekly Best Articles



parameters	masterpiece, best quality, illustration, fantasy, dark basement indoors, bookshelf, potion, gem, bottle, animal skull, plant, candle, yuri, 2girls, alchemist, skinny, grey hair, yellow eyes, hooded cloak, robe NAI WEBUI Original
negative_prompt	(worst quality, low quality, normal quality, low resolution, lowres:1.4), 3d, blurry, text, artist name, signature, multiple views, pointy ears NAI WEBUI Original
steps	28
sampler	DPM++ SDE Karras
cfg_scale	8
seed	1393659416
size	1536x1024

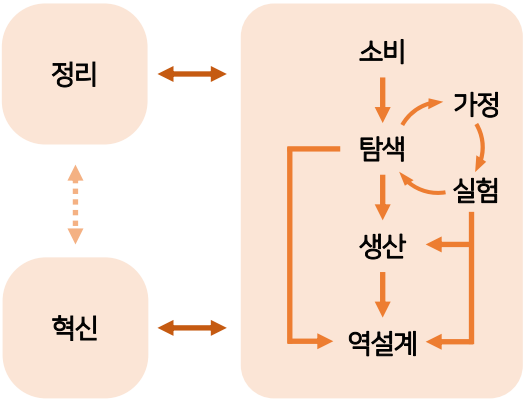
Discussion

Results

Suggestions

Results

아래는 2022년 11월 30일을 기준으로 추천 10개 이상을 받은 게시글을 1321개에서 해당되는 유형에 대한 통계와 각 보조 서비스는 어떤 유형의 행동 패턴 유형을 보조하기 위한 것인지에 대한 분석임. 행동 패턴 유형은 일련의 과정을 거치며, 혁신과 역설계를 보조하는 서비스가 부족하고, 가정과 실험, 정리, 수치적 접근 기반 탐색을 보조하는 서비스 역시 보완될 필요가 있음.



행동 패턴 유형 간의 관계 도식

	소비	탐색		생산	정리	역설계	가정	실험	혁신
		의미	수치						
Snowshell				●					
Illuminarty	●					●	●	●	
AI 이미지 EXIF 뷰어	●	●	●	●				●	
NovelAI 태그 생성기	●	●		●	●				
WebUI 태그 자동완성	●	●		●			●		
AIBooru	●	●	●	●	●		●	●	
Prompt Search	●	●	●	●	●		●	●	
NovelAI Tag Experiments	●	●	●	●	●		●		
Quick NAI		●	●	●	●		●	●	
태그 가중치 변환기		●	●	●					
Anime-Face-Detector						●			
게시글 _{1321개} 별 해당 유형 통계	856	473	107	892	160	266	22	137	60

Suggestions

사례연구를 통해 정리한 내용 외에도 이미지와 태그를 매칭한 후 학습 데이터용으로 정규화된 이미지를 다운로드 받을 수 있게 하는 서비스, 원활한 학습을 위한 서버 공유 서비스, 리터칭 서비스 등에 대한 수요가 있음을 확인함. 보완의 효율성이 클 역설계, 실험, 탐색 보조 서비스를 개선하기 위한 방안은 아래와 같음.

모델·하이퍼네트워크
학습을 위한 데이터
관리 서비스

+

시행착오를 줄여주고
공동경험을 반영한
직관적인 UI

⇒

역설계 난이도를
낮춰 모델·서비스
개선 접근성 향상

실시간 이미지
생성 과정 시각화

+

Challenge
Competition
서비스 플랫폼

⇒

가설 제시·검증
실험 활성화
초심자 학습자료

변화무쌍한
수많은 파라미터의
입력·관리가
필요한 UI

⇒

이산적으로
지정된 파라미터만
입력 가능한 UI

⇒

선택과 집중으로
프롬프트 탐색
효율성 증진

Conclusion

Limitations & Future works

Limitations & Future works

본 연구에서는 게시글 주제 분석을 통해 AI로 만화풍 이미지 생성을 위해 프롬프트를 탐색하는 사용자의 행동 패턴 유형 8가지를 도출하였고, 개선되어야 할 유형을 확인하였음. 이 과정을 통해 분석된 한계와 제언은 다음과 같음.

연구대상을 아카라이브
AI 그림 채널로 한정

연구대상을 국내외 다양한 커뮤니티로
확대한다면 더 높은 신뢰성과
시사점을 확보할 수 있을 것

게시글 필터링 기준으로
추천수의 적절성

시간에 따른 커뮤니티의 활성화 정도와
조회수 데이터를 함께 필터링에 활용하면
더 효율적인 연구가 가능해질 수도 있음

양적 연구를 동반한 인터뷰
등 사용자 조사 필요성

구체적인 불편사항과 피드백을 분석하면
AI 이미지 생성 서비스의 디자인을
사용자 친화적으로 개선 가능할 것

개선된 디자인·서비스 구현,
실험으로 사용성 평가 필요

기존 서비스 디자인을 개선하여 구현하거나
보조 서비스를 개발하여 실험하는 방식으로
실제 개선 여부에 대한 검증이 필요함

감사합니다.

**expressive
computing
lab.**

expc

2023 HCI KOREA Conference

#taggenerator: NovelAI를 이용한 이미지 생성 서비스의 텍스트 프롬프트 탐색 및 태그 생성 행태 연구

#taggenerator: How an Online Community Designs an Interface for Text Prompt Inputs Using a NovelAI's Diffusion Model

김선욱 d02reams@unist.ac.kr
표현적컴퓨팅 연구실, UNIST

권유상 yk7244@unist.ac.kr
표현적컴퓨팅 연구실, UNIST

이경호 kyungho@unist.ac.kr
조교수, 표현적컴퓨팅 연구실, UNIST